

THE INVOLVEMENT OF AGGREGAN POLYMORPHISM IN DEGENERATION OF HUMAN INTERVERTEBRAL DISC AND ARTICULAR CARTILAGE

P. Roughley^{1*}, D. Martens², J. Rantakokko², M. Alini², F. Mwale³ and J. Antoniou³

¹ Genetics Unit, Shriners Hospital for Children and Department of Surgery, McGill University, Montreal, Quebec, Canada

² Orthopaedic Research Laboratory, Royal Victoria Hospital, Montreal, Quebec, Canada

³ Lady Davis Institute, Jewish General Hospital and Department of Surgery, McGill University, Montreal, Quebec, Canada

Abstract

The functions of the intervertebral disc and of articular cartilage are intimately related to their aggrecan content. Aggrecan is a proteoglycan that interacts with hyaluronan to form large aggregates, which are responsible for the ability of the tissues to resist compressive loads. This function is related to the structure of aggrecan, and in particular to the large number of chondroitin sulphate chains present on its core protein. The chondroitin sulphate chains are present in two adjacent regions of the aggrecan core protein, termed the CS1 and CS2 domains. In the human, the region of the aggrecan gene encoding the CS1 domain exhibits size polymorphism, which can result in variation in the degree of chondroitin sulphate substitution of aggrecan in different individuals. This raises the possibility that the functional properties of aggrecan may vary between individuals, and that those individuals with an inferior aggrecan structure may be more susceptible to premature intervertebral disc or articular cartilage degeneration. Several studies have been performed to demonstrate such an association, but the results have been ambiguous. This review explains the relationship between aggrecan structure and function, describes the technique used to assess aggrecan polymorphism and the conclusions and limitations of the data obtained to date, and discusses the implications for tissue degeneration and clinical practice.

Key words: Aggrecan, polymorphism, intervertebral disc, disc degeneration, articular cartilage, osteoarthritis.

*Address for correspondence:

P.J. Roughley
Genetics Unit,
Shriners Hospital for Children,
1529 Cedar Avenue,
Montreal, Quebec,
H3G 1A6, Canada

Telephone number: 1-514-282-7156

FAX number: 1-514-842-5581

E-mail: proughley@shriners.mcgill.ca

Introduction

Structure of aggrecan

Aggrecan, a large aggregating proteoglycan, is one of the major structural components of intervertebral disc (Roughley *et al.*, 2002) and cartilage (Watanabe *et al.*, 1998). It has a long core protein with several different structural and functional regions (Doege *et al.*, 1991) (Fig. 1). At the N-terminus of the core protein, there are two globular regions (G1 and G2) separated by a short interglobular domain (IGD). These are followed by a long extended region, the glycosaminoglycan (GAG) attachment region, to which numerous keratan sulphate (KS) and chondroitin sulphate (CS) chains are attached. At the C-terminus of the core protein there is a third globular region (G3). The GAG-attachment region is divided into three adjacent domains, responsible for the attachment of KS (KS domain) or CS (CS1 and CS2 domains). The human aggrecan gene is located on chromosome 15q26 (Korenberg *et al.*, 1993), and is composed of 19 exons ranging in size from 77 to 4224 bp (Valhmu *et al.*, 1995). Most of the GAG attachment region is encoded by exon 12. In the extracellular matrix of cartilage, aggrecan exists in the form of proteoglycan aggregates which are composed of many aggrecan molecules that interact non-covalently with a central filament of hyaluronan, with each interaction being stabilized by association with a link protein (Hascall, 1988). Aggrecan structure varies with age in both intervertebral disc and articular cartilage, with the size of the KS chains increasing and that of the CS chains decreasing during juvenile development. As there is no template for KS or CS chain length, these parameters may differ between aggrecan present in the intervertebral disc and cartilage of the same individual. It is not clear at present whether these changes in GAG size are of any functional consequence.

Function of aggrecan

Aggrecan provides intervertebral disc and cartilage with the ability to resist compressive loads. The localized high concentrations of aggrecan provide the osmotic properties necessary for normal tissue function (Urban *et al.*, 2000), with the GAGs producing the swelling pressure that counters compressive loads on the tissue. This functional ability is dependent on a high GAG/aggrecan concentration being present in the tissue extracellular matrix. The formation of large proteoglycan aggregates is essential to restrict the movement of aggrecan in the

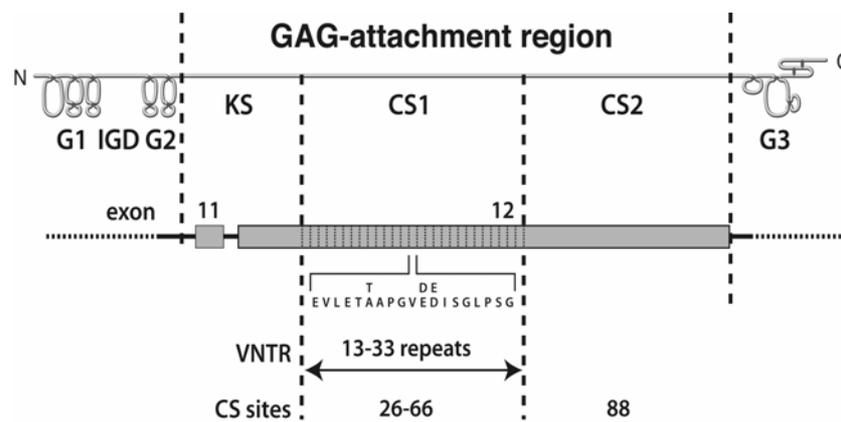


Figure 1. Aggrecan structure and polymorphism. The core protein of aggrecan is depicted in an amino (N) to carboxy (C) terminal orientation, with the location of the globular domains (G1-3), the interglobular domain (IGD) and the glycosaminoglycan (GAG)-attachment region being indicated. The GAG-attachment region is divided into the keratan sulphate-rich domain (KS) and two chondroitin sulphate-rich domains (CS1 and CS2). Beneath is indicated the exon organization of that portion of the aggrecan gene encoding the GAG-attachment region. The variable number of tandem repeat (VNTR) polymorphism in the portion of exon 12 encoding the CS1 domain is indicated, together with the 19 amino acid sequence of each repeat. The number of consensus sites for the attachment of CS chains in the CS1 and CS2 domains is also indicated.

extracellular matrix of cartilage, and so prevent diffusion away from the site under compression. Proteoglycan size does not seem to be as critical a factor in retention within the intervertebral disc, where the unique fibrocartilaginous organization permits the accumulation of proteolytic degradation products of aggrecan irrespective of hyaluronan interaction. In articular cartilage such degradation products would be lost into the synovial fluid. The extent of aggrecan degradation can vary between intervertebral disc and cartilage from the same individual, with disc aggrecan commonly having undergone more extensive degradation in the adult. Any factor that results in a decrease in aggrecan/GAG content in the extracellular matrix may result in over-compression of the tissue under load. This in turn may result in an adverse response by the disc cells or chondrocytes, with increased proteinase secretion and subsequent tissue degeneration (Walsh and Lotz, 2004).

CS1 polymorphism

The CS1 domain of the aggrecan core protein is composed of multiple repeats of 19 amino acids. Each repeat contains potential attachment sites for two CS chains. The human aggrecan gene is unique amongst species analyzed to date in possessing variable number of tandem repeat (VNTR) polymorphism in the part of exon 12 encoding the CS1 domain (Doerge *et al.*, 1997) (Fig. 1). Alleles have been identified with CS1 repeat numbers ranging from 13 to 33. The most common alleles possess 26, 27 or 28 repeats (Fig. 2). It is readily apparent that individuals possessing the longer alleles have the potential to produce aggrecan bearing a greater number of CS chains. CS substitution might vary by as many as 40 CS chains per aggrecan core protein between the products of the shortest and largest aggrecan alleles.

Polymorphism and tissue degeneration

Individuals possessing the shortest aggrecan alleles will have the lowest number of CS chains on their aggrecan molecules, and this may result in impaired aggrecan function. This could lead to a lower ability to withstand compression and an increased susceptibility to tissue degeneration. The consequences of CS1 polymorphism on the CS content of aggrecan will be most pronounced in adults due to the cumulative effects of proteolytic processing of the core protein (Roughley *et al.*, 2002). Such processing is thought to be due to the action of metalloproteinases, particularly the aggrecanases (Tortorella *et al.*, 2001), and results in removal of part or

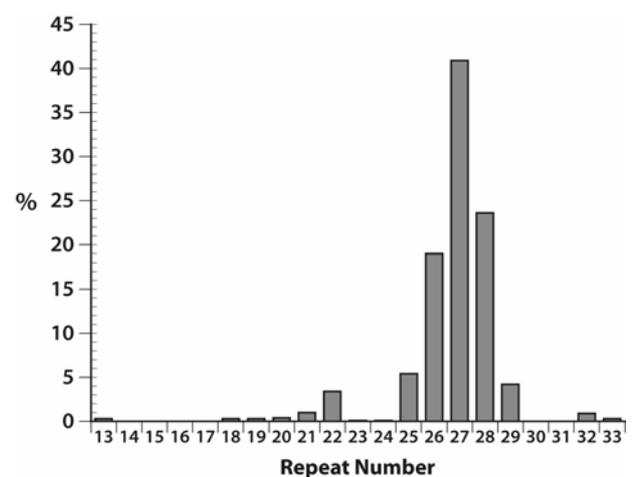


Figure 2. Aggrecan allele distribution. The figure is a histogram depicting the percentage distribution within the population of the different repeat lengths that are present within the CS1-encoding region of the human aggrecan gene. The information is an accumulation of all the published and unpublished data alluded to in the text.

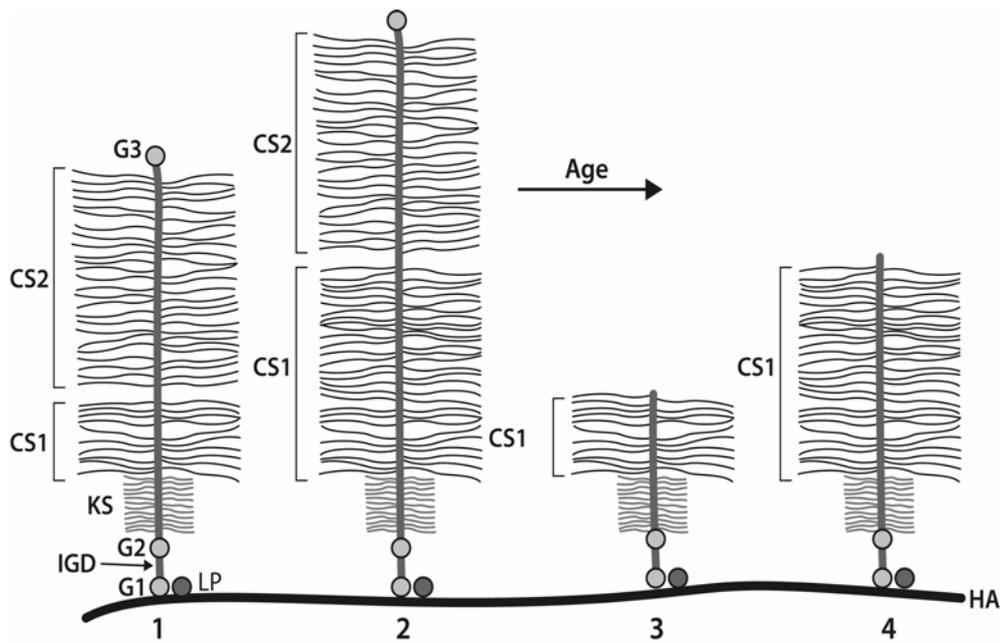


Figure 3. Aggrecan core protein processing with age. Aggrecan molecules derived from individuals possessing low (1 and 3) and high (2 and 4) repeat numbers in the CS1-domain are depicted in association with hyaluronan (HA) and a link protein (LP). The various domains of the aggrecan molecule described in Fig. 1 are also indicated. The left-hand side of the figure depicts intact aggrecan molecules (1 and 2), whereas the right-hand side depicts aggrecan molecules that have been subjected to proteolytic processing with age (3 and 4) and are devoid of their CS2 and G3 domains.

most of the CS2 region (Fig. 3). This magnifies the contribution of the CS1 domain to aggrecan function and is thought to contribute to increased tissue degeneration in the adult with age. If one accepts this scenario, it is not unreasonable to postulate that individuals with the shortest allele sizes may be at greatest risk for early age-related intervertebral disc and articular cartilage degeneration. At this time, however, the data supporting such an association are difficult to interpret, and it is therefore an opportune time to review the potential value and limitations associated with the use of aggrecan CS1 polymorphism as an indicator of intervertebral disc and cartilage degeneration.

Analysis of CS1 polymorphism

The aggrecan CS1 polymorphism is readily analyzed by polymerase chain reaction (PCR) amplification of genomic DNA and subsequent agarose gel electrophoresis.

Source of DNA

DNA can be obtained from any source of cells, but for convenience it is usually taken from fresh anticoagulated blood or buffy coat. DNA is conveniently isolated using a Qiagen DNA purification kit, and 200 μ l fresh blood usually yields about 10-20 μ g DNA.

PCR amplification

The CS1 domain of the aggrecan gene can be PCR amplified using the flanking oligonucleotide primers described by Doege *et al.* (1997):

Sense primer: 5'-TAGAGGGCTCTGCCTCTGGAGTTG

Antisense primer: 5'-AGGTCCCCTACCGCAGAGGTAGAA
Product sizes range between 700 to 2000 bp and it is therefore necessary to utilize a system that is compatible with the generation of large amplicons, such as the Expand Long Template System (Roche). A typical 50 μ l reaction mixture consists of 0.5 μ l genomic DNA (50 ng), 5 μ l Expand Long Template Buffer 2, 2.5 μ l dNTP mix (10 mmol each of dTTP, dCTP, dGTP and dATP), 1 μ l sense primer (15 pmol/ μ l), 1 μ l antisense primer (15 pmol/ μ l), and 0.75 μ l Taq DNA polymerase. Amplification is carried out for 30 cycles at an annealing temperature of 66°C followed by 2 min extension at 68°C.

Agarose gel electrophoresis

The PCR products can be analyzed for their size by electrophoresis on a 0.8% agarose gel, containing 0.25 μ g/ml ethidium bromide for visualization (Fig. 4). For each sample, 35 μ l PCR product and 7 μ l loading buffer (15% Ficoll, 0.25% Bromophenol Blue) are loaded on the gel, together with 5 μ l 100bp DNA ladder (Gibco). The electrophoresis is performed at 90 V for 2 hours. The number of repeats in the CS1 region is determined by reference to a table of repeat number versus amplicon length (Fig. 5).

Description of current data

The analysis of CS1 polymorphism has been used by several groups of investigators to study correlation between allele size and degenerative changes in the intervertebral disc or osteoarthritic changes in cartilage.

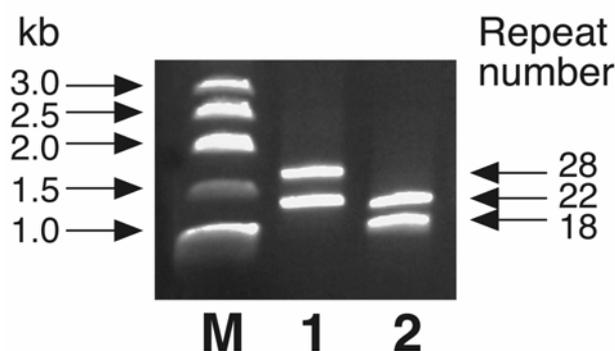


Figure 4. Agarose gel showing alleles of different sizes. The figure illustrates an ethidium bromide-stained gel showing the electrophoretic migration of molecular size markers (M) and the PCR products derived from the CS1-encoding region of the aggrecan genes present in two individuals (1 and 2). In each individual the two aggrecan genes possess CS1 domains of different size. The repeat numbers in the CS1 domains are 22 and 28 for individual 1 and 18 and 22 for individual 2.

Association with osteoarthritis

The first report of an association between the human aggrecan CS1 polymorphism and articular cartilage degeneration was in relation to bilateral osteoarthritis (OA) of the hand (Horton *et al.*, 1998). The study group consisted of 93 Caucasian men aged 60 years and above (mean age 72 years) with hand and/or knee OA. Bilateral hand OA and bilateral knee OA were present in 43 and 28 of the men, respectively, with the other individuals having either unilateral OA or no OA in their hands or knees. The allele frequency in the study group was of a similar distribution to the general population: allele 33 (0.5%), 29 (2.2%), 28 (31.7%), 27 (43.0%), 26 (16.7%), 25 (3.2%), 22 (2.2%) and 19 (0.5%). The presence of allele 27 was associated with bilateral hand OA with an odds ratio (OR) of 3.23, but there was no statistically significant association between allele 27 and bilateral knee OA. None of the other alleles showed an association with either bilateral hand or knee OA. It was concluded that this finding supports the concept that genetic factors may play a role in the development and/or progression of some forms of age-onset OA. In addition, it was suggested that a heterozygous genotype in which one copy of allele 27 was present was sufficient for the link to OA. However, it is difficult to explain why the allele with 27 repeats should be related to disease and not those with 26 or 28 repeats, where aggrecan function should be similar. One could hypothesize that alleles of the same length could have arisen independently, giving rise to base variations in the VNTR region. This could then result in a subset of individuals with 27 repeats who possess a base substitution that is linked to disease. Why such variation should not also occur with other allele sizes and why it should affect cartilage degeneration in the hand but not the knee is not clear. The inability to detect a significant association between OA and the smaller allele sizes could also relate to their low abundance in the

Repeat number	Amplicon size (bp)
13	775
14	832
15	889
16	946
17	1003
18	1060
19	1117
20	1174
21	1231
22	1288
23	1345
24	1402
25	1459
26	1516
27	1573
28	1630
29	1687
30	1744
31	1801
32	1858
33	1915

Figure 5. Allele repeat number compared to size of PCR product. The figure depicts a comparison of the various repeat numbers found in the human aggrecan CS1 domain and the corresponding amplicon size obtained upon PCR amplification. The table permits the amplicon sizes established by agarose gel electrophoresis (Fig. 4) to be converted to repeat number.

population and presence in the study group in only a few individuals. Much larger study groups may be needed to determine whether such individuals are at greater risk than their counterparts with longer alleles.

A second study consisted of 134 Australian twins aged over 50 years (including 34 monozygotic and 27 dizygotic twin pairs) who had OA of the hands, hips and/or knees (Kirk *et al.*, 2003). The allele frequency in the study group was also of a similar distribution to the general population: allele 29 (0.3%), 28 (23.6%), 27 (43.8%), 26 (20.5%), 25 (7.5%), 23 (0.3%), 22 (2.7%) and 19 (0.3%). Only alleles with 25, 26, 27 and 28 repeats were sufficiently represented for statistical analysis. Multivariate analysis was used to compare the presence of alleles with 25, 27 and 28 repeats to the allele with 26 repeats for increased risk of OA. It was reported that the alleles with 25 and 28 repeats were associated with a lower prevalence of knee and hip OA, respectively. The results were interpreted as being consistent with a protective effect of the alleles with 25 or 28 repeats in some forms of OA, in contrast to the data of Horton *et al.* (1998) where a deleterious effect of the allele with 27 repeats was suggested. It was concluded that this provides some additional evidence for an association

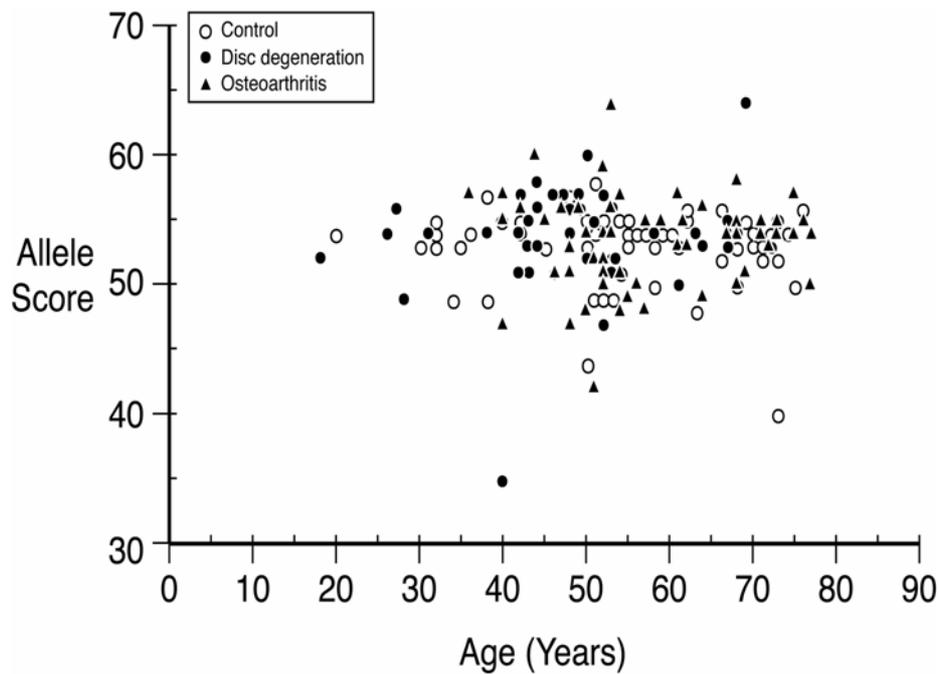


Figure 6. Allele score plotted against age for different patient groups. The figure plots allele score versus age for a group of individuals attending orthopaedic clinics for primary osteoarthritis or degenerative disc disease and compares them with a group of control individuals who show no evidence of either disorder. Allele score is defined as the sum of the CS1 repeat number found in the two aggrecan gene alleles of each individual. The individuals depicted in Fig. 4 would have allele scores of 50 and 40. In agreement with Fig. 2, the most common allele scores range between 50 and 58.

between the aggrecan CS1 polymorphism and OA. It is, however, difficult to explain why alleles with 25 or 28 repeats should be functionally different from those with 26 or 27 repeats, or why different allele sizes should be associated with different forms of OA. It is possible that evidence for an association between the aggrecan CS1 polymorphism and OA, as theoretically predicted, may not have been apparent as the short alleles with less than 25 repeats were under-represented in the study group.

Association with intervertebral disc degeneration

The first report of a correlation between aggrecan polymorphism and intervertebral disc degeneration involved a study group of 64 young women with a mean age of 21.3 years (Kawaguchi *et al.*, 1999). MRI examination showed that 32 participants had normal lumbar intervertebral discs and that 32 had disc degeneration. Among the latter group, 3 individuals had multilevel disc degeneration. Once again the allele with 27 repeats was found to be most abundant, and a normal distribution of allele sizes was observed: allele 29 (3.9%), 28 (9.4%), 27 (46.9%), 26 (27.3%), 25 (6.2%), 22 (3.1%), 21 (2.3%) and 18 (0.8%). 25 individuals were homozygous for allele 27, but none were homozygous for the short alleles. A significant difference was reported between the distribution of allele sizes and the severity of disc degeneration on MRI. As the presence of alleles with 18 or 21 repeats was over-represented in individuals with multilevel or severe disc degeneration, it was concluded that individuals with one short CS1 domain have an increased risk of disc degeneration at an early age. While

this conclusion is in agreement with theoretical predictions, the small number of individuals with the shorter alleles in the study group raises the question of whether there was a large enough representation in this subset to be able to extrapolate the conclusions to the general population.

Lack of association with osteoarthritis and degenerative disc disease

This study by the authors involved 165 individuals (both male and females, age range 18-77 years). The individuals were divided into 3 groups: 63 with primary hip osteoarthritis, 44 with degenerative disc disease, and 58 control subjects without history of either pathology. The individuals with degenerative disc disease presented to our spine clinic for lower back pain and were candidates for surgical decompression. Their average age was 55 years. (Range 18-69 years). Individuals with osteoarthritis presented to our hip arthritis clinic and were candidates for joint replacement. Their average age was 60 years (range 36-77 years). Age matched controls with no history of lower back pain or hip arthritis were healthy hospital volunteers. Their average age was 57 years (range 20-76 years). The diagnoses of osteoarthritis or degenerative disc disease were confirmed by X-ray or MRI, respectively, and none of the patients had evidence of inflammatory, traumatic, congenital or vascular etiologies associated with their disease. Upon CS1 polymorphism analysis, 13 different allele sizes were detected with the allele with 27 repeats being most abundant: allele 33 (0.3%), 32 (1.8%), 29 (8.6%), 28 (24.3%), 27 (33.8%), 26 (16.5%), 25 (6.0%), 24 (0.3%), 22 (4.5%), 21 (2.1%), 20 (1.0%), 18 (0.3%)

and 13 (0.5%). Statistical analysis did not reveal a significant difference in the allele size distribution between the disease groups and the control group. Of particular note, there was no evidence of a skewed distribution toward the lower range of allele sizes in either of the disease groups (Fig. 6). Thus, no correlation was found between aggrecan CS1 polymorphism and degenerative disc disease or primary osteoarthritis. Therefore, in this study population it could be postulated that the presence of a short CS1 domain does not predispose to these disorders and the presence of a long allele offers no protection from cartilage or disc degeneration. However, one must accept that this study has its limitations, as there may be many factors that could influence predisposition to degenerative disc disease or osteoarthritis in addition to aggrecan gene polymorphism, and the irregular age distribution of the groups makes an objective comparison difficult. As in all the above studies, there was also an under-representation of short alleles in the study groups, and individuals with two short alleles were only rarely observed. It may be these latter individuals who are at greatest risk, but who have not been the focus of any study to date.

Discussion

Based on the currently available data it is unclear whether or not there is a direct link between aggrecan CS1 length and either intervertebral disc degeneration or osteoarthritis. In the case of the intervertebral disc, an association between disc degeneration and the presence of shorter CS1 lengths has been reported, but remains to be confirmed in larger study groups. Where a link has been suggested with cartilage degeneration, it is not the individuals bearing the shorter CS1 domains who appear to be at risk, and an association of tissue degeneration with the larger CS1 lengths is difficult to reconcile on our current understanding of aggrecan function. It is possible that a lack of association with alleles of shorter length may be due to their under-representation in the study groups. Alternatively, one must accept that the reported correlations are correct, and that the association of aggrecan polymorphism with cartilage degeneration is independent of CS1 length, but is due to another feature of variable CS1 structure such as differences in glycosylation or susceptibility to proteolysis. While no evidence for such scenarios currently exists, they cannot be categorically excluded.

If one accepts that there is an inverse association between aggrecan CS1 length and intervertebral disc or cartilage degeneration, then it is important to understand what the implications of such an association may be. The aggrecan CS1 polymorphism does not indicate who will get intervertebral disc degeneration or osteoarthritis, but rather it indicates the possibility of an individual developing tissue degeneration at an earlier age than would otherwise have occurred. Such an age of onset linkage implies that individuals who develop intervertebral disc or cartilage degeneration at an early age, in the absence of any overt gene mutation, may be most likely to possess aggrecan with the shorter CS1 domains. Even though the

CS1 polymorphism is present in all intervertebral discs and cartilages, this does not necessarily imply a systemic problem with degeneration. The factor initiating tissue degeneration is likely to be independent of the aggrecan polymorphism and could be joint selective. For example, it could be biomechanical with the consequence on every joint being different depending on the load to which it is subjected. There is no reason why premature intervertebral disc and cartilage degeneration need necessarily occur together in the same individual, as the parameters regulating these events are different in the two tissues.

Another important consideration is whether the presence of one allele with a short CS1-encoding domain will be sufficient to influence cartilage or disc function in a detrimental manner. There is no definitive answer to this question, but one could certainly hypothesize that the presence of two short alleles would be more detrimental and may be needed to perturb tissue function. Certainly, in animals bearing a detrimental mutation in their aggrecan gene, a severe chondrodystrophic phenotype only results in the homozygous state, whereas the heterozygous animals are relatively normal and show only a mild phenotype (Primorac *et al.*, 1994; Watanabe *et al.*, 1997). If one arbitrarily defines a functionally impaired CS1 region as having less than 23 repeats (the mid-point in the range of CS1 repeats in the population), then only about 6% of individuals possess such an allele, and only about 0.3% will have two short alleles. It is then not surprising that linkage of aggrecan polymorphism with tissue degeneration has proved elusive to date, as there may only be one or two individuals with two short alleles in study groups of 100-200 individuals. However, if such individuals are at risk they still represent a significant proportion of the population (about 100,000 people in a population of 30 million) and hence could have considerable impact on the health care system. Future studies aimed at conclusively establishing whether aggrecan polymorphism is directly linked to intervertebral disc or cartilage degeneration would need to focus on these individuals, or on groups where tissue degeneration has occurred much earlier than expected in the normal population.

Clinical relevance

There are important clinical implications in whether or not the aggrecan polymorphism, by any mechanism, is linked with intervertebral disc or cartilage degeneration. If a linkage is present, then aggrecan allele analysis early in life can be used to screen individuals at risk, so that they can be advised on how their future lifestyle may impact on tissue degeneration. A linkage will also necessitate that donors of allogenic tissue or cells used for intervertebral disc or cartilage repair be screened for the polymorphism, as there seems little point in repairing a lesion with a tissue that is itself at risk for degeneration. It is therefore likely that the aggrecan CS1 polymorphism will continue to be a topic of study in the orthopaedic field until its relevance is fully resolved.

Acknowledgements

This study was supported by research grants from the Canadian Institutes of Health Research and the Shriners of North America.

References

- Doerge KJ, Coulter SN, Meek LM, Maslen K, Wood JG (1997) A human-specific polymorphism in the coding region of the aggrecan gene. Variable number of tandem repeats produce a range of core protein sizes in the general population. *J Biol Chem* **272**: 13974-13979.
- Doerge KJ, Sasaki M, Kimura T, Yamada Y (1991) Complete coding sequence and deduced primary structure of the human cartilage large aggregating proteoglycan, aggrecan. Human-specific repeats, and additional alternatively spliced forms. *J Biol Chem* **266**: 894-902.
- Hascall VC (1988) Proteoglycans: the chondroitin sulfate/keratan sulfate proteoglycan of cartilage. *ISI Atlas Sci Biochem* **1**: 189-198.
- Horton WE, Jr., Lethbridge-Çejku M, Hochberg MC, Balakir R, Precht P, Plato CC, Tobin JD, Meek L, Doerge K (1998) An association between an aggrecan polymorphic allele and bilateral hand osteoarthritis in elderly white men. Data from the Baltimore Longitudinal Study of Aging (BLSA). *Osteoarthritis Cart* **6**: 245-251.
- Kawaguchi Y, Osada R, Kanamori M, Ishihara H, Ohmori K, Matsui H, Kimura T (1999) Association between an aggrecan gene polymorphism and lumbar disc degeneration. *Spine* **24**: 2456-2460.
- Kirk KM, Doerge KJ, Hecht JT, Bellamy N, Martin NG (2003) Osteoarthritis of the hands, hips and knees in an Australian twin sample. Evidence of association with the aggrecan VNTR polymorphism. *Twin Res* **6**: 62-66.
- Korenberg JR, Chen XN, Doerge K, Grover J, Roughley PJ (1993) Assignment of the human aggrecan gene (AGC1) to 15q26 using fluorescence in situ hybridization analysis. *Genomics* **16**: 546-548.
- Primorac D, Stover ML, Clark SH, Rowe DW (1994) Molecular basis of nanomelia, a heritable chondrodystrophy of chicken. *Matrix* **14**: 297-305.
- Roughley PJ, Alini M, Antoniou J (2002) The role of proteoglycans in aging, degeneration and repair of the intervertebral disc. *Biochem Soc Trans* **30**: 869-874.
- Tortorella MD, Malfait AM, Deccico C, Arner E (2001) The role of ADAM-TS4 (aggrecanase-1) and ADAM-TS5 (aggrecanase-2) in a model of cartilage degradation. *Osteoarthritis Cart* **9**: 539-552.
- Urban JPG, Roberts S, Ralphs JR (2000) The nucleus of the intervertebral disc from development to degeneration. *Amer Zool* **40**: 53-61.
- Valhmu WB, Palmer GD, Rivers PA, Ebara S, Cheng J-F, Fischer S, Ratcliffe A (1995) Structure of the human aggrecan gene. Exon-intron organization and association with the protein domains. *Biochem J* **309**: 535-542.
- Walsh AJ, Lotz JC (2004) Biological response of the intervertebral disc to dynamic loading. *J Biomech* **37**: 329-337.
- Watanabe H, Nakata K, Kimata K, Nakanishi I, Yamada Y (1997) Dwarfism and age-associated spinal degeneration of heterozygote *cmd* mice defective in aggrecan. *Proc Natl Acad Sci USA* **94**: 6943-6947.
- Watanabe H, Yamada Y, Kimata K (1998) Roles of aggrecan, a large chondroitin sulfate proteoglycan, in cartilage structure and function. *J Biochem (Tokyo)* **124**: 687-693.

Discussion with reviewers

Reviewer 1: The authors argue that the associations found to-date do not have rational explanations. This is particularly true when they consider associations with higher repeats. However, the authors must be careful with their bias in thinking that lower repeats would be more likely to be associated with disease. This view is presumably based on the idea that shorter proteoglycans would perform their structural/shock-absorbing function in an inferior manner. Have the authors considered other possible mechanisms?

Authors: Current knowledge of aggrecan function implies that molecules possessing a lower sulphated glycosaminoglycan content may exhibit impaired function. If one accepts this relationship, then the functional properties of aggrecan would be expected to decrease with a decrease in the number of CS1 repeats. This relationship does, however, assume that CS chain length is independent of CS1 repeat number. If there is an inverse relationship between CS chain length and CS1 repeat number, then the functional properties of the aggrecan molecules may be less dependent on repeat number, and this could account for the lack of correlation between susceptibility to OA and low CS1 repeat number observed by some investigators. In principle, any structural parameter that results in a decrease in the CS content of aggrecan could also have detrimental functional consequences. In relation to the CS1 region, one could postulate that there may be a unique CS1 repeat structure which either impairs CS substitution of the CS1 region, or which renders the CS1 region susceptible to proteolytic cleavage. If such a unique CS1 repeat only occurred in CS1 regions of a particular length, then it could account for literature reports of a relationship between susceptibility to OA and individuals with a specific number of CS1 repeats. However, at present there is no data to support such a scenario.