A Novel Catechol-O-Methyltransferase Variant Associated with Human Disc Degeneration

Helen E. Gruber1, Wei Sha 2,3, Cory R. Brouwer 2,3, Nury Steuerwald 4, Gretchen L. Hoelscher1, and Edward N. Jr. Hanley1

1. Department of Orthopaedic Surgery, Carolinas Medical Center, Charlotte, N.C.; 2. Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, Charlotte, N.C.; 3. Bioinformatics Services Division, North Carolina Research Campus, Kannapolis, N.C.; 4. Molecular Core and Microarray Labs, Cannon Research Center, Carolinas Medical Center, Charlotte, N.C.

Abstract

Background: Disc degeneration and its associated low back pain are a major health care concern causing disability with a prominent role in this country’s medical, social and economic structure. Low back pain is devastating and influences the quality of life for millions. Low back pain lifetime prevalence approximates 80% with an estimated direct cost burden of $86 billion per year. Back pain patients incur higher costs, greater health care utilization, and greater work loss than patients without back pain.

Methods: Research was performed following approval of our Institutional Review Board. DNA was isolated, processed and amplified using routine techniques. Amplified DNA was hybridized to Affymetrix Genome-Wide Human SNP Arrays. Quality control and genotyping analysis were performed using Affymetrix Genotyping Console. The Birdseed v2 algorithm was used for genotyping analysis. 2589 SNPs were selected a priori to enter statistical analysis using logistic regression in SAS.

Results: Our objective was to search for novel single nucleotide polymorphisms (SNPs) associated with disc degeneration. Four SNPs were found to have a significant relationship to disc degeneration; three are novel. Rs165656, a new SNP found to be associated with disc degeneration, was in catechol-O-methyltransferase (COMT), a gene with well-recognized pain involvement, especially in female subjects (p=0.01). Analysis confirmed the previously association between COMT SNP rs4633 and disc degeneration. We also report two novel disc degeneration-related SNPs (rs2095019 and rs470859) located in intergenic regions upstream to thrombospondin 2.

Conclusions: Findings contribute to the challenging field of disc degeneration and pain, and are important in light of the high clinical relevance of low back pain and the need for improved understanding of its fundamental basis.

Key words: SNP, catechol-O-methyltransferase (COMT), SNP, low back pain.

Introduction

Single nucleotide polymorphisms (SNPs) are variations in the DNA sequence wherein a single nucleotide (A, C, G or T) in the sequence is altered. SNPs can occur in both coding (exon) and non-coding (intron or intergenic) regions, but even non-coding SNPs may have a functional effect on gene expression by influencing transcription factors, miRNA binding, or splicing sites [1,2]. Prediction of potentially functional SNPs in potential drug-response genes continues to be a pharmacologic research area of great interest [3,4].

There is strong interest in genetic polymor-
phisms which are associated with human intervertebral disc degeneration, and several recent reviews have provided updates and overviews [5-7]. There is also a related high interest in the role of genes and their variants which are associated with low back pain. Variants in genes related to inflammation and matrix components have been reported to show a relation with disc degeneration and also peripheral pain [8].

COMT (catechol-O-methyltransferase) is an enzyme which inactivates catechols, including the significant neurotransmitters dopamine, noradrenaline and adrenaline [9]. In clinical studies of chronic pain, the relationship between COMT variants and pain conditions, including low back pain, is an area of considerable research interest and importance [9-12].

Materials and methods

Clinical Study Population

Study of human disc specimens was approved prospectively by the Human Subjects Institutional Review Board at Carolinas Medical Center. The need for informed consent was waived by the ethical board since disc tissue was removed as part of routine surgical practice (and discarded). Scoring of disc degeneration utilized a modification of the Thompson scoring system [13] which incorporated author ENH’s radiologic, MRI and surgical findings. The Thompson system scores disc degeneration over the spectrum from a healthy disc (Thompson grade I) to discs with advanced degeneration (grade V, the most advanced stage of degeneration) [13]. Patient specimens were derived from surgical disc procedures performed on individuals with degenerative disc disease. All surgical patients presented with back pain. Surgical specimens were transported to the laboratory in sterile tissue culture medium. Annulus cells were established in monolayer culture [14], and expanded for use in as previously described [15]. During primary culture and expansion, cells were maintained in a 37°C incubator with 5% CO₂ and 95% relative humidity.

DNA isolation and preparation

DNA was isolated according to the manufacturer’s protocol for the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI). For SNP chip analysis, the DNA was further cleaned following the manufacturer’s protocol for the kit Genomic DNA Clean & Concentrator (Zymo Research, Irvine, CA).

SNP microarray processing

The Affymetrix Genome-Wide Human SNP Nsp/Sty Assay Kit 6.0 (Affymetrix, Santa Clara, CA, USA) was used to fragment and label the DNA sample. Briefly, 250 ng of total genomic DNA was digested with Nsp I and Sty I restriction enzymes and ligated to adaptors that recognize the cohesive 4 bp overhangs. All fragments resulting from restriction enzyme digestion served as substrates for adaptor ligation regardless of size. Adaptor-ligated DNA fragments were amplified using a generic primer that recognizes the adaptor sequence. Fragments in the 200 to 1,100 bp size range were preferentially PCR amplified. PCR amplification products for each restriction enzyme digest were combined and purified using the isopropanol precipitation method. The amplified DNA was then hybridized to Genome-Wide Human SNP Arrays 6.0 (Affymetrix) containing 1.8 million genetic markers, including more than 906,600 single nucleotide polymorphisms (SNPs).

SNP Data analysis methods

Quality control and genotyping analysis were performed using Affymetrix Genotyping Console 4.1 (Affymetrix, Inc., Santa Clara, CA). Affymetrix recommended thresholds of contrast QC>=0.4 and SNP call rate > 90% were used to select samples and SNPs that qualified for downstream analysis. The Birdseed v2 algorithm was used for genotyping analysis. Due to the relatively small sample size, only 2589 SNPs on the SNP array were selected to enter statistical analysis. SNPs were selected a priori based on a literature review (Linguamatics Ltd, Cambridge, UK) and pathway analysis (Ingenuity Systems, Inc., Redwood City, CA), and were selected based on QC criteria. Selected SNPs included: 1) SNPs reported to have association with disc degeneration disease, 2) SNPs that were not reported to have association with the disease, but were on the genes, or in the intergenic regions next to the genes, that were reported to have association, 3) SNPs that were on the pathways with reported association with the disease.

Statistical analyses

Statistical analyses were performed in SAS (SAS Institute Inc., Cary, NC). Logistic regression was used to examine the association between each SNP and disc degeneration grade, where disc degeneration grade was the response variable, and SNP was the predictor variable. Age, gender, ethnicity, and spine site were considered as potential confounders and adjusted in the model. Genotypes were coded 11, 12 or 22, where allele 1 and 2 indicate the major and minor allele, respectively: “11” represents the genotype that has two major alleles; “12” represents the genotype that has one major allele and one minor allele; and “22” represents the genotype that has two minor alleles. The homozygous genotype (11 or 22) that had smaller sample size was grouped together with the hetero-
zygous genotype (12) to increase sample size. Logistic regression was performed separately for each SNP to address the following question: “Does having one or two minor alleles increase (or decrease) the likelihood of having a greater disc degeneration grade?” The following logistic regression equation was employed:

\[
\text{logit}(P(y \leq g)) = \log \left( \frac{P(y \leq g)}{P(y > g)} \right) = \alpha_g + \beta_1 S + \beta_2 A + \beta_3 G + \beta_4 R + \beta_5 S
\]

where \(\alpha_g\) is the intercept for each level of \(g\); \(g\) is the disc degeneration grade (1-5); \(\beta_1\) to \(\beta_5\) are coefficients of predictor variables; \(S = \text{SNP}\) (two categories after grouping the heterozygous genotype with one of the homozygous genotypes), \(A = \text{age}\), \(G = \text{gender}\), \(R = \text{race}\), and \(S = \text{spinal site}\).

The simple M method [16] was used to calculate the effective number of independent tests, followed by applying the Bonferroni correction for multiple testing corrections.

The SNPs that were found to be significantly associated with disc degeneration grade by the analysis described above were further analyzed by a gender-stratified logistic regression to determine if the SNP effect was different between the two genders. Stratified logistic regression for age, race, and site was not performed due to small sample size (i.e., \(n < 10\) in at least one stratum).

We recognized that the allele-dose effect could be missed due to the combination of the heterozygous group with one of the homozygous groups in the analysis described above. Therefore, for the SNPs that were found to be significantly associated with disc degeneration grade, the average disc degeneration grade was calculated for each of the three genotypes (11, 12 or 22) to determine if an allele-dose effect exists in these SNPs.

Results

The clinical study population was composed of 40 subjects for whom the Affymetrix recommended threshold (contrast QC>=0.4) was met. The study population consisted of one Grade I patient (17 yr old male), 6 grade II patients (mean age 38.6 yr ± 5.9 (mean ± s.e.m.; 1 male, 5 females), 12 Grade III patients (mean age 39.9 ± 2.8; 5 males, 7 females); 17 Grade IV patients (mean age 51.1 ± 3.2; 6 males, 11 females), and 4 Grade V patients (mean age 60.0 ± 9.1; 2 males, 2 females). All surgical patients presented with back pain. Additional demographic features shown by disc grade are presented in Table 1. As shown in this table, some sample heterogeneity in age, gender, ethnicity and spinal site was observed; therefore these factors were adjusted in the logistic regression model to remove potential confounding effects. Because the sample size in this study is small, three steps in data analysis were taken to increase statistical power. These are discussed in detail in the Discussion section below.

Based on the literature review, pathway analysis, and quality control (SNP call rate >90%), 2589 SNPs were selected a priori. These SNPs were then analyzed by logistic regression. Four SNPs (SNPs rs2095019 and rs4708592, located upstream of the THBS2 gene, and SNPs rs165656 and rs4633, located within the COMT gene) were found to be significantly associated with disc degeneration disease after adjusting for confounders, and correction for multiple testing. Note that the Bonferroni correction for multiple testing used here was based on 38 (not 2589) tests, because the Simple M method [16] revealed that there were only 38 independent tests among the 2589 tests. The Bonferroni corrected P values for association with disc degeneration grade are: 0.032, 0.049, 0.043, and 0.043, respectively. SNPs rs2095019, rs4708592 and rs165656 are novel SNPs reported here for the first time in patients with degenerative disc disease. Detailed results are shown in Table 2.

The novel pain-related SNP reported here for the first time in association with disc degeneration and pain, rs165656, lies within the COMT gene region. SNPs rs2095019 and rs4708592, novel findings in the present study, both lie in intergenic regions located upstream to THBS2 (thrombospondin 2), a matricellular protein which our laboratory has previously shown to be present in the human disc [17]. Rs2095019 is positioned 51,000 base pairs from the starting position of THBS2, and rs4708592 lays 106,329 base pairs from the starting position.

For SNPs that were found to be significantly associated with disc degeneration, the average disc degeneration grade of each genotype was calculated. For rs2095019, allele A was found to be significantly associated with lower disc degeneration grade, and an allele-dose effect was observed. The average disc degeneration grade was 2.4 for AA (\(n=5\)), 3.14 for AC (\(n=14\)), and 3.86 for CC (\(n=21\)). No clear allele-dose effect was observed for the other three SNPs. For example, for the two COMT SNPs, the highest disc degeneration grade was observed in the heterozygote group, but not in one of the homozygous groups (result not shown).

Stratified logistic regression analyses were also performed. For both the novel COMT SNP rs165656 reported here, and the previously recognized COMT SNP rs4633, a significant effect was present for female gender (\(p = 0.0104\)) (but not for male gender (\(p = 0.3783\))).
Discussion

The lifetime prevalence of disc degeneration and its associated low back pain approximates 80% with estimated health care costs exceeding $100-200 billion/year. Chronic low back pain takes a high socio-economic toll in terms of lost wages, reduced productivity; chronic, long term low back pain also takes a high personal toll, also, and can lead to disability and job loss [18]. Our identification here of a new SNP associated with a major pain gene is especially interesting in light of the high clinical relevance of low back pain and the need for an improved understanding of the fundamental basis of low back pain.

This study investigated the association between SNPs and disc degeneration. Due to the small sample size (N=40), several steps were taken to increase the power of this study. First, instead of analyzing 906,600 SNPs genotyped by the array, we analyzed 2589 SNPs that were known or have the potential to be associated with disc degeneration based on literature review and pathway analysis. Detailed SNP selection criteria are described in the Methods section. Secondly, for each SNP, we combined the heterozygote group with one of the homozygote group to increase sample size in each subgroup. Thirdly, we used the simple Monte Carlo simulation of the homozygote group to increase sample size in each subgroup. Therefore the search for the actual number of independent tests was necessary, and it greatly increased statistical power. With the three steps discussed above, we were able to detect significant associations with a relative small sample size. Additional details on these three steps are described in the Methods section.

In a recent genome-wide analysis of pain- and neurotrophin-related gene expression in the degenerating human disc (which did not include SNPs), our laboratory identified significant differences in ex-
pression of COMT in more degenerated (grades IV and V) discs vs. less degenerated discs (grades I-III) (p = 0.002), and also in surgical compared to control human discs (p = 0.035) [19]. Other pain-related genes identified in our analyses were bradykinin receptor B1 and calcitonin gene-related peptide (CGRP).

Kim and Schwartz have presented a review of pain and spinal disease [20]. Chronic low back pain is a complicated event which is related to a number of factors, including structural, psychosocial, genetic and occupational factors [7]. As noted by Tegeder and Lotsch [7], genetic profiles of individual patients are of interest since this may open an opportunity for identifying patients at high risk for chronic pain, and in the future potentially opening the opportunity for individualized pain management.

Catechol-O-methyltransferase (COMT) codes for a protein which is important in catabolic pathways of a number of pain-relevant neurotransmitters, including noradrenalin, adrenaline and dopamine [21], and is a protein which has been attributed to related issues including pain perception, mood, cognition, and responses to physical/emotional stressors [22]. Patients with a specific COMT polymorphism identified by Zubieta et al. showed higher sensory and effective pain ratings [23], and Dai et al. have identified an association between specific COMT SNPs associated with pain sensitivity and outcomes following surgery for low back pain [24], thus pointing to the potential importance of further exploration of COMT SNPs in patients with disc degeneration.

Identification of the COMT SNP rs165656 in association with disc degeneration was a novel finding in this report. This SNP has previously been reported in association with mental retardation in a Chinese population in the Qinba region [25]. Although not always included in reviews of low back pain and SNPs, (such as reviews of Eskola et al. [6] and of Mayer et al. [5]), Tegeder and Lotsch have provided a recent 2009 review of COMT genetic variants related to back pain [7]. Connections between the COMT SNP rs4633, identified in our present study, and pain have been previously reported and provide evidence for modulation of low back pain by COMT variants [7], and the relationship to disc degeneration and pain [10;24]. The most common and well-studied COMT polymorphism associated with pain appears to be V158M (val158met; rs4680). This SNP was not identified in the present study population.

In humans, associations have been reported between COMT functional haplotypes and thermal pain perception and capsaicin-induced pain [22]. COMT variants have been seen to influence pain phenotypes, especially in women with chronic pain reminding us that in the future COMT genetic differences may be of clinical relevance in tailored pain management [26]. COMT variants have also been noted to influence the efficacy of morphine therapy in cancer patients with pain [27].

Our finding of a significant effect related to female gender status in both the novel COMT SNP reported here and the COMT SNP rs4633 is similar to the significant relationship seen with female gender in the work of Belfer et al. in an average pain sensitivity COMT haplotype [22]. These authors commented that such findings reflect a fundamental contribution of COMT to the complex area of pain processes. Belfer and Segall also remind us that the pain experience by humans has contributions from both genetic and environmental factors [26].

For the two COMT SNPs, it is interesting that the highest disc degeneration grade was observed in the group of patients with one minor allele (CG for rs165656 and CT for rs4633), but not in the group with two minor alleles (CC for rs165656 and TT for rs4633). The lack of allele-dose effect could be due to small sample size in this study, but it is also important to note that previous studies have also observed a lack of allele-dose effect in COMT SNPs [10;28;29].

Note should also be made concerning the two novel disc degeneration-associated SNPs reported here (rs2095019 and rs4708592) which lie in the intergenic region near the THBS2 gene. Genes in noncoding regions, and even silent SNPs, can have effects on cellular processes, including DNA transcription, RNA splicing, mRNA stability or mRNA transport/translation [9]. THBS2 is an important matricellular protein which is present within the disc [17], and which is known to modulate collagen fibrillogenesis [30] and matrix metalloproteinase 2 interactions [31;32], biologic functions critically important to the disc. It is also interesting to note that our previous work with THBS2 null mice showed disruption of the lamellar morphology in the disc [33]. Related important observations include the previous work by Hirose et al. which identified a THBS2 SNP associated with disc degeneration [34], and the work of Valdes et al which showed an association of a thrombospondin SNP with the radiographic progression of disc degeneration [35]. These observations increase our interest in the relationship of disc degeneration to the two intergenic SNPs near the THBS2 gene.

In summary, the findings reported here contribute to the challenging field of disc pain by identifying a new COMT SNP associated with disc degeneration. COMT variants continue to be of high interest, especially in women with chronic pain. The present work also identified two novel disc degeneration-related SNPs located upstream of the THBS2 gene. Findings are important in light of the high clin-
Axial relevance of low back pain and the need for improved understanding of the fundamental basis of low back pain.

Acknowledgements

We thank Synthia Bethea for expert technical assistance in DNA isolation. HEG and ENH wish to thank the Brooks Back Pain Research Endowment for general laboratory support.

Competing Interests

The authors have declared that no competing interest exists.

References

1. Mooney S. Bioinformatics approaches and resources for single nucleotide polymorphism functional analysis. Brief Bioinform 2015; 6:44-56.
2. Wasserman WW, Sandelin A. Applied bioinformatics for the identification of regulatory elements. Nat Rev Genet 2004; 5:276-287.
3. Pang GSY, Wang J, Wang Z, Lee CGL. Predicting potentially functional SNPs in drug-response genes. Pharmacogenomics 2009; 10:639-653.
4. Shastry BS. SNPs in disease gene mapping, medicinal drug development and evolution. J Human Genetics 2007; 52:871-880.
5. Mayer JE, latridis JC, Chan D, Qureshi SA, Gottesman O, Hecht AC. Genetic polymorphisms associated with intervertebral disc degeneration. The Spine Journal 2013; 13:299-317.
6. Escola PJ, Lemmelä S, Kjaer P, Solovieva S, Männikkö M, Tommerup N et al. Genetic association studies in lumbar disc degeneration: A systematic review. PLoS ONE 2012; 7:e49995.
7. Tegeder I, Lötsch J. Current evidence for a modulation of low back pain by human genetic variants. J Cell Mol Med 2009; 13:1605-1619.
8. Omair A, Holder M, Lie BA, Reikeras O, Bros JL. Treatment outcome of chronic low back pain and radiographic lumbar disc degeneration are associated with inflammatory and matrix degrading gene variants: a prospective genetic association study. BMC Musculoskeletal Dis 2013; 14:105.
9. Andersen S, Skorpen F. Variation in the COMT gene: implications for pain perception and pain treatment. Pharmacogenomics 2009; 10:669-684.
10. Omair A, Lie BA, Reikeras O, Holden M, Bros JI. Genetic contribution of catechol-O-methyltransferase variants in treatment outcome of low back pain; a prospective genetic association study. BMC Musculoskeletal Dis 2012; 13:76.
11. Jacobsen LM, Schistad EI, Storesund A, Pedersen LM, Rygh LJ, Roe C et al. The COMT rs4680 Met allele contributes to long-lasting low back pain, sciatica and disability after lumbar disc herniation. Eur J Pain 2012; 16:1004-1009.
12. Kambur O, Männistö PT. Catechol-O-methyltransferase and pain. Intern Rev Neurobiol 2010; 95:227-279.
13. Thompson JP, Pearce RH, Schechter MT, Adams ME, Tsang IKY, Bishop PB. Preliminary evaluation of a scheme for grading the gross morphology of the human intervertebral disc. Spine 1990; 15:411-415.
14. Gruber HE, Fisher EC, Jr., Desai B, Stasky AA, Hoelscher G, Hanley EN. Human intervertebral disc cells from the annulus: Three-dimensional culture in agarose or alginate and responsiveness to TGF-β1. Exp Cell Res 1997; 235:13-21.
15. Gruber HE, Hoelscher GL, Leslie K, Ingmar JA, Hanley EN Jr. Three-dimensional culture of human disc cells within agarose or a collagen sponge: assessment of proteoglycan production. Biomaterials 2006; 27:371-376.
16. Gao X, Starmer J, Martin ER. A genetic testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet Epidemiol 2008; 32:361-369.
17. Gruber HE, Ingmar JA, Hanley EN Jr. Immunolocalization of thrombospondin in the human and sand rat intervertebral disc. Spine 2006; 31:2556-2561.
18. Katz JN. Lumbar disc disorders and low-back pain: Socioeconomic factors and consequences. J Bone Joint Surg [Am] 2006; 88A, Suppl 2:W21-22.
19. Gruber HE, Hoelscher GL, Ingmar JA, Hanley EN. Genome-wide analysis of pain-, nerve- and neurotrophin-related gene expression in the degenerating human annulus. Molecular Pain 2012; 8:63.
20. Kim DL, Schwartz CE. The genetics of pain: implications for evaluation and treatment of spinal disease. The Spine Journal 2010; 10:827-840.
21. Andersen S, Skorpen F. Variation in the COMT gene: implications for pain perception and pain treatment. Pharmacogenomics 2009; 10:669-684.
22. Belfer I, Segall SK, Laviviere WR, Smith SB, Dai F, Slade GD et al. Pain modality- and sex-specific effects of COMT genetic functional variants. Pain 2013; 154:1368-1376.
23. Zubiate J-K, Heitzeg MM, Smith YR, Bueller JA, Xu K, Xu Y et al. COMT genotype affects μ-opioid neurotransmitter responses to a pain stressor. Science 2003; 299:1240-1243.
24. Dai F, Belfer I, Schwartz CE, Banco R, Martha JF, Tighiougart H et al. Association of catechol-O-methyltransferase genetic variants with outcome in patients undergoing surgical treatment for lumbar degenerative disc disease. The Spine Journal 2010; 10:949-957.
25. Zhang K, Gao J, An C, Gao X, Zheng Z, Li R et al. An association study between catechol-O-methyltransferase gene and mental retardation in the Chinese Han population. Neurosci Lett 2007; 419:83-87.
26. Beller I, Segall SK. COMT genetic variants and pain. Drugs of Today 2011; 47:457-467.
27. Rakvåg TT, Ross JR, Sato H, Skopfen F, Kaasa S, Klepstad P. Genetic variation in the catechol-O-methyltransferase (COMT) gene and morphine requirements in cancer patients with pain. Molecular Pain 2008; 4:44.
28. Gütosy S, Erdal E, Herken H, Madersbacher E, Aheeleib E, Erdal N. Significance of catechol-O-methyltransferase gene polymorphism in fibromyalgia syndrome. Rheumatol Int 2002; 23:104-107.
29. Kim H, Mittal DP, Iadarola MJ, Dionne RA. Genetic predictors for acute experimental cold and heat pain sensitivity in humans. J Med Genet 2006; 43:e10.