Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Song, YQ, Karasugi T, Cheung KMC et al. A functional variant in CHST3 associated with the susceptibility for lumbar disc degeneration

Contents

Supplementary Methods:

Subjects and assessment

For the linkage analysis of families with early-onset LDD, two sets of families were recruited independently, comprising 10 families (n = 89) and 8 families (n = 37). All subjects are of Southern Chinese origin and each family was recruited on the basis of having at least 1 early-onset LDD case (age ≤ 30). For all subjects, genomic DNA was isolated from the leukocytes extracted from blood.

For the GWAS, we recruited 366 cases (J1) for stage 1, and 544 cases (J2) for stage 2. The control groups consisted of 3,331 (J1) and 15,800 (J2) individuals registered in the Leading Project for Personalized Medicine in the Ministry of Education, Culture, Sports, Science and Technology, Japan as the subjects with diseases unrelated to LDH and the volunteers in the Osaka-Midosuji Rotary Club, Osaka, Japan [1], or from the general population[2]. For stage 3, we recruited Japanese (J3) and Northern Chinese (NC) (Table 1) LDH case-control groups. The Japanese case-control group consisted of 242 cases and 622 healthy controls. The controls were unrelated volunteers. We have used a subset of the cases in J1, J2 and J3 in previous publications, as well as a subset of the controls in J3, J1'and J2' in the same publications [2, 3].

The Northern Chinese case-control group (NC) consisted of 572 cases and 776 controls. Patients were recruited from 2008 to July 2010 in the Department of Orthopedic Surgery at the Beijing JiShuiTan Hospital. All patients underwent magnetic resonance imaging (MRI). Sagittal MRI (GE 1.5T, Milwaukee) was performed with a slice thickness of 4 mm. A T2-weighted image with a repetition of 2,500 ms and an echo time of 90 ms of the IVD was taken. More than 95% patients have positive MRI finding for LDH. All patients had discogenic low-back pain and/or unilateral pain radiating from the back along the femoral or sciatic nerve to the corresponding dermatome of the nerve root with duration of at least 2 months. Primary exclusion criteria included synovial cyst, spondylolisthesis, spinal tumor, spondylosis, trauma and inflammatory disease. All patients underwent surgery to palliate the symptoms. Individual who had known environmental risk factors, including heavy physical loading, occupational driving or cigarette smoking, were excluded. The control group consisted of 776 unrelated individuals from Nanjing.

For stage 4, we recruited a Finnish LDH case-control group, a Finnish population-based sample (NFBC 66) and a Chinese cohort (SC-1S) of LDD patients. The Finnish case-control group consisted of 281 cases and 393 controls [4]. All patients presented to the Oulu University Hospital due to sciatica symptoms. Inclusion criteria of the patient group were unilateral pain radiating from the back to below the knee that had been published

elsewhere[4]. All patients were confirmed using MRI to have LDH concordant with the symptoms. All individuals are from the same region of Finland. The methodology to assess LDD with sciatica in the NFBC 66 has been published previously [38]. The Southern Chinese cohort with severe phenotype (SC-S) is a subset of the SC cohort, an extension of the cohort from a previous publication (SC-1) [4]. LDD classification is based on an age-adjusted LDD score assessed by MRI, with severity of LDD graded with the Schneiderman's classification [5]. Following normalization for age, this cohort was divided into top 20% and bottom 20% groups according to the adjusted LDD score. As a result, the cohort consisted of 270 cases and 271 controls.

Genome-wide linkage analysis

In the first stage, the PRISM® human linkage mapping set v2.5-MD10 (Applied Biosystems) was used, with 400 microsatellite markers spanning the genome with an average resolution of 10 cM. In the second stage, 19 microsatellite markers were added within 5 cM interval of 5 selected regions from the initial genome-wide scan (Supplementary Table S1). Linkage map was based on the sex-averaged Marshfield map (Center for Medical Genetics website) and physical positions were based on NCBI Human Genome Build 36.1 (National Center for Biotechnology Information website). PCR were performed with a MJ Research PTC-225 thermal cycler, and pooled PCR products analyzed using an ABI 3730xl DNA analyzer.

Genotype data collection and allelic assignment were performed by using GeneMapper v3.7 (Applied Biosystems). Any erroneous genotypes were removed following assessment for Mendelian inheritance errors with Pedstats [6] and Pedcheck [7].

SNP Genotyping

For GWAS, the J1 cohort was genotyped using an Illumina HumanHap550v3 Genotyping BeadChip in stage 1. After excluding three cases with call rate of <0.98, we applied SNP quality control (call rate of <0.99 in both cases and controls and *P* value of Hardy-Weinberg equilibrium test of $>1.0\times10^{-5}$ in controls). 464,775 SNPs on autosomal chromosomes passed the QC filters. In stage 2, 1,500 selected SNPs were genotyped in J2 using the Affymetrix GeneChip Custom 10K array or a multiplex PCR-based Invader assay (Third Wave Technologies), or direct sequencing of PCR products using ABI 3700 DNA analyzers (Applied Biosystems) according to manufacturers' protocols. The 15,800 control samples were genotyped using an Illumina HumanHap550vs Genotyping BeadChip. After excluding one cases with call rate of <0.98, we applied SNP quality control (call rate of <0.99 in both cases and controls and *P* value of Hardy-Weinberg equilibrium test of $>1.0\times10^{-5}$ in controls). 1,349 SNPs on autosomal chromosomes passed the QC filters. In stages 3-4, SNPs were genotyped using the multiplex PCR-based invader assay or direct sequencing of PCR products using ABI 3700 DNA analyzers (Applied Biosystems), SNaPshot Multiplex System (Applied Biosystems),

or the Sequenom platform (Sequenom) according to manufacturers' protocols.

Collection of surgically retrieved intervertebral disc (IVD) specimen

All disc tissues were taken postoperatively with informed consent. Tissues from surgically obtained IVD were separated into NP, AF and EP portions, frozen in liquid nitrogen immediately and stored at -80°C until further processing.

Quantitative RT-PCR (Q-PCR) and analyses of CHST3 and miR-513a expression

Total RNA was extracted from IVD tissues using the Trizol reagent (Invitrogen). First-strand cDNA was synthesized using SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen). Amplications were carried out using the primers specified for *CHST3*, *COL2A1* and GAPDH as indicated in Supplementary Table S4, monitored using SYBR Green fluorescence (SYBR Green Master mix, Applied Biosystems). Expression of *CHST3* and *COL2A1* were normalized against *GAPDH* by the comparative threshold cycle method (ΔCt (*CHST3* or *COL2A1*– *GAPDH*)). Statistical analysis for a difference in ΔCt was compared by Mann-Whitney U tests.

Pyrosequencing was used to determine the relative levels of the *CHST3* rs4148941 allelic mRNA products in disc tissues obtained from individuals with a heterozygous genotype. A fragment flanking rs4148941was amplified from the cDNA using primers

CHST3-3 and biotin-labeled CHST3-4 (Supplementary Table S4). The amplicons were purified using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and subject to pyrosequencing using primer CHST3-Pyro (Supplementary Table S4). The data were analyzed in PSQ software, and difference in the percentage composition of the two alleles was compared using Wilcoxon signed rank test.

Expression of miR-513a-5p was detected using TaqMan microRNA assays as per manufacturer's primer design and protocol (Applied Biosystems). In brief, total RNA extracted from IVD tissues was reverse-transcribed separately using the primers of hsa-miR-513-5p and small nuclear RNA RNU6B as endogenous control. The RT products were subjected to quantitative PCR. Relative expression of miR-513a-5p to RNU6B was determined by comparative threshold cycle method (ΔCt (miR-513a-5p –RNU6B)) and the difference was compared by Mann-Whitney U test.

miRNA binding prediction of rs4148941 and rs4148949

PolymiRTS database [8] was used to look for predicted miRNA binding sites where rs4148941 and rs4148949 were located within the 3' UTR of *CHST3*. RNAhybrid [9] was used to analyze the minimum free energy hybridisation of binding miRNA candidates from PolymiRTS search. Candidate with the lowest minimum free energy was used for further analyses.

Luciferase constructs for the 3'UTR of the human CHST3

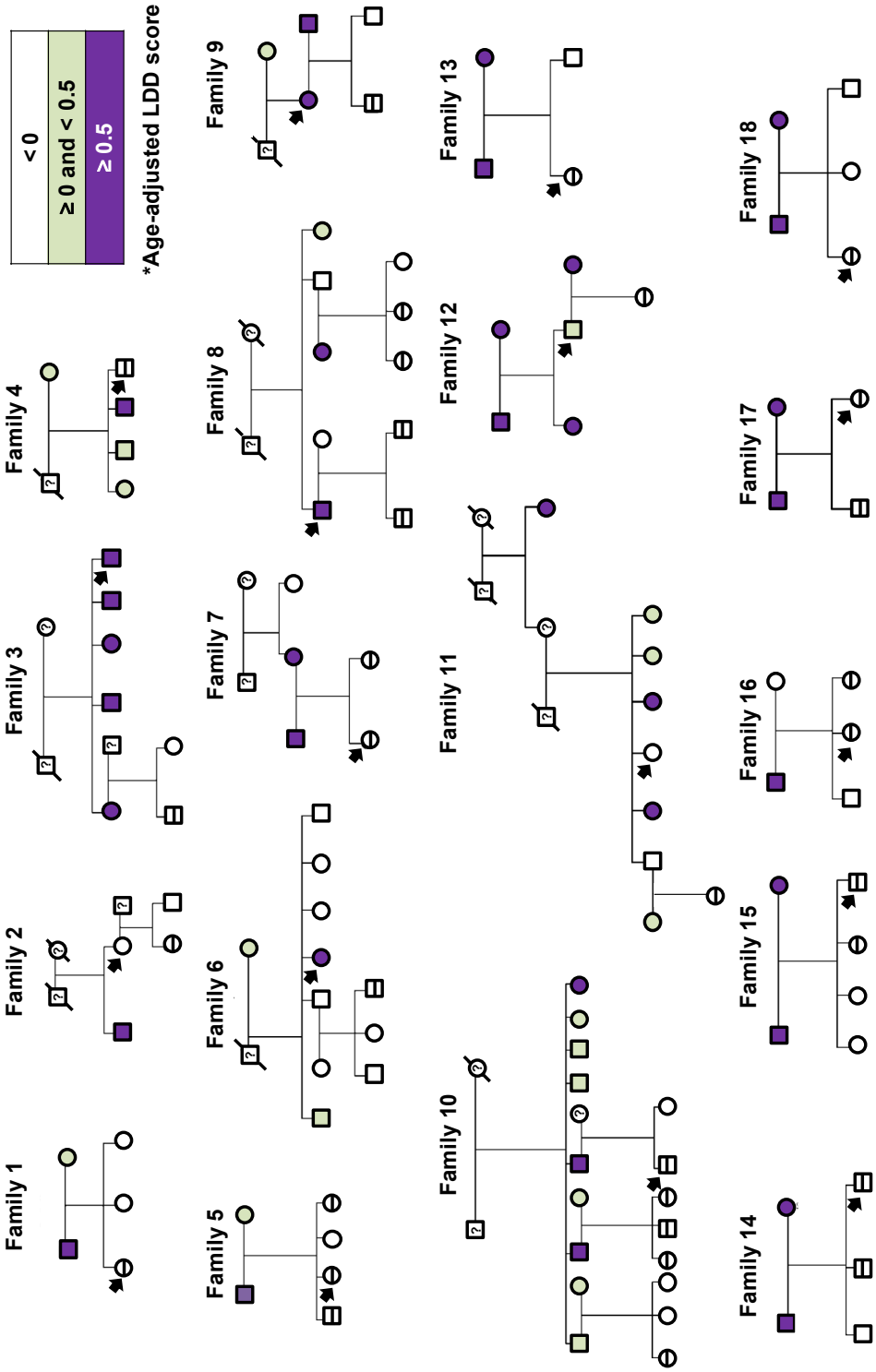
Primers for both allelic strands at the miR-513 (CHST3-rs4148941A/C) and miR-626 (CHST3-rs4148949C/T) target sites were synthesized (Supplementary Table S4), annealed and cloned into the Xba-I site downstream of the firefly luciferase gene in a pmirGLO construct (Promega). Human *CHST3* 3'UTR bearing the polymorphic alleles at rs4148941 (A/C) and rs4148949 (C/T) were generated using the same approach. Detection and quantification of hsa-miR-513 expression in IVD tissue was performed using the TaqMan method.

Cell transfection and luciferase assay for miRNA binding

The C28I2 chondrocyte line (gift from Dr. Mary Goldring; Weill Cornell Medical College, New York, USA) was maintained in Dulbecco's modified Eagle's medium (DMEM) (GIBCO, Invitrogen) supplemented with 10% fetal bovine serum (FBS) in a 5% CO_2 incubator at 37°C. C28I2 cells at a density of 1x10^{4} cells were seeded into 24-well plate and incubated for 24 h before transfection. Cells were cotransfected with 0.2 μg reporter constructs, with or without 5 nM of hsa-miR-513 or hsa-miR-626 (Ambion), using Lipofectamin 2000 (Invitrogen) according to the manufacturer's protocol. Forty-eight hours after transfection, cells were lysed and Renilla luciferase activities were determined using the Dual Luciferase Assay reagents (Promega) and a Sirius luminometer (Berthold Technologies Gmb H & Co.KG). The relative reporter activity was normalized to the Renilla luciferase activity. Three

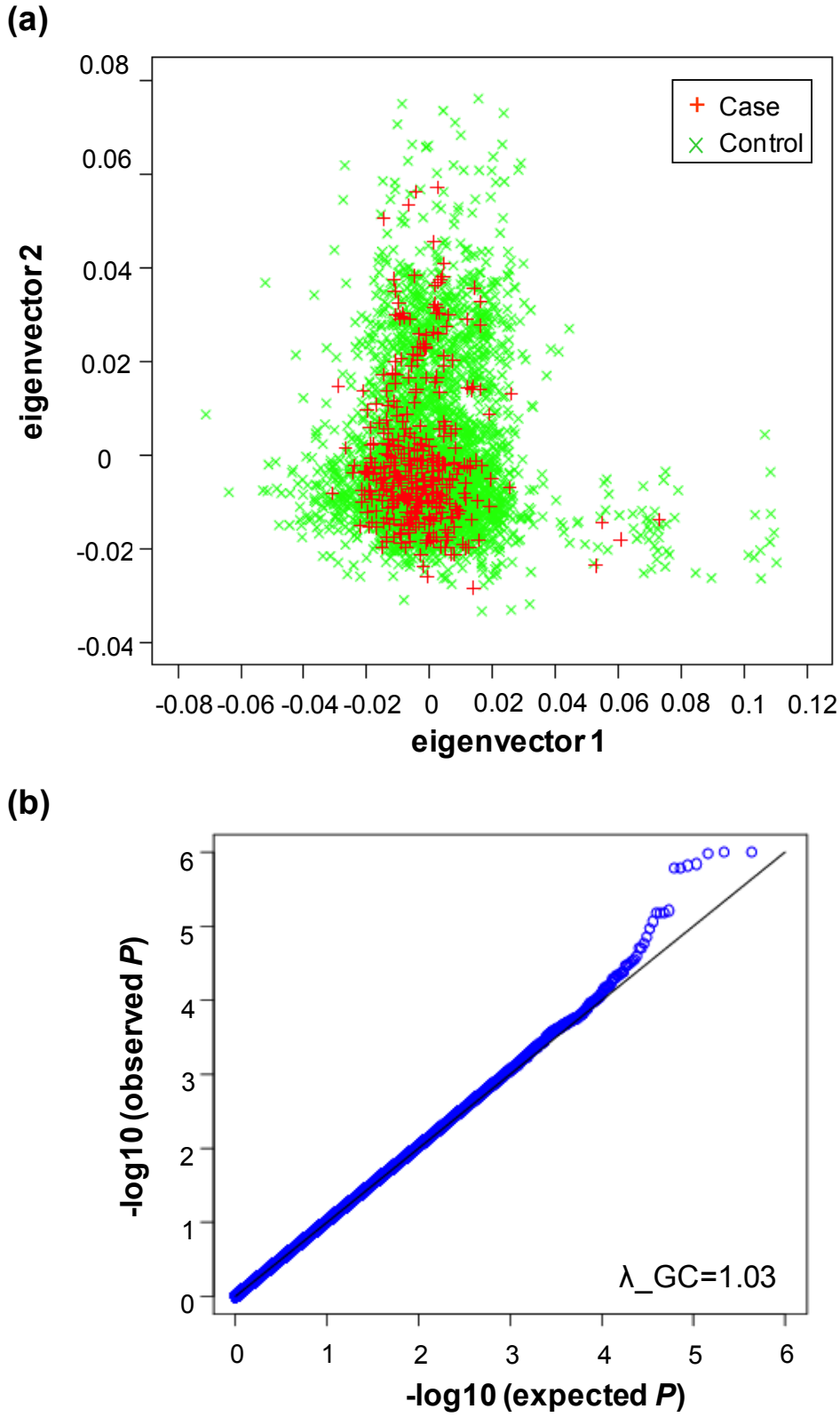
independent experiments were performed in duplicate. Independent samples t-test was performed.

Supplementary Figures:



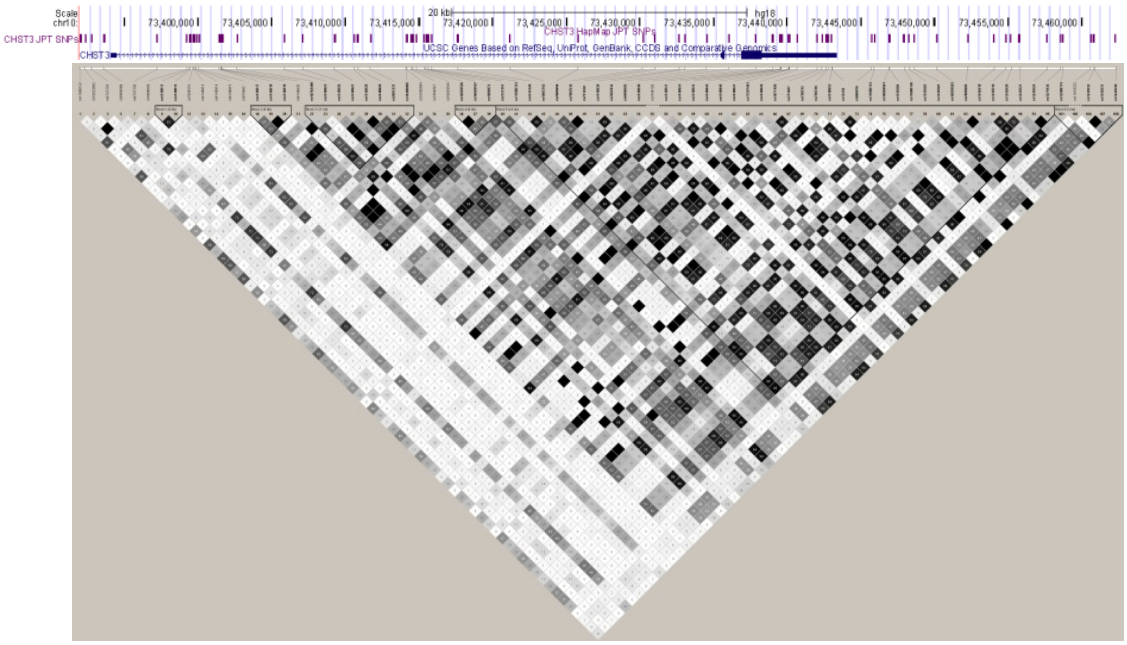
Supplementary Figure S1. Families in linkage analysis. Families 1 to 10 represent families used for first-stage genome-wide linkage analysis, while families 11-18 represent those used for second-stage linkage analysis. Arrows (➔) indicate probands who are sufficiently symptomatic to require surgery. Stripes within individuals (-) indicate individuals with LDD at the age of 30 or under. Question marks (?) indicate unknown phenotype. Color intensities of individuals indicate age-adjusted LDD score range.

* The youngest descendents are not colored as they were selected based on their age, not by age-adjusted LDD score.

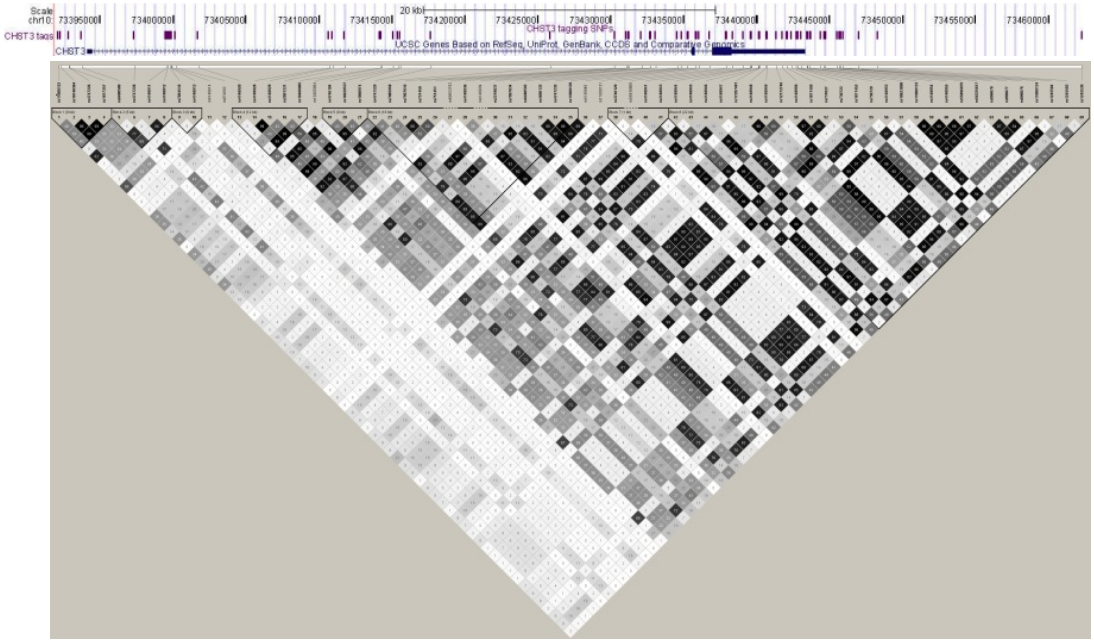


Supplementary Figure S2. Analyses of Japanese GWAS data. (a) Principal components for GWAS of Japanese cohort (J1). (b) Q-Q plot of J1 GWAS.

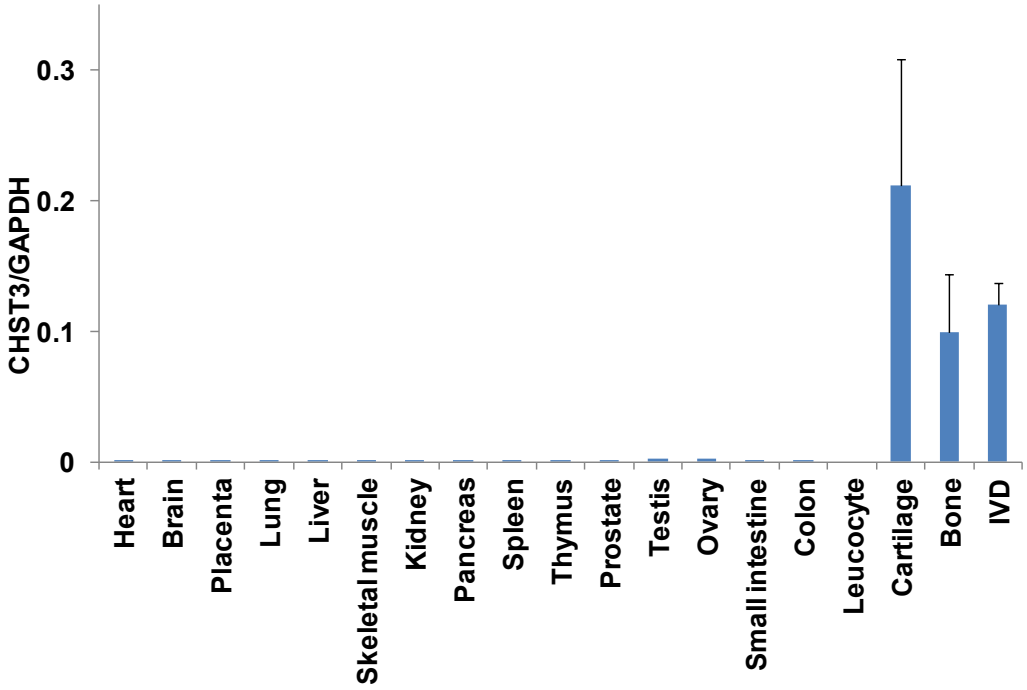
a



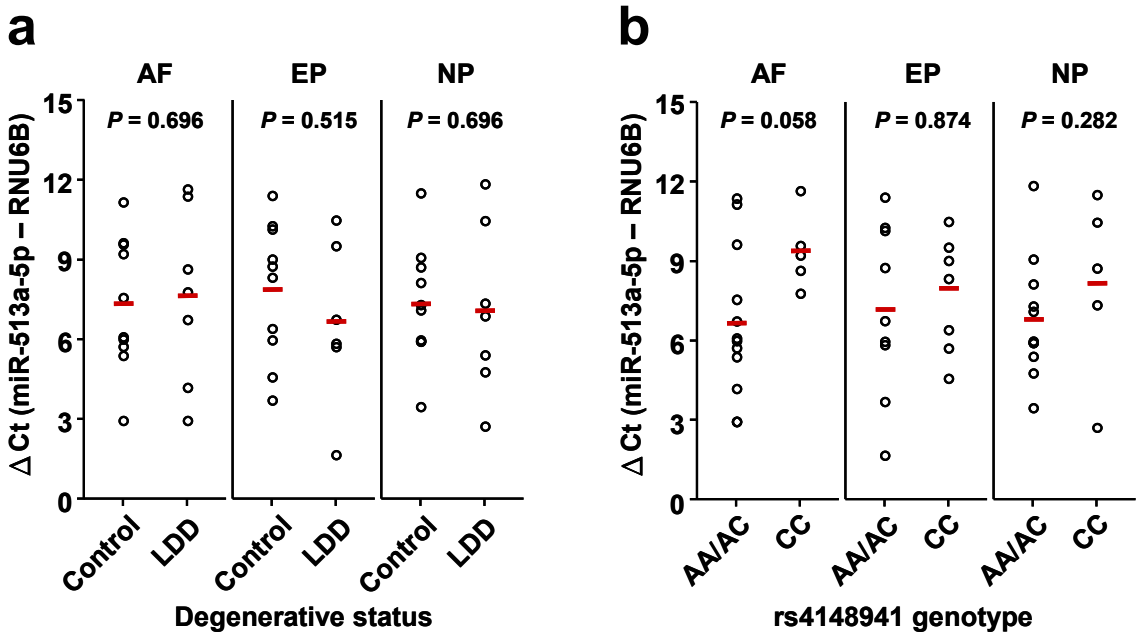
b



Supplementary Figure S3. (a) LD map of *CHST3* on the basis of the Japanese (JPT) information in the International HapMap Project database (release 24). (b) LD map of the same region generated using data from 69 tagging SNPs genotyped in a Southern Chinese cohort (SC1). The LD measures are shown as r^2.



Supplementary Figure S4. Quantitative RT-PCR of *CHST3* expression in various human tissues



Supplementary Figure S5. Quantitative analysis of miR-513a-5p expression in the human intervertebral disc. Comparisons of miR-513a-5p expression by quantitative RT-PCR between; (a) control and degenerated samples; and (b) samples differentiated according to genotypes of AA/AC and CC for rs4148941. Levels are expressed as ΔCt values above the endogenous house keeping gene for mRNA, RNU6B**.** The corresponding *P* values are indicated with <0.05 as the significant threshold. AF, annulus fibrosus; EP, end plate; NP, nucleus pulposus

Supplementary Tables:

Supplementary Table S1. Summary for NPL Z-scores of microsatellite markers in chromosome 10 for all 18 families in the second-stage linkage analysis.

New markers were selected around the region of the microsatellite (*) with the best NPL score in the first-stage linkage analysis of 10 families.

Genetic position (cM)	Marker	NPL Z-score (Age-adjusted LDD score with 0.5 as cut-off)
80.77	D10S1652	2.18
88.41	D10S1647	2.54
91.13	D10S537	3.02
93.92	D10S1146	3.04
98.96	D10S569*	3.72
105.04	D10S1686	1.72

Supplementary Table S2. Linear regression analysis of 69 SNPs of CHST3 in Southern Chinese cohort SC1.

SNP	SNP location in *CHST3*	Physical position (chr. 10)	Allele 1 (a_1)	Allele 2 (a_2)	Genotype count ($a_1a_1/a_1a_2/a_2a_2$)	Allele frequency (a_1)	β (95% CI)*	*P**
rs11000122	5' upstream	73392029	T	C	70/446/841	0.216	0.066 (-0.021 - 0.153)	1.36×10^{-1}
rs10740394	5' upstream	73392199	C	T	213/635/462	0.405	0.052 (-0.023 - 0.128)	1.75×10^{-1}
rs4747226	5' upstream	73392769	C	G	220/666/459	0.411	0.075 (0.000 - 0.150)	4.95×10^{-2}
rs1817357	5' upstream	73392786	G	A	218/655/458	0.410	0.061 (-0.015 - 0.136)	1.15×10^{-1}
rs6480589	5' upstream	73393644	C	G	49/423/891	0.191	0.026 (-0.067 - 0.118)	5.89×10^{-1}
rs4747228	intron 1	73397233	T	G	53/415/894	0.191	0.037 (-0.054 - 0.129)	4.26×10^{-1}
rs4148911	intron 1	73399423	G	C	62/472/825	0.219	0.043 (-0.045 - 0.132)	3.36×10^{-1}
rs4148912	intron 1	73399504	C	T	81/509/772	0.246	0.039 (-0.045 - 0.124)	3.61×10^{-1}
rs7894516	intron 1	73399656	C	A	42/380/941	0.170	-0.008 (-0.104 - 0.087)	8.63×10^{-1}
rs4148913	intron 1	73399758	G	A	64/472/817	0.222	0.062 (-0.026 - 0.150)	1.67×10^{-1}
rs4148914	intron 1	73399812	A	G	2/108/1255	0.041	0.127 (-0.056 - 0.310)	1.74×10^{-1}
rs874692	intron 1	73400079	A	G	48/420/893	0.190	0.022 (-0.071 - 0.115)	6.44×10^{-1}
rs4148920	intron 1	73401657	T	C	43/350/966	0.160	0.024 (-0.073 - 0.121)	6.28×10^{-1}
rs4148928	intron 1	73410619	A	G	61/471/822	0.219	0.102 (0.013 - 0.190)	2.45×10^{-2}
rs4148929	intron 1	73410843	C	A	60/463/813	0.218	0.108 (0.019 - 0.197)	1.73×10^{-2}
rs2091331	intron 1	73411712	G	A	215/633/512	0.391	0.060 (-0.013 - 0.134)	1.07×10^{-1}
rs10400065	intron 1	73414124	A	G	60/473/822	0.219	0.096 (0.007 - 0.185)	3.44×10^{-2}
rs10400064	intron 1	73414188	T	C	110/528/726	0.274	0.016 (-0.065 - 0.096)	7.02×10^{-1}
rs4746104	intron 1	73414993	C	T	62/469/799	0.223	0.073 (-0.016 - 0.162)	1.07×10^{-1}
rs16929537	intron 1	73415360	G	A	14/285/1064	0.115	0.093 (-0.022 - 0.208)	1.14×10^{-1}
rs1006974	intron 1	73415482	T	C	59/472/826	0.217	0.087 (-0.001 - 0.176)	5.41×10^{-2}
rs4747229	intron 1	73417597	T	C	12/270/1067	0.109	0.012 (-0.107 - 0.130)	8.48×10^{-1}
rs7085958	intron 1	73425784	A	G	60/473/820	0.219	0.069 (-0.020 - 0.157)	1.29×10^{-1}
rs7907616	intron 1	73430182	G	A	224/655/484	0.405	0.069 (-0.005 - 0.143)	6.79×10^{-2}
rs751450	intron 1	73431021	G	A	227/664/469	0.411	0.070 (-0.004 - 0.144)	6.40×10^{-2}
rs751451	intron 1	73431118	T	G	119/554/650	0.299	0.070 (-0.010 - 0.151)	8.69×10^{-2}
rs60653762	intron 1	73431175	T	C	2/59/1252	0.024	0.035 (-0.202 - 0.273)	7.69×10^{-1}
rs4148936	intron 1	73432062	T	A	12/263/1055	0.108	0.009 (-0.111 - 0.130)	8.77×10^{-1}
rs4148939	intron 1	73432668	T	C	3/110/1229	0.043	-0.009 (-0.187 - 0.169)	9.21×10^{-1}
rs2219837	intron 1	73432679	C	T	227/638/462	0.412	0.068 (-0.007 - 0.142)	7.42×10^{-2}
rs7097834	intron 1	73433034	T	G	16/261/1058	0.110	0.002 (-0.115 - 0.119)	9.74×10^{-1}
rs6480592	intron 1	73434515	T	C	220/655/485	0.403	0.074 (0.000 - 0.148)	5.00×10^{-2}
rs4600132	intron 1	73434785	A	G	16/251/1014	0.111	0.073 (-0.047 - 0.192)	2.32×10^{-1}
rs4747235	intron 1	73435219	C	T	10/251/1047	0.104	0.026 (-0.097 - 0.150)	6.77×10^{-1}
rs11000129	intron 2	73435779	G	T	10/246/1083	0.099	0.040 (-0.083 - 0.164)	5.22×10^{-1}
rs4148940	intron 2	73436001	G	A	12/262/1050	0.108	0.024 (-0.096 - 0.144)	6.96×10^{-1}
rs11000131	intron 2	73436734	T	C	0/31/1305	0.012	0.205 (-0.137 - 0.548)	2.40×10^{-1}
rs3740129	exon3	73437865	A	G	10/232/1123	0.092	0.078 (-0.048 - 0.204)	2.27×10^{-1}
rs55939908	3' UTR	73438301	T	G	0/35/1288	0.013	0.325 (0.002 - 0.649)	4.89×10^{-2}
rs4148941	3' UTR	73438998	A	C	224/652/484	0.404	0.076 (0.002 - 0.150)	4.31×10^{-2}
rs4148943	3' UTR	73439513	T	C	14/268/1043	0.112	0.047 (-0.071 - 0.166)	4.32×10^{-1}
rs4148944	3' UTR	73439549	A	G	103/558/690	0.283	0.081 (0.000 - 0.162)	5.12×10^{-2}
rs4148945	3' UTR	73439596	T	C	12/254/1073	0.104	0.061 (-0.060 - 0.182)	3.26×10^{-1}
rs4148946	3' UTR	73440079	T	C	215/628/473	0.402	0.089 (0.014 - 0.164)	1.96×10^{-2}
rs4148947	3' UTR	73440123	C	T	14/278/1046	0.114	0.058 (-0.058 - 0.174)	3.25×10^{-1}
rs17297481	3' UTR	73440134	A	G	11/273/1079	0.108	0.084 (-0.035 - 0.203)	1.68×10^{-1}
rs4148948	3' UTR	73440584	G	A	223/648/444	0.416	0.075 (-0.001 - 0.150)	5.19×10^{-2}
rs4148949	3' UTR	73440657	C	T	213/625/472	0.401	0.090 (0.015 - 0.165)	1.95×10^{-2}
rs12172746	3' UTR	73441288	T	C	115/555/661	0.295	0.078 (-0.002 - 0.158)	5.76×10^{-2}
rs4148950	3' UTR	73441712	A	G	10/260/1079	0.104	0.071 (-0.051 - 0.192)	2.54×10^{-1}
rs1871450	3' UTR	73442020	A	G	17/242/1065	0.104	0.085 (-0.034 - 0.204)	1.63×10^{-1}
rs731027	3' UTR	73442342	C	T	13/266/1059	0.109	0.042 (-0.076 - 0.161)	4.83×10^{-1}
rs730722	3' UTR	73442656	G	C	126/557/678	0.297	0.085 (0.007 - 0.164)	3.39×10^{-2}
rs1871452	3' UTR	73442667	A	T	223/621/480	0.403	0.091 (0.017 - 0.165)	1.58×10^{-2}
rs730720	3' UTR	73442768	T	C	17/270/1039	0.115	0.028 (-0.088 - 0.143)	6.39×10^{-1}
rs4148953	3' UTR	73443005	A	G	46/391/885	0.183	0.067 (-0.028 - 0.162)	1.67×10^{-1}
rs10823898	3' downstream	73443405	T	G	11/251/1055	0.104	0.067 (-0.054 - 0.189)	2.78×10^{-1}
rs11000133	3' downstream	73443451	A	C	111/536/653	0.292	0.082 (0.001 - 0.164)	4.86×10^{-2}
rs4148954	3' downstream	73443608	T	C	114/537/659	0.292	0.073 (-0.008 - 0.154)	7.79×10^{-2}
rs4148955	3' downstream	73444324	T	C	112/547/659	0.293	0.080 (0.000 - 0.161)	5.15×10^{-2}
rs55896879	3' downstream	73444587	A	G	104/559/661	0.290	0.076 (-0.006 - 0.158)	7.03×10^{-2}
rs56325617	3' downstream	73444670	C	T	109/552/659	0.292	0.076 (-0.006 - 0.157)	6.87×10^{-2}
rs896078	3' downstream	73445465	T	G	15/275/1036	0.115	0.057 (-0.060 - 0.174)	3.36×10^{-1}
rs896077	3' downstream	73445609	T	C	13/251/1073	0.104	0.082 (-0.038 - 0.203)	1.81×10^{-1}
rs896076	3' downstream	73445713	G	A	16/266/1032	0.113	0.070 (-0.047 - 0.187)	2.43×10^{-1}
rs11000134	3' downstream	73445889	T	C	49/399/914	0.183	0.059 (-0.034 - 0.152)	2.17×10^{-1}
rs1245584	3' downstream	73446949	A	C	15/260/1085	0.107	0.054 (-0.064 - 0.171)	3.73×10^{-1}
rs1245582	3' downstream	73448273	T	C	221/650/484	0.403	0.083 (0.009 - 0.157)	2.74×10^{-2}
rs1245530	3' downstream	73462274	A	G	9/144/1180	0.061	0.206 (0.058 - 0.354)	6.53×10^{-3}

P* values and **β for linear regression were determined using age-adjusted LDD score. SNPs for the follow-up analysis are highlighted.

Supplementary Table S3. Linear regression analysis of selected SNPs of CHST3 in the Southern Chinese cohort SC.

SNP[†]	Physical position (chr. 10)	Allele 1 (a_1)	Allele 2 (a_2)	Genotype count ($a_1a_1/a_1a_2/a_2a_2$)	Allele frequency (a_1)	β (95% CI)*	*P**
rs4747226	73392769	C	G	449/1423/910	0.417	0.055 (0.003 - 0.107)	3.90×10^{-2}
rs4148928	73410619	A	G	119/954/1877	0.202	0.076 (0.015 - 0.138)	1.44×10^{-2}
rs4148929	73410843	C	A	117/925/1788	0.205	0.060 (-0.001 - 0.121)	5.56×10^{-2}
rs10400065	73414124	A	G	123/959/1870	0.204	0.076 (0.015 - 0.137)	1.41×10^{-2}
rs6480592	73434515	T	C	455/1374/1121	0.387	0.087 (0.037 - 0.137)	6.03×10^{-4}
rs4148941	73438998	A	C	437/1326/1052	0.391	0.093 (0.043 - 0.143)	3.02×10^{-4}
rs4148946	73440079	T	C	444/1322/1059	0.391	0.093 (0.043 - 0.143)	2.86×10^{-4}
rs4148949	73440657	C	T	447/1318/1051	0.393	0.094 (0.044 - 0.144)	2.42×10^{-4}
rs730722	73442656	G	C	249/1162/1540	0.281	0.079 (0.025 - 0.133)	3.98×10^{-3}
rs1871452	73442667	A	T	447/1316/1057	0.392	0.093 (0.043 - 0.143)	2.66×10^{-4}
rs11000133	73443451	A	C	233/1140/1523	0.277	0.077 (0.022 - 0.132)	6.03×10^{-3}
rs1245582	73448273	A	G	434/1321/1050	0.390	0.094 (0.044 - 0.144)	2.60×10^{-4}
rs1245530	73462274	A	G	37/420/2253	0.091	0.127 (0.043 - 0.212)	3.29×10^{-3}

**P* values and β for linear regression were determined using the age-adjusted LDD scores.

[†]rs55939908 was excluded from the analysis due to a significant deviation from Hardy-Weinberg equilibrium.

Supplementary Table S4. Primers for PCR and functional analyses

Primers	Sequence (5' to 3')
Quantitative RT-PCR	
CHST3-1	GCAAATACGCCCTTTTCTTG
CHST3-2	TGTTGGCATCTGCTAGAGCTT
COL2A1-1	ACACTGGGACTGTCCTCTGC
COL2A1-2	CCAGGTTCTCCTTTCTGTCC
GAPDH-1	CGACCACTTTGTCAAGCTCA
GAPDH-2	AGGGGTCTACATGGCAACTG
Pyrosequencing at rs4148941	
CHST3-3	CTTATGCCCACAGGGTTTTTC
CHST3-4	CTGCCCTCGAAATAATCCTACAAA
CHST3-Pyro	ACCTCAGAGGAGCCTGT
Luciferase assay	
miR-513	CTAGAATGACACCTCCCT***GTG***AAggatccATGACACCTCCCT*GTG*AAT
miR-513-M*	CTAGAATGACACCTCCCT***CAC***AAggatccATGACACCTCCCT*CAC*AAT
miR-626	CTAGAAAGACATTTTCAGA***CAG***CTggatccAAGACATTTTCAGA*CAG*CTT
miR-626-M*	CTAGAAAGACATTTTCAGA***GTC***CTggatccAAGACATTTTCAGA*GTC*CTT
CHST3-rs4148941A	CTAGAGTATGACACACCTCAGAGGAGCCTGTG***A***TTAggatccGTATGACACACCTCAGAGGAGCCTGTG*A*TTAT
CHST3-rs4148941C	CTAGAGTATGACACACCTCAGAGGAGCCTGTG***C***TTAggatccGTATGACACACCTCAGAGGAGCCTGTG*C*TTAT
CHST3-rs4148949C	CTAGACTAGGGGCCCTGCTAATGTGGACAG***C***AGAggatccCTAGGGGCCCTGCTAATGTGGACAG*C*AGAT
CHST3-rs4148949T	CTAGACTAGGGGCCCTGCTAATGTGGACAG***T***AGAggatccCTAGGGGCCCTGCTAATGTGGACAG*T*AGAT

* Corresponding mutant (M) primers with the base changes highlighted (underlined) between the normal and mutated sequences.

References

1. Nakamura Y. The BioBank Japan Project. Clin Adv Hematol Oncol 2007; 5(9):696-697.

2. Karasugi T, Semba K, Hirose Y et al. Association of the tag SNPs in the human SKT gene (KIAA1217) with lumbar disc herniation. J Bone Miner Res 2009; 24(9):1537-1543.

3. Hirose Y, Chiba K, Karasugi T 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(5):1122-1129.

4. Virtanen IM, Song YQ, Cheung KMC et al. Phenotypic and population differences in the association between *CILP* and lumbar disc disease. J Med Genet 2007; 44(4):285-288.

5. Schneiderman G, Flannigan B, Kingston S, Thomas J, Dillin WH, Watkins RG. Magnetic resonance imaging in the diagnosis of disc degeneration: correlation with discography. Spine 1987; 12(3):276-281.

6. Wigginton JE, Abecasis GR. PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. Bioinformatics 2005; 21(16):3445-3447.

7. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 1998; 63(1):259-266.

8. Bao L, Zhou M, Wu L et al. PolymiRTS Database: linking polymorphisms in microRNA target sites with complex traits. Nucleic Acids Res 2007; 35(Database issue):D51-D54.

9. Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R. Fast and effective prediction of microRNA/target duplexes. RNA 2004; 10(10):1507-1517.