
Genetics of Lumbar Disc Degeneration: Technology, Study Designs, and Risk Factors

Authors: ¹ Patrick Yu-Ping Kao, PhD, ¹ Danny Chan, PhD, ² Dino Samartzis, DSc,
³ Pak Chung Sham, BM Bch, PhD, ¹ You-Qiang Song, PhD

Affiliations: ¹ Department of Biochemistry, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong SAR, China; ² Department of Orthopaedics and Traumatology, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong SAR, China; and the ³ Department of Psychiatry, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong SAR, China

Keywords: genetics, lumbar disc degeneration, association study, twin studies, familial aggregation

Correspondence: You-Qiang Song
Department of Biochemistry
Faculty of Medicine Bldg
21 Sassoon Road
Pokfulam
Hong Kong SAR, China
Email: songy@hku.hk

ABSTRACT

Lumbar disc degeneration (LDD) is a common musculoskeletal condition. Genetic risk factors have been suggested to play a major role in its etiology. The following article reviews the main research strategies that have been used to study the genetics of LDD, and the genes that thus far have been identified to influence susceptibility to LDD. With the rapid progress in genomic technologies, further advances in the genetics of LDD are expected in the next few years.

Introduction

Low back pain (LBP) is a common musculoskeletal disorder and a global burden. Approximately 70-80% of people have experienced LBP at some point in their life.¹ The annual prevalence of LBP ranges from 15% to 45%, but is largely dependent on population under study and surveillance methods^{1,2}. Low back pain can lead to diminished function, loss of productivity, loss of work, and psychological distress, and increased health-care costs. As such, LBP poses considerable financial burden to nations through the direct cost spent on LBP treatments and the indirect cost due to loss in working productivity.^{3,4} In fact, in United States, one of the most frequent reasons for medical consultation is LBP^{1,3,4}. Therefore, the impact of LBP upon society is substantial and understanding of its development is imperative (**see article by Karppinen *et al***).

Although LBP is a prevailing disorder, the factors that contribute to LBP are still not fully understood. Indeed a variety of factors, such as occupation, biopsychological factors, cardiovascular disorders, smoking, and obesity among others have been suggested to be associated with LBP⁵⁻⁹. Conversely, British twin studies showed that there is excessive concordance of LBP for monozygotic (MZ) twins over dizygotic (DZ) twins, resulting in heritabilities ranging from 52% to 68% for various LDD phenotypes¹⁰. In addition, according to a study by Battie *et al*¹¹ consisting of Finnish Twins noted that heritability estimates ranged from 30% to 39% for various definitions of back pain problems over the prior year. Apparently, genetic factors play a vital role to the aetiology of LBP.

Intervertebral disc (IVD) degeneration is one proxy by which genetic risk factors act to influence LBP^{10,11}. This is revealed by studies that found IVD degeneration to be associated with LBP¹¹⁻²⁰. Although LBP is a multifactorial condition, discogenic origin can be mediated by such mechanisms as nerve in-growth in degenerated discs, nerve root compression from herniated discs and inflammatory responses of surrounding tissues in response to disc pathology^{11,21}. Due to the close relationship between LBP and LDD and the fact that both LDD and LBP have a large genetic component in their etiology, it is worthwhile to investigate the genetics of LDD to unveil the mechanism by which LDD and thus LBP develop. Thus, the purpose of this article is to discuss and raise awareness of genetics factors related to LDD. Furthermore, this article will also address the genetic-based technologies, methodologies and study designs that have been developed to assess LDD.

Genetic Component of Lumbar Disc Degeneration

The first step in studying genetics of a condition is to ask the following -- is there a genetic component for a disease or its symptoms? The initial step in approaching this issue is to determine if there is **familial aggregation**, that is if a relatively higher number of occurrences exist of diseased individuals among family members than in that of non-diseased individuals. On that basis, various studies were carried out to decipher if familial aggregation is indeed present in LDD. Varlotta *et al*²² found that in adolescent patients with disc herniation that a larger proportion of individuals had a positive family history of disc herniation than in control subjects. This finding was subsequently further supported by other investigators, such as Frino *et al*²³ and Matsui *et al*²⁴. According to a study by Postacchini and colleagues²⁵ consisting of discogenic LBP patients, the authors noted a large proportion of first relatives that had experienced discogenic LBP and who undergone disc surgery. Moreover, based on a study by Simmons *et al*,²⁶ patients who had undergone spine surgery had a greater frequency of relatives having experienced LBP and sciatica. Other reports have also noted a positive family history of LBP²⁷ and disc degeneration^{28,29} in spine surgery patients. In short, these reports and others have indicated the presence of familial aggregation for LDD and as such the possibility of a genetic component influence on LDD.

Although studies have illustrated an association between familial aggregation and LDD, they only suggest a "possibility" of a genetic component. The reason being is that we still need to differentiate between familial aggregation that is due to social-behavioural factors from that attributed to genetic factors. This can be achieved through **twin studies** in which assess differences in disease concordance rates between MZ twins and DZ twins. If there is a strong genetic background for a disease, MZ twins should have a greater concordance rate in disease status than do DZ twins. From the twin studies, heritability, defined as the proportion of a population's phenotypic differences due to genetic variation, can be estimated. In 1995, based on 155 identical male twins of the Finnish Twin Spine Study, Battie *et al*³⁰ found that genetic risk factors play a major role in the development of LDD as noted on magnetic resonance imaging (MRI), explaining 77% of the variability. Soon thereafter in 1999 and based on UK Twin Study, Sambrook *et al*³¹ assessed LDD based on MRI of 172 MZ twins and 154 DZ twins, and noted a 74% heritability rate of LDD. The findings from these seminal studies based on Finnish and UK twins greatly substantiated the role of genetics as an etiological risk factor in the development of LDD.

Although the importance of genetics can be established, there are two interesting points to note. First, there is no single gene alone that is substantial for causing LDD. Based on a study by Livshits *et al*³² assessing an Arabic pedigree, the authors reported that a family history was found to be a risk factor for the development of LDD. Interestingly, a simple monogenic Mendelian pattern of inheritance was rejected, suggesting that there could be a more complex

mode of inheritance for LDD involving multiple genes. Secondly, genetic factors may interact with environmental factors in the etiological framework of LDD. For example, based on early reports of the Finnish Twin Study, cigarette smoking status was found to greatly increase the ability of predicting LDD.³³ Therefore, since gene-environment interaction effects may be present and oftentimes may be complex to disseminate, this presents a challenge to the “hunting down” of the disease mechanism of LDD.

Defining the Phenotype of Disc Degeneration

Throughout the years, the integrity of the intervertebral discs has been assessed by various imaging modalities, such as plain radiographs, discography, computed tomography (CT) and MRI.³⁴⁻³⁸ Standing plain radiographs or x-rays are commonly utilized to assess disc space height and sclerosis of the endplate as well as other spinal and alignment abnormalities; however, this imaging lacks the ability to assess the soft tissue of the disc. Discography entails injection of provocative material directly into the disc to identify discogenic origin of the pain source; however, this methodology is oftentimes uncomfortable to the individual and may accelerate LDD due to the direct disc injury that may be induced by the needle puncture.³⁹ Computed tomography is a method widely used for chest and abdominal imaging. However, it is less useful for examining soft tissues of this disc and is associated with excessively high-levels of ionizing radiation exposure. Since the hallmark of disc degeneration entails the loss of water and proteoglycan content in the nucleus pulposus with structural changes of the disc and adjacent endplate (**see article by Chan *et al***), this process entails various stages of degeneration that are best assessed with imaging sensitive to such alterations. As such, MRI is a non-invasive method of imaging that allows direct evaluation of the soft tissues of the disc and as such is a desirable method, or rather a more sensitive method than the alternative imaging modalities, for intervertebral disc imaging in assessing the phenotype of LDD. Throughout the years, MRI technology has further been developed to assess the integrity of the disc in a more sensitive and quantitative manner (**see article by Majumdar *et al***). Nonetheless, due to the variety of imaging modalities in existence and the presence of numerous classification schemes in assessing disc degeneration, this has in fact posed a dilemma in adopting a universal methodology to assess the phenotype of LDD.^{13,40} Therefore, tremendous heterogeneity exists in the imaging phenotype of LDD between various observational cohort and genetic studies.

Integration of Knowledge of the Human Genome for Genetic Studies

The study of genetics was pioneered by Gregor Mendel who developed Mendel's laws of genetics that allowed us to relate differences in genetic material (i.e. DNA) of organisms to appearances or conditions that are observable (i.e. phenotypes). Two major factors have propelled genetics even further: our knowledge of all the variants present in our genome, and the advancement in capturing these variations, or genotyping, technologies.

Each unit of DNA consists of an alkaline structure, known as a base pair. Single nucleotide polymorphisms (SNPs) are single base-pair changes among individuals. Microsatellites are short repeated DNA fragments. Copy number variations (CNVs) are longer (> 1 kb, kilo base pair) DNA fragments that have a different number of copies compared with a reference genome. Before the first decade of the 21st century genetic studies mainly focused on microsatellite and SNP markers. Nowadays, considerable attention has been given to the importance of CNVs. Presently, in comparison with SNP and CNVs, microsatellites are relatively seldom used for genetic studies.

Genome Projects for Human Genetics

Improvements in genotyping technologies have enriched our knowledge of the relationship between genetic sequence and disease. The large-scale study of this relationship was made possible by the Human Genome Project⁴¹, which provided the first reference human genome sequence. Later on, the HapMap project was initiated, with the main goal of capturing the pattern of SNPs, the most common type of genetic variation among individuals, within the human genome^{42,43}. The initial stages of the project included Yoruban individuals from Africa, Caucasian individuals from the United States, Chinese, and Japanese individuals. Around 3.1 million SNPs were successfully genotyped⁴². The success of the HapMap project was a big step forward in supporting genetic studies. It helped us understand the linkage disequilibrium (LD) pattern within the human genome, enabling us to choose only one or a small number of "tagging" SNPs within a high LD (low recombination rate) region to genotype to obtain almost complete common variant information on a particular genomic region. This process helps save a huge amount of resources for genetic studies. Recently, the 1000 Genome Consortium has completed its initial sequencing of 179 individuals⁴⁴, signifying the next era of genetic studies, which consists of sequencing each base-pair of DNA -- known as "deep resequencing" as opposed to genotyping, which captures only known varying regions of genome. In terms of disease genetic study, sequencing information enables us to study rare and previously unknown sequence variants. The 1000 Genome Project data are also valuable to allow us to "predict" the genotypes of SNPs not genotyped in the actual experiment by a computational process known as imputation⁴⁵. With the denser map, locations of disease-causing variants can be determined with more precision.

Genotyping Technologies

Modern genotyping technologies can be divided into two main categories: (1) those that are suitable for genotyping a larger number of SNPs and are thus suitable for genome-wide association studies (GWAS), and (2) those that are optimal for large sample size and are thus suitable for replication studies⁴⁶. The most popular GWAS genotyping platforms are Illumina's Infinium BeadChip and Affymetrix GeneChip, while other assays such as Perlegen and Invader are also available⁴⁶. Illumina BeadChip and Affymetrix GeneChip are now capable of genotyping over 1 million SNPs and CNVs in the genome. Because of the design of the products, Illumina BeadChip has a higher genomic coverage, meaning that the SNPs of the product is more capable of capturing information of the genome⁴⁷. However, both Affymetrix and Illumina, can also be used for genotyping smaller, more confined genotypic regions.

Sequenom MassArray is one platform used for a larger sample number with a small number of SNPs. I-plex assays for MassArray can be used for genotyping up to 40 SNPs. Alternatively, Taqman assay (from Applied Biosystems) uses a real-time PCR-based method for SNP detection. Different alleles have different fluorescent probes. Therefore, the detection of different alleles can be achieved from signals given by their corresponding probes. The accuracy of this assay can reach up to 99%⁴⁶. Other genotyping assays, such as Pyrosequencing, are also available⁴⁸.

Main Genetic Study Strategies

Linkage and association strategies of genetic studies

Genetic studies are divided into two main streams: (1) family linkage analysis and (2) case-control association approach. In **family linkage analysis**, the pattern of inheritance of a disease is compared with the pattern of disease markers. If a particular region of markers can be observed to be inherited in the same way as the disease, then the disease causing variant may lie within the region. Family data are required for carrying out linkage studies.

In the **case-control association approach**, unrelated individuals with extreme disease statuses, namely, the case group (those with the phenotype) and the control (those without the phenotype) group, are recruited. The allele frequencies of the genetic variants between cases and controls are compared. Disease-causing variants are identified by searching for markers that show statistically significant differences in allele frequencies between the two groups.

Association within family trios can also be carried out using transmission disequilibrium test (TDT). This test is conducted by observing in heterozygous parents if there is one allele with higher transmission frequency to the offspring than the other⁴⁹. This design is advantageous than the case-control design in that it avoids false positives caused by the differences in population composition between the case and control groups. However, the

power of this test is lower than that of the case-control design because the two parents are approximately equivalent to only one unrelated control subject in information⁴⁹.

Depending on the SNP selection strategy, the case-control association approach can be divided further into the candidate gene and genome-wide approaches. The **candidate gene approach** has the main advantage that it is more specific, as the selection of SNPs is based on the prior knowledge on the biology of the disease. When a high-throughput genotyping platform is not so readily available, this approach has been shown to be successful in identifying a number of genetic risk variants for some complex diseases. Currently due to the availability of whole-genome genotyping platforms, the **genome-wide approach** for association studies has been made possible. The main advantage of this approach is that regions of interest need not be confined in particular genomic regions, relying on investigators' knowledge. Therefore, novel genetic variants are more readily identified.

Strengths and weaknesses of linkage and association approaches (Table 1)

Linkage studies were attractive at the time when genetic marker maps and genotyping technologies allowed few markers to be studied. This is because linkage can be detected over long distances from a disease susceptibility locus. Alternatively, an important disadvantage of linkage studies is that they are powerful only when the mutation has a very high probability of causing disease (high penetrance). Association studies rely on linkage disequilibrium between genetic markers and disease susceptibility locus, and this typically requires the two loci to be very close to each other. In consequence, association studies were for a long time confined to the study of specific candidate genes supported by a strong hypothesis, and a genome-wide screen using association was not feasible until the recent developments in genotyping technologies mentioned above. An important advantage of association over linkage is that association has much greater power to detect a locus with small or modest effect size. However, linkage is more robust to allelic heterogeneity, which refers to the presence of multiple risk alleles at a locus. With modern sequencing technology, linkage and association are best considered to be two complementary ways of using genetic markers to study the genetic basis of simple and complex diseases.

Findings on the Genetics of Lumbar Disc Degeneration

Candidate gene studies

Genetic risk factors of LDD have been quite extensively studied in light of the findings from family and twins studies, which show that LDD has a substantial genetic component. Several articles have reviewed these genetic risk factors⁵⁰⁻⁵³. Today most of the genetic studies of LDD adopted the **candidate gene approach**. Genetic studies on LDD have focused on the genes that code for the functioning molecules in the disc, including the components of collagen IX⁵⁴⁻⁵⁶, aggrecan^{57,58}, degrading enzymes such as matrix metalloproteinase II⁵⁹ and III^{60,61}, inflammatory signaling molecules such as interleukin I (IL-1)⁶²⁻⁶⁴. Other genes identified associated with LDD are susceptibility genes related to other diseases. For instance, asporin⁶⁵ was first identified in rheumatoid arthritis⁶⁶, osteoarthritis (OA) and knee OA^{65,67,68}; collagen I⁶⁹ and vitamin D receptor (VDR)⁷⁰⁻⁷² are related to osteoporosis (collagen I: ⁷³⁻⁷⁵; VDR: ^{76,77}), and CILP⁷⁸ is found in cartilage⁷⁹.

However, most of these studies need to be confirmed with a larger sample size or a meta-analysis. A linkage study is also being carried out by our group (unpublished). The results show that both association and linkage approaches can identify genetic risk factors of LDD.

Current and Future Trends

As the field of genetics is rapidly developing, genetic studies on LDD will become more comprehensive and more powerful. In the following section, we discuss areas of advancement and future trends in genetics of LDD.

Marker selection and analysis

Genetic studies on LDD have previously concentrated on the candidate gene approach. Although this approach has fulfilled its mission successfully in the past, it is heavily dependent on what we know about the disease, making finding novel genes difficult. Therefore, our study group has initiated genome-wide linkage and association approaches, to fill the “knowledge gap” and lead to the identification of novel genes in relation to LDD. In our linkage study, we recruited early-onset LDD patients and their family members for whole-genome linkage scan. For our GWAS, we selected individuals with extreme disc degeneration status, after having corrected for age. The synergy of the two approaches can help locate genetic risk factors with different properties, thus making our knowledge base for the genetics of LDD more complete. In the future, the genomes and exomes of individuals are expected to be sequenced with next-generation sequencing methodologies, enabling us to look directly for rare variants in individuals and helping us decipher the genetic mechanism related to LDD.

Aside from genes that increase the risk of LDD, the focus can be placed in searching for protective genes. This broadening of the mind-set will enable us to discover gene sets that may influence LDD through other mechanisms.

In terms of statistical methodologies, multivariate approaches are currently being developed. These statistical approaches can make use of information from several genetic variants and phenotypes. By combining information from different genes, we will be able to consider variants within genes as whole or genetic information across different genes. The combined effects of different genes can then be considered simultaneously. Moreover, disc degeneration phenotypes can be combined similarly. This combination may help us identify if certain combinations of genes or phenotypes are related and therefore, share similar pathologic mechanisms. Considering the sets of phenotypes and genes can also reduce the number of statistical tests required and increase the statistical power.

Study design

Most LDD genetic studies are cross-sectional, which means only disease condition at a certain time point can be observed. As databases of volunteers develop, a longitudinal study approach can be adopted by recruiting the volunteers for a second imaging. The possibility of a dynamic assessment of the progression of LDD condition would allow us to investigate the influence of genetic risk factors to LDD and validate hypotheses about their effects.

Increasing sample size by international collaboration

In genetic studies, statistical power is one of the most important considerations for likely success. The most efficient and economic way (from each research centre's point of view) to increase sample size is to combine study results from different research groups. Meta-analysis methodologies can directly combine the significance and effects from multiple studies. With this method, samples from different collaborators can be virtually combined in a simple way. Currently, many of the GWASs are carried out by meta-analysis across different populations. This trend is expected to continue for polygenic diseases like LDD, and is probably a good way to "ride the wave" of improving technologies.

Summary

Genetics plays a role to the etiology of LDD. However it may either exert effect by itself or with other environmental factors. Moreover, our knowledge of the genetic risk factors of LDD was not yet very complete. In future, together with our knowledge and advancement in genotyping technologies, we can extend our study approach in both the directions of finding novel genes that is related to LDD, and analyze LDD etiology with other behavioral and environmental factors.

References

1. Andersson GB. Epidemiological features of chronic low-back pain. *Lancet* 1999;354:581-5.
2. Manchikanti L, Singh V, Datta S, et al. Comprehensive review of epidemiology, scope, and impact of spinal pain. *Pain Physician* 2009;12:E35-70.
3. Deyo RA, Tsui-Wu YJ. Descriptive epidemiology of low-back pain and its related medical care in the United States. *Spine* 1987;12:264-8.
4. Hart LG, Deyo RA, Cherkin DC. Physician office visits for low back pain. Frequency, clinical evaluation, and treatment patterns from a U.S. national survey. *Spine* 1995;20:11-9.
5. Samartzis D, Karppinen J, Mok F, et al. A population-based study of juvenile disc degeneration and its association with overweight and obesity, low back pain, and diminished functional status. *J Bone Joint Surg Am* 2011;93:662-70.
6. Shiri R, Karppinen J, Leino-Arjas P, et al. The association between obesity and low back pain: a meta-analysis. *Am J Epidemiol* 2010;171:135-54.
7. Shiri R, Solovieva S, Husgafvel-Pursiainen K, et al. The association between obesity and the prevalence of low back pain in young adults: the Cardiovascular Risk in Young Finns Study. *Am J Epidemiol* 2008;167:1110-9.
8. van Tulder M, Koes B, Bombardier C. Low back pain. *Best Pract Res Clin Rheumatol* 2002;16:761-75.
9. Videman T, Nurminen M, Troup JD. 1990 Volvo Award in clinical sciences. Lumbar spinal pathology in cadaveric material in relation to history of back pain, occupation, and physical loading. *Spine* 1990;15:728-40.
10. MacGregor AJ, Andrew T, Sambrook PN, et al. Structural, psychological, and genetic influences on low back and neck pain: a study of adult female twins. *Arthritis Rheum* 2004;51:160-7.
11. Battie MC, Videman T, Levalahti E, et al. Heritability of low back pain and the role of disc degeneration. *Pain* 2007;131:272-80.
12. Luoma K, Riihimaki H, Luukkonen R, et al. Low back pain in relation to lumbar disc degeneration. *Spine (Phila Pa 1976)* 2000;25:487-92.
13. Chou R, Qaseem A, Owens DK, et al. Diagnostic imaging for low back pain: advice for high-value health care from the American College of Physicians. *Ann Intern Med* 2011;154:181-9.
14. de Schepper EI, Damen J, van Meurs JB, et al. The association between lumbar disc degeneration and low back pain: the influence of age, gender, and individual radiographic features. *Spine (Phila Pa 1976)* 2010;35:531-6.

15. Kjaer P, Leboeuf-Yde C, Korsholm L, et al. Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. *Spine (Phila Pa 1976)* 2005;30:1173-80.
16. Kjaer P, Leboeuf-Yde C, Sorensen JS, et al. An epidemiologic study of MRI and low back pain in 13-year-old children. *Spine (Phila Pa 1976)* 2005;30:798-806.
17. Samartzis D, Karppinen J, Chan D, et al. The association of disc degeneration based on magnetic resonance imaging and the presence of low back pain. World Forum for Spine Research: Intervertebral Disc. Montreal, Canada, 2010.
18. Savage RA, Whitehouse GH, Roberts N. The relationship between the magnetic resonance imaging appearance of the lumbar spine and low back pain, age and occupation in males. *Eur Spine J* 1997;6:106-14.
19. Takatalo J, Karppinen J, Niinimäki J, et al. Does lumbar disc degeneration on MRI associate with low back symptom severity in young Finnish adults? *Spine (Phila Pa 1976)* 2011;[Epub ahead of print].
20. Visuri T, Ulaska J, Eskelin M, et al. Narrowing of lumbar spinal canal predicts chronic low back pain more accurately than intervertebral disc degeneration: a magnetic resonance imaging study in young Finnish male conscripts. *Mil Med* 2005;170:926-30.
21. Freemont AJ, Peacock TE, Goupille P, et al. Nerve ingrowth into diseased intervertebral disc in chronic back pain. *Lancet* 1997;350:178-81.
22. Varlotta GP, Brown MD, Kelsey JL, et al. Familial predisposition for herniation of a lumbar disc in patients who are less than twenty-one years old. *J Bone Joint Surg Am* 1991;73:124-8.
23. Frino J, McCarthy RE, Sparks CY, et al. Trends in adolescent lumbar disk herniation. *J Pediatr Orthop* 2006;26:579-81.
24. Matsui H, Terahata N, Tsuji H, et al. Familial predisposition and clustering for juvenile lumbar disc herniation. *Spine (Phila Pa 1976)* 1992;17:1323-8.
25. Postacchini F, Lami R, Pugliese O. Familial predisposition to discogenic low-back pain. An epidemiologic and immunogenetic study. *Spine (Phila Pa 1976)* 1988;13:1403-6.
26. Simmons ED, Jr., Guntupalli M, Kowalski JM, et al. Familial predisposition for degenerative disc disease. A case-control study. *Spine (Phila Pa 1976)* 1996;21:1527-9.
27. Saftic R, Grgic M, Ebling B, et al. Case-control study of risk factors for lumbar intervertebral disc herniation in Croatian island populations. *Croat Med J* 2006;47:593-600.
28. Richardson JK, Chung T, Schultz JS, et al. A familial predisposition toward lumbar disc injury. *Spine (Phila Pa 1976)* 1997;22:1487-92; discussion 93.

29. Matsui H, Kanamori M, Ishihara H, et al. Familial predisposition for lumbar degenerative disc disease. A case-control study. *Spine (Phila Pa 1976)* 1998;23:1029-34.
30. Battie MC, Videman T, Gibbons LE, et al. 1995 Volvo Award in clinical sciences. Determinants of lumbar disc degeneration. A study relating lifetime exposures and magnetic resonance imaging findings in identical twins. *Spine (Phila Pa 1976)* 1995;20:2601-12.
31. Sambrook PN, MacGregor AJ, Spector TD. Genetic influences on cervical and lumbar disc degeneration: a magnetic resonance imaging study in twins. *Arthritis Rheum* 1999;42:366-72.
32. Livshits G, Cohen Z, Higla O, et al. Familial history, age and smoking are important risk factors for disc degeneration disease in Arabic pedigrees. *Eur J Epidemiol* 2001;17:643-51.
33. Battie MC, Haynor DR, Fisher LD, et al. Similarities in degenerative findings on magnetic resonance images of the lumbar spines of identical twins. *J Bone Joint Surg Am* 1995;77:1662-70.
34. de Schepper EI, Damen J, van Meurs JB, et al. The association between lumbar disc degeneration and low back pain: the influence of age, gender, and individual radiographic features. *Spine (Phila Pa 1976)* 2010;35:531-6.
35. Molinari RW, Bridwell KH, Lenke LG, et al. Complications in the surgical treatment of pediatric high-grade, isthmic dysplastic spondylolisthesis. A comparison of three surgical approaches. *Spine (Phila Pa 1976)* 1999;24:1701-11.
36. Pfirrmann CW, Metzdorf A, Zanetti M, et al. Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine* 2001;26:1873-8.
37. Schneiderman G, Flannigan B, Kingston S, et al. Magnetic resonance imaging in the diagnosis of disc degeneration: correlation with discography. *Spine* 1987;12:276-81.
38. Tilson ER, Strickland GD, Gibson SD. An overview of radiography, computed tomography, and magnetic resonance imaging in the diagnosis of lumbar spine pathology. *Orthop Nurs* 2006;25:415-20; quiz 21-2.
39. Carragee EJ, Don AS, Hurwitz EL, et al. 2009 ISSLS Prize Winner: Does discography cause accelerated progression of degeneration changes in the lumbar disc: a ten-year matched cohort study. *Spine (Phila Pa 1976)* 2009;34:2338-45.
40. Fourney DR, Andersson GBJ, Arnold PM, et al. Chronic low back pain: a heterogeneous condition with challenges for an evidence-based approach. *Spine* (In Press).
41. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409:860-921.

42. Frazer KA, Ballinger DG, Cox DR, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007;449:851-61.
43. International_HapMap_Consortium. The International HapMap Project. *Nature* 2003;426:789-96.
44. The_1000_Genomes_Project_Consortium, Durbin RM, Abecasis GR, et al. A map of human genome variation from population-scale sequencing. *Nature* 2010;467:1061-73.
45. The_1000_Genomes_Project_Consortium. About the 1000 Genomes Project, 2010. Available from: <http://www.1000genomes.org/about>.
46. Ragoussis J. Genotyping technologies for genetic research. *Annu Rev Genomics Hum Genet* 2009;10:117-33.
47. Magi R, Pfeufer A, Nelis M, et al. Evaluating the performance of commercial whole-genome marker sets for capturing common genetic variation. *BMC Genomics* 2007;8:159.
48. Fakhrai-Rad H, Pourmand N, Ronaghi M. Pyrosequencing: an accurate detection platform for single nucleotide polymorphisms. *Hum Mutat* 2002;19:479-85.
49. Cardon LR, Bell JI. Association study designs for complex diseases. *Nat Rev Genet* 2001;2:91-9.
50. Zhang Y, Sun Z, Liu J, et al. Advances in susceptibility genetics of intervertebral degenerative disc disease. *Int J Biol Sci* 2008;4:283-90.
51. Kalichman L, Hunter DJ. The genetics of intervertebral disc degeneration. Associated genes. *Joint Bone Spine* 2008;75:388-96.
52. Chan D, Song Y, Sham P, et al. Genetics of disc degeneration. *Eur Spine J* 2006;15 Suppl 3:S317-25.
53. Ala-Kokko L. Genetic risk factors for lumbar disc disease. *Ann Med* 2002;34:42-7.
54. Jim JJ, Noponen-Hietala N, Cheung KM, et al. The TRP2 allele of COL9A2 is an age-dependent risk factor for the development and severity of intervertebral disc degeneration. *Spine (Phila Pa 1976)* 2005;30:2735-42.
55. Paassilta P, Lohiniva J, Goring HH, et al. Identification of a novel common genetic risk factor for lumbar disk disease. *JAMA* 2001;285:1843-9.
56. Annunen S, Paassilta P, Lohiniva J, et al. An allele of COL9A2 associated with intervertebral disc disease. *Science* 1999;285:409-12.
57. Roughley P, Martens D, Rantakokko J, et al. The involvement of aggrecan polymorphism in degeneration of human intervertebral disc and articular cartilage. *Eur Cell Mater* 2006;11:1-7; discussion
58. Kawaguchi Y, Osada R, Kanamori M, et al. Association between an aggrecan gene polymorphism and lumbar disc degeneration. *Spine (Phila Pa 1976)* 1999;24:2456-60.

59. Dong DM, Yao M, Liu B, et al. Association between the -1306C/T polymorphism of matrix metalloproteinase-2 gene and lumbar disc disease in Chinese young adults. *Eur Spine J* 2007;16:1958-61.
60. Takahashi M, Haro H, Wakabayashi Y, et al. The association of degeneration of the intervertebral disc with 5a/6a polymorphism in the promoter of the human matrix metalloproteinase-3 gene. *J Bone Joint Surg Br* 2001;83:491-5.
61. Goupille P, Jayson MI, Valat JP, et al. Matrix metalloproteinases: the clue to intervertebral disc degeneration? *Spine (Phila Pa 1976)* 1998;23:1612-26.
62. Le Maitre CL, Freemont AJ, Hoyland JA. The role of interleukin-1 in the pathogenesis of human intervertebral disc degeneration. *Arthritis Res Ther* 2005;7:R732-45.
63. Solovieva S, Leino-Arjas P, Saarela J, et al. Possible association of interleukin 1 gene locus polymorphisms with low back pain. *Pain* 2004;109:8-19.
64. Solovieva S, Kouhia S, Leino-Arjas P, et al. Interleukin 1 polymorphisms and intervertebral disc degeneration. *Epidemiology* 2004;15:626-33.
65. 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;82:744-7.
66. Torres B, Orozco G, Garcia-Lozano JR, et al. Asporin repeat polymorphism in rheumatoid arthritis. *Ann Rheum Dis* 2007;66:118-20.
67. Kizawa H, Kou I, Iida A, et al. An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis. *Nat Genet* 2005;37:138-44.
68. Ikegawa S, Kawamura S, Takahashi A, et al. Replication of association of the D-repeat polymorphism in asporin with osteoarthritis. *Arthritis Res Ther* 2006;8:403; author reply
69. Tilkeridis C, Bei T, Garantziotis S, et al. Association of a COL1A1 polymorphism with lumbar disc disease in young military recruits. *J Med Genet* 2005;42:e44.
70. Kawaguchi Y, Kanamori M, Ishihara H, et al. The association of lumbar disc disease with vitamin-D receptor gene polymorphism. *J Bone Joint Surg Am* 2002;84-A:2022-8.
71. Videman T, Gibbons LE, Battie MC, et al. The relative roles of intragenic polymorphisms of the vitamin d receptor gene in lumbar spine degeneration and bone density. *Spine (Phila Pa 1976)* 2001;26:E7-E12.
72. 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;23:2477-85.
73. Pluijm SM, van Essen HW, Bravenboer N, et al. Collagen type I alpha1 Sp1 polymorphism, osteoporosis, and intervertebral disc degeneration in older men and women. *Ann Rheum Dis* 2004;63:71-7.

74. Grant SF, Reid DM, Blake G, et al. Reduced bone density and osteoporosis associated with a polymorphic Sp1 binding site in the collagen type I alpha 1 gene. *Nat Genet* 1996;14:203-5.
75. Uitterlinden AG, Burger H, Huang Q, et al. Relation of alleles of the collagen type I alpha 1 gene to bone density and the risk of osteoporotic fractures in postmenopausal women. *N Engl J Med* 1998;338:1016-21.
76. Riggs BL. Vitamin D-receptor genotypes and bone density. *N Engl J Med* 1997;337:125-6.
77. Morrison NA, Qi JC, Tokita A, et al. Prediction of bone density from vitamin D receptor alleles. *Nature* 1994;367:284-7.
78. Seki S, Kawaguchi Y, Chiba K, et al. A functional SNP in CILP, encoding cartilage intermediate layer protein, is associated with susceptibility to lumbar disc disease. *Nat Genet* 2005;37:607-12.
79. Lorenzo P, Bayliss MT, Heinegard D. A novel cartilage protein (CILP) present in the mid-zone of human articular cartilage increases with age. *J Biol Chem* 1998;273:23463-8.
80. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005;6:95-108.

Table 1: Advantages and disadvantages of various genetic study methodologies

Table was adopted with modification from ⁸⁰. ✓ here means the methodology has an advantage in this aspect, and the opposite for ✗.

| | Association (candidate gene approach) | Association (genome-wide approach) | Linkage |
|---|--|---|----------------|
| No knowledge of gene functions required | ✗ | ✓ | ✓ |
| Localization to small genomic region | ✓ | ✓ | ✗ |
| Cost | ✓ | ✗ | ✓ |
| Families not required | ✓ | ✓ | ✗ |
| Not easily affected by population stratification | ✗ | ✗ | ✓ |
| Power to detect common alleles (MAFs>5%) of modest effect | ✓ | ✓ | ✗ |
| Power to detect rare alleles (MAFs<1%) | ✗ | ✗ | ✓ |

