Proteoglycan genes and anterior cruciate ligament injury susceptibility

Sasha Mannion

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN

In fulfilment of the requirements for the degree

Master of Science in Medicine

Human Genetics

Faculty of Health Science

UNIVERSITY OF CAPE TOWN

2013

Supervisors:

Dr Alison September

Prof Malcolm Collins

Declaration

I, Sasha Mannion, hereby declare that the work on which this dissertation is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this, or any other university.

I authorise the university to reproduce, for the purpose of research, either the whole or any portion of the contents of this dissertation in any manner whatsoever.

Signature:…….

Date:……

Table of Contents

Acknowledgements

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

The South African Medical Research Council (MRC) is acknowledged for their financial contribution to the UCT/MRC Research Unit for Exercise Science and Sports Medicine (ESSM).

Thank you to the University of Cape Town for the financial assistance in the form of the University Research Committee scholarship, which made this research possible.

Participants: Thank you for donating your time and DNA samples for the studies presented in this dissertation.

Dr Alison September (*Lady September*): Thank you for the best post-graduate experience I could have asked for. You have been such a great mentor to me, far surpassing the help with my dissertation, manuscripts and presentations, always giving sound advice beyond my current work. I feel as though your guidance has prepared me for my next chapter and I can't express how greatly I appreciate your constant support over the past few years.

Prof Malcolm Collins (*The Chief*): Thank you for all your input over the past years and always pushing me to deliver my best work. I feel very lucky to have been a part of your team at ESSM that is able to learn and make a significant contribution to research, while having too much fun!

Dr Michael Posthumus: Thank you for all your support and guidance. I much appreciate your assistance with the manuscript and liaising with doctors to help us with recruitment.

Dr W van der Merwe, Dr H Hobbs, Dr D O'Cuinneagain, (The Sports Science Orthopeadic Clinic) Dr P Rouw, Dr I Coller, Dr R Dacks, and Dr R von Bormann: Thank you for your assistance with the recruiting of participants and your willingness to help.

Prof Raj Ramesar, Prof Jacquie Greenberg and A/Prof Collet Dandara: Thank you for allowing me the opportunity to do my MSc Human Genetics while being a part of two different divisions. I have thoroughly enjoyed my experience.

ESSM Research Unit at the Sports Science Institute of South Africa, and Prof Tim Noakes: Thank you for giving me the opportunity to do my Masters at ESSM. I have truly enjoyed being a part of such a diverse group of researchers and a department that is constantly striving for excellence.

Masouda Rahim: This dissertation would not be possible without you! I feel as though we made such a great team, constantly supporting each other over the past three years. Thank you for everything!

Kevin O'Connell and Marelize Burger: Thank you for all your help with absolutely anything and everything. From taking bloods, to help in the lab, stats and writing up, I know I asked a lot of the two of you and you were always there and willing to help. I much appreciate it.

Neezaam and Trevino: Thank you for all your support in the lab and your continual willingness to help.

Asanda Mtintsilana: Thank you for your assistance with the genotyping of the polymorphisms within the *FMOD* and *LUM* genes.

All the '*Genettes*', '*Genette-men*' and my friends at ESSM: Thank you for all the love, support, fun and Vida-runs over the years. I have loved being a part of such a great group of people and truly value all your help and friendship along the way.

To my family: Mom and Dad, thank you for all your support, both emotionally and financially, throughout the years. Thanks for always listening to me stress about work and sitting through presentations, even though you didn't always understand. I know how blessed I am to have had the opportunity to further my studies and I am eternally grateful to you both for that. Chris, thank you for being there for me through the good and the crazy times, in what has been an amazing year. I love you all for the support you have given me.

Scientific Outputs

Publications

Mannion S, Mtintsilana A, Posthumus M, van der Merwe W, Hobbs H, Collins, M, September AV. Genes encoding proteoglycans are associated with the risk of anterior cruciate ligament ruptures. The British Journal of Sports Medicine, 2013 (*in press*).

Rahim M, **Mannion S**, Hobbs H, O'Cuinneagain, van der Merwe W, Collins M, September AV. Investigation of genes involved in the matrix remodelling pathway with risk of anterior cruciate ligament rupture, 2013 (Manuscript in preparation).

Conference Presentations

Mannion S, Mtintsilana A, Posthumus M, van der Merwe W, O'Cuinneagain D, Collins M, September AV. Genes encoding proteoglycans are associated with the risk of anterior cruciate ligament ruptures. The 18th European College of Sports Science annual congress, Catalonia, Barcelona, Spain, 2013 (Conference oral presentation).

Mannion S, Mtintsilana A, Posthumus M, van der Merwe W, O'Cuinneagain D, Collins M, September AV. Genes encoding proteoglycans are associated with the risk of anterior cruciate ligament ruptures. The University of Cape Town Postgraduate Research Day, Cape Town, South Africa, 2013 (Poster presentation).

Rahim M, **Mannion S**, Dandara C, Posthumus M, Collins M, September AV. The apoptosis pathway and risk of anterior cruciate ligament ruptures. South African Society for Human Genetics, Johannesburg, South Africa, 2013 (Poster presentation).

Abbreviations

˚C	degrees Celsius
A	adenine
ACAN	the gene encoding aggrecan
ACL group	participants with surgically diagnosed ACL ruptures
ACL	anterior cruciate ligament
ADAMTS	a disintegrin and metalloproteinase with thrombospondin motifs
AMB	anteromedial bundle
ANOVA	one-way analysis of variance
BGN	the gene encoding biglycan
BMI	body mass index
C	cytosine
CNS	central nervous system
COB	country of birth
COL12A1	the gene encoding the alpha 1 chain of type XII collagen
COL1A1	the gene encoding the alpha 1 chain of type I collagen
COL5A1	the gene encoding the alpha 1 chain of type V collagen
CON group	control participants
CS	chondroitin sulfate
DAMP	damage-associated molecular pattern
DCN	the gene encoding decorin
DNA	deoxyribonucleic acid
DS	dermatan sulfate
ECM	extracellular matrix
EDS	Ehlers–Danlos syndrome
EDTA	ethylenediaminetetraacetic acid
FACIT	fibril associated collagens with interrupted triple helicies
FMOD	the gene encoding fibromodulin
G	guanine
GAG	glycosaminoglycan

GPI	glycosylphosphoinositide
HA	hyaluronic acid
HS	heparan sulfate
HWE	Hardy-Weinberg equilibrium
IL	interleukin
KS	keratan sulfate
LCL	lateral collateral ligament
LD	linkage disequilibrium
LRR	leucine-rich repeat
LUM	the gene encoding lumican
MCL	medial collateral ligament
MGB	minor groove binding
miRNA	micro ribonucleic acid
MMP	gene encoding a matrix metalloproteinase
MMP	matrix metalloproteinase
mRNA	messenger ribonucleic acid
NCBI	National Centre for Biotechnology Information
NFL	National Football League
NON subgroup	ACL injured participants with a non-contact mechanism of injury
OA	osteoarthritis
PCL	posterior cruciate ligament
PCR	polymerase chain reaction
PG4	proteoglycan 4
PLB	posterolateral bundle
PRELP	proline/arginine-rich end leucine-rich repeat protein
PRP	platelet-rich plasma
Q angle	quadriceps angle
RA	rheumatoid arthritis
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
SLRP	small leucine-rich proteoglycan
SNP	single nucleotide polymorphism

STREGA	STrengthening the REporting of Genetic Association studies
STROBE	STrengthening the Reporting of OBservational Studies in Epidemiology
T	thymine
TGF-β	transforming growth factor-β
TIMP	tissue inhibitor metalloproteinase
TLR	toll-like receptor
TNF-α	tumour necrosis factor-α
USA	United States of America
UTR	untranslated region

List of Figures

List of Tables

Abstract

Background: Genetic variants within genes involved in fibrillogenesis, including a polymorphism in the *COL5A1* gene, have previously been associated with anterior cruciate ligament (ACL) injury susceptibility, specifically in females. Proteoglycans have also been implicated in having important functions in fibrillogenesis and maintaining the structural integrity of ligaments. Moreover, their content appears to be lowered in ruptured ACL tissue in comparison to non-ruptured controls. Genes encoding proteoglycans are therefore plausible candidates to be investigated for an association with ACL injury susceptibility. Sequence variants within genes encoding the proteoglycans aggrecan (*ACAN:* rs2351491, rs1042631, rs1516797), biglycan (*BGN:* rs1126499, rs1042103), decorin (*DCN:* rs13312816, rs516115), fibromodulin (*FMOD:* rs7543148, rs10800912) and lumican (*LUM:* rs2268578) were consequently investigated.

Methods: A case-control genetic association study was conducted where 227 participants with surgically diagnosed ACL ruptures (ACL group), including a non-contact mechanism of injury subgroup (NON: 126 participants), and 234 controls without any history of ACL injury were genotyped for ten polymorphisms in the five proteoglycan genes, using standard PCR based methods. Inferred haplotypes were constructed for specific regions encompassing the *ACAN*, *BGN, DCN, FMOD* and *LUM-DCN* genes, and inferred allele constructs were created to investigate gene-gene interactions between the proteoglycans and the previously associated *COL5A1* gene, using the genotype data obtained for the respective loci. Statistical analyses determined whether genotype, allele, haplotype and/or allele construct frequencies differed significantly between cases and controls. Sex-specific associations, as well as associations by mechanism of injury were also investigated.

Results: Independent associations were noted for two of the polymorphisms investigated. The G allele of *ACAN* rs1516797 was significantly under-represented in the controls (p=0.024; OR=0.72; 95% CI:0.55-0.96) compared to the ACL group. For *DCN* rs516115, the GG genotype was significantly over-represented in female controls (p=0.015; OR=9.23; 95% CI:1.16-73.01) compared to the ACL group; and the AA genotype was significantly under-represented in the controls (p=0.013; OR=0.33; 95% CI:0.14-0.78) compared to the female NON subgroup. No further independent associations were identified for the investigated variants.

Inferred haplotype analyses implicated regions overlapping the *ACAN*, *BGN*, and *LUM-DCN* genes in ACL injury susceptibility. The T-C-T and T-C-G *ACAN* inferred haplotypes (rs2351491 C>T - rs1042631

T>C - rs1516797 T>G) were significantly over-represented (p=0.001) and under-represented (p=0.005) in the CON group respectively, in comparison to the ACL group. The *BGN* (rs1126499 C>T - rs1042103 G>A) C-G inferred haplotype was significantly over-represented (p=0.027) in the CON group in comparison to the ACL group for female participants only. The T-A-G *LUM-DCN* inferred haplotype (rs2268578 T>C - rs13312816 A>T - rs516115 A>G) was significantly under-represented (p=0.038) in the CON group in comparison to the ACL group.

Several gene-gene interactions were also noted in this dissertation. Inferred allele constructs identified gene-gene interactions between (i) *BGN* and *DCN,* (ii) *BGN* and *COL5A1*, (iii) *DCN* and *COL5A1,* and (iv) all three of these genes collectively, which may potentially be involved in the modulation of ACL injury susceptibility.

Conclusion: The evidence suggests that *ACAN* rs1516797 and *DCN* rs516115 are independently associated with ACL injury risk. In addition, the inferred haplotype analyses have implicated defined genomic boundaries overlapping the (i) *ACAN*, (ii) *BGN* and (iii) *LUM-DCN* genes which require further interrogation to identify the potential causal genomic motifs associated with ACL injury susceptibility. The potential gene-gene interactions noted between the various proteoglycan genes and *COL5A1* further highlight the complex myriad of interactions between the structural components of the extracellular matrix of ligaments, which may collectively contribute to our understanding of the pathobiology of ACL injury susceptibility. The novel results presented are collectively providing evidence contributing to the current hypothesis that genetic sequence variability within genes, such as the proteoglycan encoding genes, may potentially modulate the ligament fibril properties thereby effecting ACL injury susceptibility.

Introduction and scope of dissertation

The regular participation in physical activity is an important part of maintaining a healthy lifestyle.[1,2] There is however an increased risk of sustaining a musculoskeletal soft tissue injury when participating in either competitive or recreational activities.[3] Musculoskeletal soft tissue injuries collectively refer to acute and overuse injuries of the tendons, ligaments and skeletal muscles, and may occur in either active or sedentary individuals.[4,5] Some of the extensively studied common tendon and ligament injuries include the Achilles tendon of the ankle,[6] the patella tendon of the knee, the extensor carpi radialis brevis tendon of the elbow,[7] the rotator cuff tendons of the shoulder,[8] the anterior cruciate ligament (ACL) of the knee [9] and the ankle ligaments.[10] The exact aetiology of these injuries remains unknown,[11-16] but several intrinsic and extrinsic risk factors have been implicated in the mechanism of injury for each.[14,16-21]

The ACL is a major intra-articular ligament which stabilises the knee joint and contributes to its normal function.[22] Rupture to the ACL accounts for the majority of ligament injuries in the knee joint [23] and although relatively uncommon in the general population, they are of high prevalence in individuals participating in sporting activities,[24] particularly female athletes.[25-28] It has been estimated that 250,000 new ACL ruptures occur annually in the United States alone.[29] Rupture of the ACL is considered one of the most severe and detrimental injuries sustained in sports;[30] costly reconstructive surgery and rehabilitation is required to treat the injury, and an individual who has sustained an ACL rupture is at greater risk of developing osteoarthritis (OA) of the knee joint.[31,32]

Intrinsic risk factors for ACL injuries can be broadly classified into four categories: anatomical, hormonal, neuromuscular, and more recently established genetic risk factors.[21,33] Harner *et al.* (1994) was the first to suggest that there may be a familial predisposition to ACL injuries,[34] and in 2005 Flynn *et al.* showed that individuals with an ACL tear were twice as likely to have a first-degree relative with an ACL tear, in comparison to controls.[9] Since then, several genetic variants within candidate genes have been implicated as contributing factors to the risk of sustaining an ACL injury,[35-39] highlighting the polygenic nature of this multifactorial condition.

Interestingly, polymorphisms within several genes (*COL1A1*, *COL5A1* and *COL12A1*) encoding proteins implicated in the regulation of fibrillogenesis, have been associated with ACL injury risk in several candidate gene association studies.[35-38] Proteoglycans are a unique family of proteins within the extracellular matrix (ECM) having decidedly diverse functions which include playing a role in

fibrillogenesis.[40] Appropriately regulated fibrillogenesis is essential for the development of regular collagen fibrils and normal fibrillar organisation.[41-44] This is vital for the proper functioning of ligaments and thus important components of the fibril, such as proteoglycans, warrant investigation.

Increasing our understanding of the biological consequences of the numerous complex interactions between the components of the ECM in regulating fibrillogenesis, should assist us in elucidating the aetiology of ACL injuries. The identification of additional genetic risk factors that contribute to sustaining an ACL rupture will further add to the understanding of the mechanisms underlying this injury. This is essential for the development of preventative measures so as to reduce the incidence of this detrimental injury. Comprehending the pathobiology of ACL injuries may also lead to the development of improved treatments after injury. Following the approach of a candidate gene based genetic association study; the primary aim of this dissertation was (i) to identify novel genetic loci within proteoglycan encoding genes, which may predispose an individual to sustaining an ACL rupture. In addition, this dissertation aimed to identify (ii) genomic intervals overlapping the proteoglycan genes, and (iii) potential gene-gene interactions between proteoglycans and the *COL5A1* gene, which may highlight the biological mechanisms involved in the modulation of ACL injury susceptibility.

In preparation for the experimental chapters of this dissertation, Chapter 1 reviews the current literature. This review provides an overview of the anatomy and function of the ACL (Section 1.1), injury to the ACL (Section 1.2) and the risk factors associated with ACL injuries (Section 1.3). This section also details the genetic risk factors that have previously been associated with ACL injury risk (Section 1.3.2). The proteoglycans, the focus of this dissertation, are also introduced in Chapter 1 (Section 1.4), and the novel candidate genes to be implicated in ACL injury susceptibility are identified and discussed (Section 1.5). The structure and functions of proteoglycans are explored, as well as the structural role of proteoglycans in ligaments (Section 1.6) and proteoglycans in the injured state (Section 1.7).

Subsequent experimental chapters present the candidate gene approach followed to fulfil the aims of (i) investigating loci within proteoglycan encoding genes, and (ii) regions encompassing the proteoglycan encoding genes, for an association with ACL injury susceptibility in a Caucasian South African population (Chapter 2). Chapter 3 aimed (iii) to explore the possible interaction of proteoglycan encoding genes with the previously ACL risk associated *COL5A1* rs12722 polymorphism in the modulation of ACL injury susceptibility.

Chapter 4 concludes the novel findings of this dissertation and the limitations. It further highlights the role of genetics in ACL ruptures and the possible clinical implications of the findings of this dissertation. This chapter also presents some of the potential future directions of research in this area of ACL injuries.

Chapter 1: Literature review

1.1 The anterior cruciate ligament

Ligaments are dense bands of collagenous tissue which extend over a joint and are attached to the bone at either end.[45] The primary function of ligaments is to mechanically stabilise joints and guide them through their normal range of motion when a tensile load is applied.[45] The viscoelastic behaviour of ligaments also enables them to provide joint homeostasis. Lastly, ligaments function in joint proprioception: the conscious perception of limb position in space.[45]

Throughout the human body ligaments differ in size, shape, location and orientation. They are organised hierarchically into interconnected groups of parallel fibre bundles which run along the long axis of the ligament. These fibrous bundles, or fascicles, are made up of collagen fibres, which in turn consist of smaller collagen fibrils.[45] The collagen bundles run parallel to the long axis, and exhibit waviness or 'crimp' along the length. Crimp refers to the regular pattern of the matrix and has a biomechanical role concerning the ligaments loading state: when an appropriate load is applied to the ligament it can 'uncrimp', elongate and lengthen to a point of stiffness or restraint without sustaining fibrous damage. Applying further strain beyond the point of stiffness may result in the rupture of the ligament.[45,46]

Ligaments are more complex at the microscopic level: cells, called fibroblasts, compose a small part of the total structure, and are responsible for synthesising and maintaining the ECM which surrounds the cells.[47] Biomechanically ligaments are composed of two-thirds water and one-third solid components. The water is responsible for the cellular function and viscoelastic behaviour of ligaments.[45] The solid component of the ECM consists primarily of collagen, which accounts for approximately 75% of the dry weight (85% of which is type I collagen and the remainder is made up of type III, V, VI, XI and XIV). The balance of the solid component consists of elastin, glycoproteins, and glycosaminoglycans (GAGs).[45]

The ACL is a major intra-articular ligament, which connects the femur to the tibia (Figure 1.1) and together with the posterior cruciate ligament (PCL) stabilises the knee joint as it moves through its normal range of motion. It functions in preventing excessive anterior translation of the tibia on the femur, as well as limiting internal rotation of the tibia and providing posteroanterior knee stability.[48] Because of the position and functions of this ligament, it is one of the most frequently injured structures during sporting activities.[49]

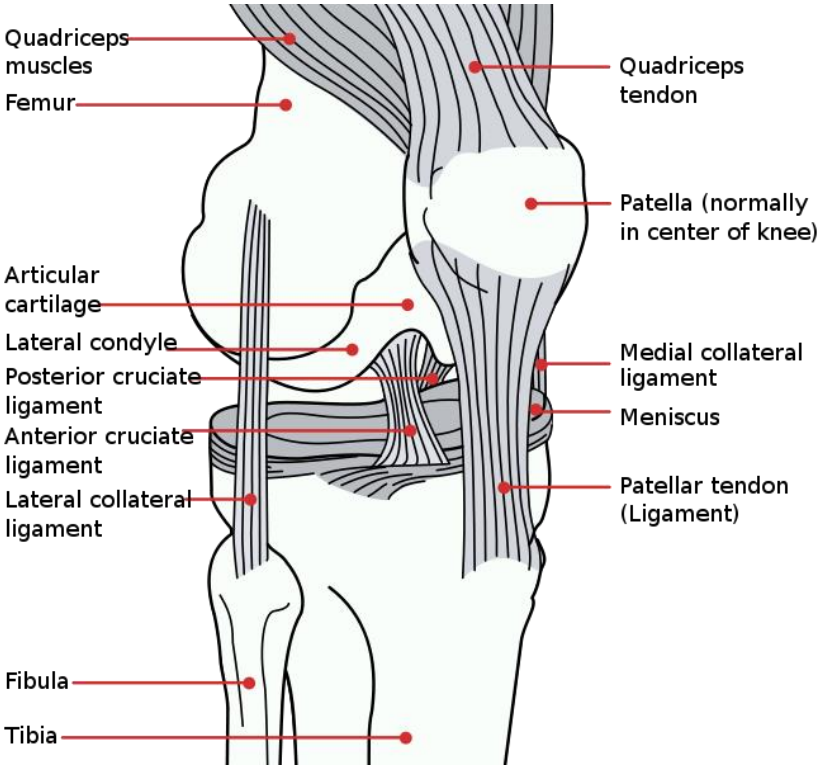


Figure 1.1: Basic anatomy of the right knee joint displaying the anterior cruciate ligament and surrounding structures. The femur and tibia, as well as the posterior cruciate ligament, medial collateral ligament and lateral collateral ligament are shown.[50]

Proximally the ACL originates from the posterior medial surface of the lateral femoral condyle where it is attached in a semi-circle;[51] the ACL fans out to the larger and stronger distal attachment which is on the anterior portion of the intercondylar eminence of the tibia.[52,53] The tibial insertion is approximately 120% broader than that of the femoral insertion.[54] The ACL orientates in a spiral across the joint due to its attachments: from the femur it runs anteriorly, medially, and distally to the tibia.[55] The PCL attaches from the posterior region of the tibial intercondylar eminence to the femoral medial condyle. The ACL and PCL are completely covered in synovium, hence why the ACL is classified as extra-synovial but intra-articular.[56] The medial collateral ligament (MCL) and lateral collateral ligament (LCL) are extra-articular (Figure 1.1) and are important for knee stability during inward (varus) and outward (valgus) angulation of the joint.[52]

The ACL has a mean length of 32 mm and its width ranges from 7 to 12 mm.[55] Collagen fibrils composing the ACL are grouped to form 1 to 20 µm fibres surrounded by other ECM proteins. The fibres subsequently merge to form larger sub-fasicular units that can fan out into functional bands.[45] The ACL contains two distinct functional bands or bundles of fibres which differ in tension

during varying ranges of motion: the anteromedial bundle (AMB) lengthens and tightens during knee flexion, while the posterolateral bundle (PLB) shortens and becomes slack; the PLB tightens during knee extension.[53,57,58] The PLB controls anterior translation between 0 and 30˚ flexion and the AMB controls translation after 30˚ flexion. These bundles are named after the locations of their tibial insertions. The fibres of the AMB have their origin in the most proximal part of the femur and insert at the anteromedial part of the tibia, while fibres of the PLB originate from the distal part of the femur and insert on the posterolateral part of the tibia.[59] These collagenous bands attach to the bones by transitional zones of fibrocartilage and mineralized cartilage.[56,60] The forces that are transmitted through these bundles differ with varying knee-joint positions.[61] Controversy exists in the literature as to whether the ACL is composed of three bundles rather than two. To date the understanding is that the ACL has two distinct functional bundles [53,55,59,62,63] but three distinct bundles when examined anatomically.[64-66]

Macroscopically, the ACL is composed of multiple fascicles surrounded by a connective tissue sheath. The fascicles are composed of sub-fasciculi, ensheathed by loose connective tissue, which are in turn composed of sub-fascicular units also surrounded by a loose connective tissue consisting of type II collagen. These sub-fascicular units are made up of fibres of collagen fibrils, the basic unit of the ACL.[59,67] Two types of fibrils compose the ACL: large inhomogeneous fibrils resist high tensile stress (approximately 50%), and small homogeneous fibrils maintain the structural organisation of the ligament (approximately 44%). These fibrils thus compose approximately 94% of the ACL tissue while the remaining 6% constitutes the cells and matrix components.[67] The ACL matrix consists of collagens (with type I collagen being a major component), proteoglycans, GAGs, glyco-conjugates and elastic components.[59,67] Proteoglycan content of the ACL is highly variable across the fibrocartilage and at the bone attachment sites.[68]

1.2 Injury to the ACL

ACL injuries are the most common ligamentous injury of the knee.[56] Traumatic joint injuries are commonly responsible for ligament tears, which may be either partial or complete ruptures. These ligament tears are usually owed to the application of an excessive external load to the limb; however the ACL can also be injured without an external force being directly or indirectly applied to the knee joint, but rather due to internal forces generated within the joint.[52] Ultimately the ACL will fail when a biomechanical load placed upon it exceeds the maximum load that it can withstand.[69] Since the ACL acts to resist anterior tibial translation and internal tibial rotation, it is predominantly ruptured when the femur is externally rotated in relation to a fixed lower extremity or if the tibia is internally rotated relative to the femur; the ACL is also at greater risk of injury during excessive knee extension.[56,70]

Most ACL ruptures involve the middle substance of the ligament and infrequently the femoral or tibial attachments.[56] When an ACL rupture occurs it is likely that a secondary injury to the menisci, articular cartilage or other ligamentous structures will take place. Meniscal injuries are reported in approximately 50-70% of patients with ACL injuries and MCL injuries occur in approximately 18-22% of patients with ACL injuries; however these can also occur in individuals without ACL injuries.[56,71]

It has been estimated that 250,000 new ACL ruptures occur annually in the United States,[29,33] however it is difficult to determine ACL injury incidence rates in many countries because not all individuals with an ACL injury seek medical care.[72] ACL ruptures are considered one of the most severe and detrimental injuries sustained in sports,[30] as the ligament does not heal and reconstructive surgery and rehabilitation is required to treat this injury,[73] a costly expense. Following an ACL rupture, an athlete will be forced to miss several months of competitive sport which may result in the loss of contracts or sponsorship, and the athlete may never regain their pre-injured state of fitness or skill.[74] Without reconstructive surgery the knee remains unstable and is predisposed to further injury. An estimated 125,000 ACL reconstructions were performed in the United States in 2006, and the number of ACL reconstructions taking place appears to be increasing annually.[75] The Swedish National ACL Register reported 15,387 primary ACL reconstructions during the period from 2005 to 2010,[76] while the Danish ACL reconstruction registry reported 12,193 primary reconstructions for the same period.[77]

When a complete ACL rupture occurs, the ligament rapidly degenerates and thus primary repair does not usually provide a pleasing end result.[78,79] The poor healing capacity of the ACL (relative to the extra-articular MCL) may be due to the biochemical properties of the ECM, biomechanical forces, intra-synovial environment, blood supply or nutrient supply.[78,79] Cellular proliferation, migration, collagen synthesis, response to growth factors and variable gene expression may also be responsible for the poor healing capability of the ACL.[78-81]

An individual who has sustained an ACL injury is at a greater risk of developing OA of the knee joint because of the degenerative changes that take place following the injury.[31,32,82] Lohmander *et al.* (2004) examined female soccer players twelve years after ACL injury and discovered that 82% of the participants had radiographic changes in their knee index while 51% met the criteria for radiographic knee OA.[83] Even after ACL reconstructive surgery is performed, patients still have degeneration of the articular cartilage and symptoms of early onset OA. Emphasis thus needs to be placed on increasing research efforts in order to prevent ACL injuries as well as to improve treatment interventions.

1.2.1 Mechanisms of ACL injury

ACL injuries may occur through contact or non-contact mechanisms, both of which have been described as the mechanism of injury in the large majority of participants investigated in the experimental chapters of this dissertation. The mechanism of injury for downhill alpine skiing is completely different to that of contact and non-contact, and the ACL injury in this case may result from several different causes described in detail by Silvers *et al.* (2007).[52] Contact injuries are due to the application of an external force, such as collision with another player or an object in the area of activity. These injuries can further be divided into two different mechanisms: direct or indirect contact. Direct contact is when an external force is applied directly to the knee that is injured, whereas indirect contact is when an external force is applied to the individuals' body but not directly to the injured knee. On the other hand, non-contact ACL injuries are a result of internal forces created within the knee joint as a result of the athletes own movements.[33,52] The classification of the mechanisms of ACL injury is in accordance with the American Orthopaedic Society of Sports Medicine.[33]

The majority of reported ACL injuries are caused by non-contact mechanisms.[28,33,84] The frequency of non-contact injuries however varies with type of activity: sports that involve landing and pivoting have high incidence of non-contact ACL injuries, while contact sports tend to have a lower frequency of non-contact injuries and higher frequency of contact injuries.[84] Non-contact injuries usually involve sudden deceleration in a single step or stopping action, or quick direction changes, as well as cutting or twisting movements. Injury by a non-contact mechanism typically occurs with the knee in an extended position, and thus landing from a jump with either near or complete knee and hip extension or insufficient flexion, is another cause.[33] If an individual attempts to rapidly decelerate or change direction when the foot is fixed and pronated, the tibia is rotated inward or the knee is near or at complete extension, it is likely that the ACL will partially or completely rupture due to the undesirable torsional force placed on the knee joint.[33,52]

1.3 Risk factors associated with ACL injury

Although the precise aetiology remains unknown, multiple risk factors have been implicated in ACL injuries.[9,85] Susceptibility to an ACL rupture is understood to be multifactorial, with complex interactions between risk factors. Males and females are believed to have separate risk models for ACL injury,[86] as females have a 3 fold greater risk of sustaining an ACL rupture in comparison to their male counterparts.[28]

Risk factors for multifactorial disorders, including ACL injuries, can be classified into two groups: intrinsic and extrinsic. Intrinsic risk factors influence the individual from within and are generally non-modifiable, with the exception of weight and body mass index (BMI). Extrinsic risk factors impact the individual from the external environment and are generally modifiable.[4]

To date, five different groups of risk factors, one extrinsic and four intrinsic, have been implicated in an individual sustaining an ACL injury. The four intrinsic risk factors include: anatomy, hormones, neuromuscular factors and genetics.[4,33] These risk factors are not mutually exclusive, but rather work in combination to modulate ACL injury susceptibility. One can hypothesise that the interaction of risk factors, for example the extrinsic environment and genetics, contribute to the observed inter-individual variation noted with regards to sustaining an ACL injury.

Individuals or athletes having one or more intrinsic risk factors are considered to be predisposed to injury. Exposure of these individuals to extrinsic risk factors increases their susceptibility to sustaining an injury. Ultimately the occurrence of an inciting event results in an injury, such as an ACL rupture (Figure 1.2).[4,5,87] Note, it is not the risk factors which cause the injury; they simply contribute to an individual being more susceptible to injury. The identification of risk factors predisposing an individual to injury will aid in understanding the aetiology and biological mechanisms underlying musculoskeletal soft tissue injuries such as ACL ruptures.

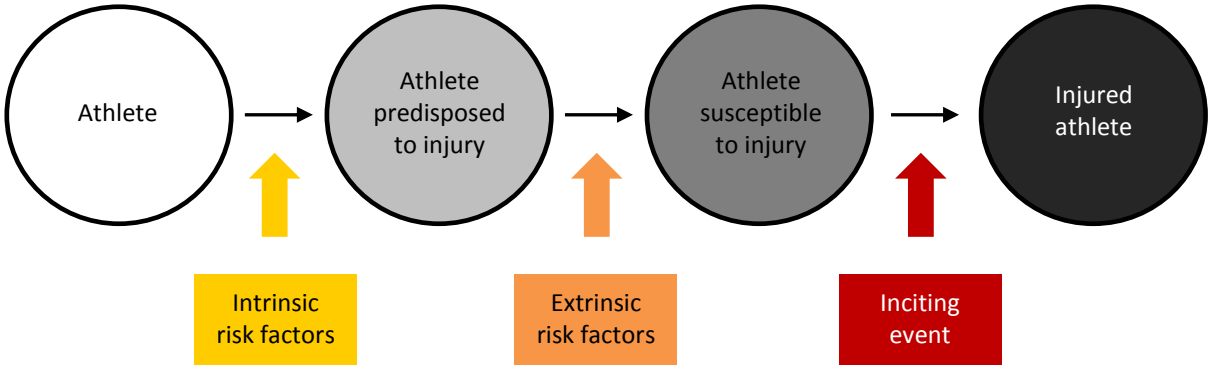


Figure 1.2: A schematic diagram illustrating the complex multifactorial relationship between intrinsic and extrinsic risk factors, and an inciting event that results in an anterior cruciate ligament rupture. Intrinsic risk factors can be broadly classified as anatomical, hormonal, neuromuscular, and genetic, while extrinsic risk factors involve the environment the athlete is placed in. These risk factors predispose an individual to sustaining an ACL rupture. Ultimately an inciting event results in the injury of the athlete. Adapted from Meeuwisse *et al.* (1994)[4] and Collins *et al.* (2009).[5]

1.3.1 Extrinsic and intrinsic risk factors for ACL injury

An extensive evidence-based review of intrinsic risk factors for ACL injuries has recently been published.[21] The intrinsic risk factors will therefore be summarised in this dissertation. Since the extrinsic risk factors are not the primary focus of this dissertation, they will also only be summarised in the following section.

1.3.1.1 Extrinsic risk factors

Extrinsic risk factors that may affect susceptibility to ACL injuries include the weather, playing surface, footwear and prophylactic/functional knee bracing (Table 1.1).[52]

In the analysis of National Football League (NFL) games, it was found that when playing in outdoor stadiums, cold weather is associated with a lower risk of ACL injuries in comparison to hot weather; this was suggested to be due to lower shoe-surface traction.[88] Evidence suggests that friction is the key element playing a role in injury risk with regards to playing surface. Artificial surfaces have been found to have higher shoe-surface friction and thus a higher risk for non-contact injuries.[89,90]

When considering footwear, it is important to note that although increased friction may enhance performance and allow more rapid sprinting, stopping and changes in direction, it will also unintentionally increase the risk of injury during these movements. It is important to identify optimal footwear that will decrease rotational traction in order to try avoid injury, but enhance intermediate friction for ideal performance.[91]

Controversy exists regarding the efficacy of the use of knee braces.[92] A study involving a major American college football team demonstrated that knee injury incidence, including ACL injury incidence, was greater when laterally placed prophylactic knee braces were worn.[93] Another study on American football players showed that there is no statistically significant difference in injury rates or severities when wearing prophylactic knee braces or not.[94]

1.3.1.2 Intrinsic risk factors for ACL injury

The intrinsic risk factors will be discussed in four categories: anatomical, hormonal, neuromuscular (Table 1.1) and genetic risk factors (Table 1.2). Researchers have overwhelmingly shown that female athletes have an excessively higher incidence of ACL injuries in comparison to males,[25,27,28,86,95-99] and several studies have provided evidence that the female sex is an intrinsic risk factor for ACL ruptures.[37,38,100,101] Anatomy, anthropometric, hormonal and neuromuscular disparities may be the reasons for the difference in ACL injury susceptibility between men and women. Gwinn *et al.* (2000) showed that overall, female participants had a 2.4 to 1 relative risk of ACL injury in sports and a 9.5 to 1 relative risk in military exercises in comparison to males. This study was conducted on a United States military training cohort with identical facilities for males and females, it was thus suggested that intrinsic factors could be the only reason for the increased risk of ACL injury in females.[99]

Anatomical risk factors

Anatomical risk factors include ACL size and geometry, quadriceps angle (Q angle), notch width, tibial slope, foot pronation, pelvic tilt, BMI and knee laxity.[33,86] The ACL has sex-based variations in: size, females have smaller ACLs;[102] structural properties, females have lower load to failure, elongation to failure and energy absorbed at failure;[103] and ultrastructure, females have more

densely packed collagen fibrils.[104] Thus, anatomical risk factors for ACL injuries differ between males and females.

It has been suggested that males and females with ACL ruptures have smaller ACL diameters.[105] Several studies have also shown that the ACL diameter and cross-sectional area is smaller for females in comparison to males,[52,106,107] perhaps a reason for the higher injury incidence observed in females. In addition, the Q angle which is believed to reflect the alignment of the leg, is greater in females in comparison to males; this could be another reason why females are at an increased risk of ACL ruptures,[108] however there is no definitive evidence of this.

Increased risk of ACL ruptures has further been associated with a smaller femoral intercondylar notch width and femoral notch width index.[33,105,109-111] Interestingly, females have a smaller femoral intercondylar notch width in comparison to males. Men have a wider 'U-shaped' intercondylar notch, whereas women have a narrower 'A-shaped' notch.[106] Moreover, several other studies have also implicated notch opening,[112] lateral condylar width [113] and total condylar width [107] in ACL injury susceptibility. There is also an anatomical correlation between height and total condylar width in males and females; ACL area is likewise correlated to height in males but not females,[107] thus indirectly implicating height in ACL injury risk.

The tibial slope of the ACL is thought to be a risk factor for ACL injuries because an increased tibial slope is associated with increased anterior tibial translation which may increase strain on the ACL.[114]

Foot and pelvic alignment has also been associated with the risk of ACL ruptures. Increased subtalar joint pronation of the foot has been suggested to be associated with an increased risk of ACL ruptures [115] because increased pronation results in greater rotation of the knee joint.[116] Females characteristically have excessive subtalar pronation in comparison to males.[106] An increase in anterior pelvic tilt is also thought to be associated with ACL ruptures because of its effect on the alignment on the lower leg.[117]

The risk of ACL injury was previously associated with an increased BMI. One of the reasons for this is that individuals with an increased BMI tend to have a more extended knee position when landing.[33] Increased BMI has been significantly associated with the risk of ACL ruptures in females, but not with males.[105]

Generalised joint laxity, an overall measurement of whole body joint laxity or flexibility, has been associated with an increased risk of ACL ruptures in both males and females.[105,117] Females generally have greater generalised joint laxity in comparison to males.[118] Increased anterior knee laxity or increased anterior tibial translation has also been suggested as an anatomical risk factor for ACL injuries.[105]

Hormonal risk factors

The female menstrual cycle phase was previously highlighted as one of the hormonal risk factors for ACL injury.[19] Receptor sites for oestrogen, progesterone and relaxin have been found in the ACL,[52,119] and sex hormones have been suggested to affect ACL structure and composition and may play a role in the biological processes of the ACL.[119] The fluctuations in oestrogen and progesterone during the menstrual cycle may therefore influence the properties of the ACL and a female may thus be more susceptible to injury at different stages of the cycle.[26,120-124]

Neuromuscular risk factors

Neuromuscular control is defined as the unconscious efferent response to an afferent signal about dynamic joint stability. Efferent signals either involve feedback mechanisms: reactions due to an afferent input (force to the joint), or feed forward mechanisms: results of pre-activated preparatory activation of muscle. Proprioceptive activities may play a role in injury reduction because they are vital for optimal motor performance.[52] Other neuromuscular risk factors proposed for ACL injuries include reduced hip and knee flexion angles, increased knee valgus, internal rotation of the femur on the tibia, high quadriceps muscle activity in comparison to the hamstrings and inadequate trans-knee muscle stiffness.[33] These factors pose risks because they cause kinematic and kinetic differences during cutting and pivoting or landing from a jump.[33]

Muscular recruitment and strength are critical for knee stability, and an imbalance in quadriceps to hamstring strength ratios, especially a weak hamstring, puts the ACL at an increased risk of injury.[52] Furthermore, the ACL is strained when landing from a jump with insufficient knee and hip flexion due to the increased load on the knee and force from the quadriceps. During flexion in a squat position, the loaded hamstring muscle functions together with the ACL to reduce anterior tibial

translation and internal tibial rotation.[125] Valgus loading was further implicated as a likely cause of injury during sidestepping-cutting manoeuvres, particularly in females.[126,127]

Table 1.1: A summary of the risk factors implicated in anterior cruciate ligament injury susceptibility.

Extrinsic risk factors	Intrinsic risk factors			
	Anatomical	Hormonal	Neuromuscular	Genetic
physical activity	sex	oestrogen concentration	flexion angles	*COL1A1*
training errors	age	progesterone concentration	knee valgus	*COL5A1*
weather	BMI	menstrual cycle phase	rotation	*COL12A1*
playing surface	ACL size and geometry		muscular activity	*MMPs*
footwear	knee joint geometry			
knee bracing	knee joint laxity			
medication	pelvis and foot alignment			

Abbreviations: ACL, anterior cruciate ligament; BMI, body mass index; *COL1A1*, collagen 1 alpha 1 gene; *COL5A1*, collagen 5 alpha 1 gene; *COL12A1*, collagen 12 alpha 1 gene; *MMPs*, the genes encoding the matrix metalloproteinases.

Genetic risk factors

In 2005 Flynn *et al.* suggested that there is a familial predisposition toward tearing the ACL,[9] as had previously been alluded to by Harner *et al.* (1994).[34] Flynn's study demonstrated that participants with an ACL rupture were twice as likely to have a first-degree relative with an ACL rupture in comparison to those participants who had no history of ACL injuries, thus proposing genetic factors as risks for ACL injuries. Genetic loci within several genes have since been identified as being associated with ACL injury risk. The genes and their associated variants are listed in Table 1.2. These case-control genetic association studies were successful in exploiting genetics as a tool for identifying plausible pathways that may be contributing to the risk of ACL injuries. Understanding these pathways aids in the understanding of the molecular mechanisms that underlie ACL injury susceptibility.

Table 1.2: A summary of the genetic loci and respective polymorphisms previously associated with anterior cruciate ligament injuries.

Gene	Function	Polymorphism accession number	Associated injury	Reference
COL1A1	Encodes the pro-α1 polypeptide of type I collagen	rs1800012	ACL ruptures and shoulder dislocations ACL ruptures	Khoschnau *et al.* (2008)[35] Posthumus *et al.* (2009) [36] Ficek *et al.* (2013)[128]
COL5A1	Encodes the pro-α1 polypeptide of type V collagen	rs12722	ACL ruptures in females	Posthumus *et al.* (2009)[37]
COL12A1	Encodes the pro-α1 polypeptide of type XII collagen	rs970547	ACL ruptures in females	Posthumus *et al.* (2010)[38]
***MMP* chromosomal region 11q22**	Encodes for matrix metalloproteinases 1, 3, 10 and 12	*MMP1* *rs1799750,* *MMP3* *rs679620,* *MMP10* *rs486055,* *MMP12* rs2276109	ACL ruptures	Posthumus *et al.* (2011)[39]

Abbreviations: ACL, anterior cruciate ligament; *COL1A1*, collagen 1 alpha 1; *COL5A1*, collagen 5 alpha 1; *COL12A1*, collagen 12 alpha 1; *MMP*, matrix metalloproteinase.

Although genetic risk factors have been considered as a separate category in this dissertation, it is important to note that individually the anatomical, hormonal and neuromuscular risk factors are all multifactorial phenotypes determined by both genetic and non-genetic factors. Genetic risk factors are therefore probably not independent factors but rather part of the traditional three categories of mentioned intrinsic factors. For example flexibility is an anatomical risk factor for musculoskeletal soft tissue injuries that is influenced by genetics; the *COL5A1* gene has been associated with flexibility in a sit and reach test.[130]

1.3.2 Genetic loci previously associated with ACL injury risk

1.3.2.1 *COL1A1*

The *collagen I α1* gene (*COL1A1*) encodes the α1 chain of type I collagen and regulates the production of this fibril forming collagen.[131] Type I collagen is the major protein constituent of ligaments, comprising 70-80% of the dry weight.[45] This collagen is a key structural component of ligaments, tendons and other non-cartilaginous connective tissues, and is arranged in bundles of parallel fibres to make up the tissue.[45] The Sp1 transcription factor binding site polymorphism (rs1800012) within intron 1 of the *COL1A1* gene is a G>T substitution that has previously been associated with cruciate ligament injuries as well as shoulder dislocations in the Caucasian population.[35,36,128] The rare TT genotype of this polymorphism was shown to be significantly under-represented in participants with cruciate ligament ruptures, predominantly ACL, in comparison to controls. Although this genotype is rare in the general population, this is still an interesting association.[5,35,36,128] The Sp1 transcription factor binding site polymorphism has also been associated with several other complex multifactorial phenotypes including OA,[132] osteoporotic fractures,[133] lumber disc disease,[134] myocardial infarction [135] and stress urinary incontinence.[136] This polymorphism has also been investigated in Achilles tendon injuries but no association was identified between rs1800012 and Achilles tendinopathy or Achilles tendon ruptures.[129] The T allele of rs1800012 has been proposed to increase the binding affinity of the Sp1 transcription factor, causing an increase in *COL1A1* gene expression and the construction of a weaker type I collagen homotrimer, composed of three α1(I) chains.[133,137] In addition, Ficek *et al.* (2013) recently implicated a haplotype within the promoter of the *COL1A1* gene, containing rs1800012 and rs1107946, in ACL injury susceptibility.[128]

1.3.2.2 *COL5A1*

The *collagen V α1* gene (*COL5A1*) encodes the α1 chain of type V collagen. Type V collagen, a minor fibrillar collagen, is another structural component of connective tissues which makes up approximately 10% of the collagen content in ligaments.[138] Type V collagen intercalates with type I collagen and is believed to regulate the collagen fibril diameter.[138] The *Bst*UI restriction fragment length polymorphism (RFLP) (rs12722) within *COL5A1*, is a C>T substitution which localises to the 3'-

untranslated region (UTR). The CC genotype of this polymorphism was found to be significantly over-represented in a control group of female participants, while the T allele (CT and TT genotypes) was significantly over-represented in a female group with ACL ruptures, when Caucasian females from ACL rupture and control groups were compared.[37] The 3'-UTR contains important regulatory elements; the T allele of *COL5A1* rs12722 is believed to increase *COL5A1* mRNA stability, suggesting that more α1(V) chain is synthesised.[139] Laguette *et al.* (2011) have recently identified three polyadenylation signals and a functional microRNA (miRNA) binding site for Hsa-miR-608 within the 3'-UTR of *COL5A1*.[139] The *COL5A1* gene has also been associated with phenotypes other than soft tissue injuries, including range of motion and endurance running performance.[140]

1.3.2.3 *COL12A1*

The *collagen XII α1* gene (*COL12A1*) encodes a structural component of ligament fibrils, the α1 chain of type XII collagen. Type XII collagen mediates fibril interactions with extracellular and cell surface molecules.[141] This collagen associates with the surface of collagen fibrils, in order to mediate interactions, and is a member of the fibril associated collagens with interrupted triple helicies (FACIT) family. Sequence variants within *COL12A1* have also been investigated for an association with ACL ruptures.[38] The *COL12A1 Alu*I RFLP (rs970547) is a G>A substitution that is significantly associated with ACL ruptures in Caucasian females; the AA genotype of this polymorphism was significantly over-represented in a female ACL injured group in comparison to controls in a study by Posthumus *et al*. (2010).[38] This polymorphism (rs970547) lies within exon 65 of *COL12A1* and the non-synonymous G>A substitution (the GG genotype is rare) results in a glycine to serine amino acid change. Although this substitution results in a non-polar neutral glycine amino acid being substituted by a neutral polar amino acid, the variation has no known functional significance. However, it has been suggested that this polymorphism may cause an altered type XII collagen to be produced, which may change the interactions of this peptide with other proteins and modify the biomechanical properties of the collagen fibril.[38]

1.3.2.4 *MMPs* (11q22)

Matrix metalloproteinases (MMPs) are the key regulators of ECM degradation and remodelling and are thus critical to ligament homeostasis and integrity.[142] Four polymorphisms within genes

encoding MMPs,(*MMP1* rs1799750, *MMP3* rs679620, *MMP10* rs486055 and *MMP12* rs2276109) which are all positioned on chromosome 11q22, have been investigated for an association with ACL ruptures in Caucasian individuals. Genetic variants within this region have previously been associated with multiple complex phenotypes including OA,[143] rheumatoid arthritis (RA),[144] lumbar disk degeneration,[145] idiopathic scoliosis,[146] aseptic prosthetic loosening [147] and Achilles tendinopathy.[142] *Matrix metalloproteinase 1 (MMP1)* encodes a secreted matrix enzyme which breaks down the interstitial collagen types I, II, and III.[148] The *MMP1* rs1799750 single nucleotide polymorphism (SNP) is a deletion/insertion of a G nucleotide, which is known to be functional; the 2G allele causes significantly higher transcription of the gene in comparison to the 1G allele.[149] *Matrix metalloproteinase 3 (MMP3)* encodes the MMP3 enzyme which degrades collagen types II, III, IV, IX, and X, proteoglycans, fibronectin, laminin and elastin and can also activate other MMPs.[150] The *MMP3* gene has previously been associated with Achilles tendinopathy.[142] *Matrix metalloproteinase 10* (*MMP10*) encodes MMP10 which degrades proteoglycans and fibronectin.[151] Both *MMP3* rs679620 and *MMP10* rs486055 are non-synonymous polymorphisms which result in an amino acid change; however the functional significance of both these SNPs is yet to be defined.[152] *Matrix metalloproteinase 12* (*MMP12*) encodes MMP12 which degrades soluble and insoluble elastin,[153] and the rs2276109 T>C substitution polymorphism within this gene was found to be independently associated with ACL ruptures.[39] This polymorphism lies within the upstream regulatory promoter region of the *MMP12* gene and is proposed to have functional effects on the gene. The A allele is believed to increase the activity of the transcription factor activator protein-1 and increase *MMP12* gene expression.[154] Two inferred haplotypes constructed using different combinations of the variants in the *MMP1, MMP3, MMP10* and *MMP12* genes listed above, were also found to be associated with ACL injury risk.[38]

The genetic association studies discussed have provided cumulative evidence that genetic risk factors are an important component of the intrinsic risk factors that may predispose an individual to sustaining an ACL injury. These studies have also been vital in identifying structural components of the ACL that contribute to the molecular mechanisms of injury risk. Injury to the ACL is however a multifactorial condition subjected to multiple risks; it is thus probable that there are multiple genes contributing to the aetiology of this injury.

The collagen genes are vitally important in fibrillogenesis as well as the formation of collagen fibrils of regular organisation. Mutations within the collagen genes are associated with several connective tissue disorders which result in the loss of mechanical functions of tendons and ligaments.[140,155,156]

Mutations in the *COL5A1* gene, for example, have been shown to cause classic Ehlers–Danlos syndrome (EDS).[157] The proteoglycans, found within the ECM, are also important structural constituents of connective tissues, such as the ACL, having significant roles in fibrillogenesis, ECM remodelling as well as cell signalling within the ligament.[158] Mutations within proteoglycans have also been implicated in connective tissue disorders,[159,160] and murine tissues deficient in certain proteoglycans have similar physical phenotypes to humans with classic EDS: fibrillogenesis is compromised resulting in collagen fibrils of highly irregular diameters and abnormal fibrillar organisation.[41-43]

Genes biologically relevant to ACL injuries, encoding proteins and enzymes involved in the structure and function of the ACL tissue, as well as signalling molecules within the tissue, should be investigated for associations with ACL injury risk. The genes encoding proteoglycans are therefore viable candidates to be investigated for an association with ACL injury susceptibility.

1.4 Introduction to proteoglycans

Proteoglycans, also known as mucoproteins, are glycosylated proteins that vary in size and structure. They are found throughout the human body, on the cell surface and within the ECM. Proteoglycans are composed of a core protein peptide, ranging from 10 kDa to >500 kDa in size, to which one or more GAGs are covalently attached. GAGs are linear carbohydrate polysaccharide chains formed from repeating disaccharide sugar units. The most common GAG structures include chondroitin sulfate (CS), dermatan sulfate (DS), heparin, heparan sulfate (HS), keratan sulfate (KS) and a non-sulfated GAG, hyaluronic acid (HA).[161] Either one or both of the sugars in the GAG unit have a sulfate group, rendering the proteoglycan highly negatively charged. This is essential for the function of proteoglycans which bind water giving the ECM gel like properties, thus improving the biomechanical properties of the connective tissue and providing the collagen fibrils with high capacity to resist compressive and tensile forces.[162] Proteoglycans have vast molecular diversity and biological functions because the various protein cores can have one or more different GAG chains attached.[163] They play an important role in maintaining the structural integrity of tissues by stabilising the collagenous network. Proteoglycans are also involved in cell signalling pathways as well as the regulation of ionic homeostasis and collagen fibrillogenesis, which is essential for development and tissue repair.[158,162]

Proteoglycans can be broadly classified into three major groups according to their protein core, structure and function: the small leucine-rich proteoglycans (SLRPs), the large modular proteoglycans, and the cell-surface proteoglycans (Table 1.3).[163] It must however be noted that it is possible to have overlapping with this classification.

Several proteoglycans such as versican and decorin are widely expressed in various tissues, however some proteoglycans are tissue specific. Neurocan, brevican and neuroglycan C for example are specific to the central nervous system (CNS).[164] The distribution of proteoglycans in tendons and ligaments is subject to which areas are affected by different mechanical forces. In tendons for example, regions subjected to longitudinal/tensional forces are rich in the SLRP decorin although there are other members of the SLRP family present.[165]

Table 1.3: Classification of the proteoglycans identified in ligaments and tendons, as reviewed in current literature.

Proteoglycan	GAGs	Ligament	Tendon
SMALL LEUCINE-RICH PROTEOGLYCANS (SLRPs)			
Class I			
Biglycan	CS/DS	Somerman 1990;[166] Yukawa 2001;[167] Ilic 2005;[168] Young 2011 [169]	Vogel 1993;[170] Berenson 1996;[171] Waggett 1998;[172] Samiric 2004 [173]
Decorin	CS/DS	Larjava 1992;[174] Yukawa 2001;[167] Ilic 2005;[168] Henninger 2007;[175] Young 2011 [169]	Vogel 1993;[170] Berenson 1996;[171] Waggett 1998;[172] Samiric 2004 [173]
Asporin		Yamada 2001[176]	
Class II			
Fibromodulin	KS	‡ Liu 2003 [177]	Waggett 1998 [172]
Lumican	KS	Chatuparisute 2012 [178]	Waggett 1998 [172]
Keratocan	KS		˜ Rees 2009 [179]
Class III			
Epiphycan (DSPG3)	CS/DS	Deere 1996 [180]	
Class IV ***			
Class V ***			
MODULAR PROTEOGLYCANS			
Hyalectans/Lecticans (HA/lectin binding)			
Versican	CS/DS	Larjava 1992;[174] Ilic 2005;[168] Young 2011 [169]	Waggett 1998;[172] Corps 2004;[181] Samiric 2004 [173]
Aggrecan	CS/KS	Ilic 2005;[168] *Nelson 2006;[182] Young 2011 [169]	Vogel 1994;[183] Berenson 1996;[171] Samiric 2004 [173]
Non-hyaluronic acid-binding ***			
CELL-SURFACE PROTEOGLYCANS			
Syndecans (1-4)	CS/HS	Worapamorn 2000 [184]	Sawaguchi 2006 [185]
Glypicans (1-6)	HS	Worapamorn 2000 [184]	Sawaguchi 2006 [185]
Betaglycan	CS/HS	Worapamorn 2000 [184]	Sawaguchi 2006 [185]
CD44			Sawaguchi 2006 [185]
Other proteoglycans			
Proteoglycan 4 (Lubrican)	CS/KS	†Elsaid [186]	Marcelino 1999;[187] ®Rees 2002 [188]

*** These classes of proteoglycans have not been described in tendons and ligaments
‡ Detected fibromodulin in the knee cartilage of canines following unilateral ACL transection
˜ Keratocan identified in bovine tendons
* Detected increase in aggrecan in the articular cartilage adjacent to the ACL following injury to the ligament
† Detected decrease in lubrican concentration in synovial fluid of knee following an ACL rupture
® Detected proteoglycan 4 in bovine tendon by immunohistochemical analyses
Abbreviations: CS, chondroitin sulfate; DS, dermatan sulfate; DSPG3, dermatan sulfate proteoglycan 3; GAGs, glycosaminoglycans; HA, hyaluronic acid; HS, heparan sulfate; KS, keratan sulfate.

The SLRPs can be divided into five subfamilies (Classes I-V) based on homology, cysteine-rich regions and chromosomal organisation.[163] SLRPs are composed of a small protein core that contains leucine-rich repeats (LRRs), N-terminal cysteine clusters and a minimum of one GAG side chain attached (Figure 1.3).[163,189,190] A distinct feature of Class I-III SLRPs is that they contain an 'ear' repeat. This is the second to last LRR (the longest LRR) that is proposed to maintain the conformation of the core protein.[191]

SLRPs not only have structural roles, but constitute a network of signalling roles as well.[60] They regulate collagen fibril growth, fibril organisation and ECM assembly. They are therefore critical for the structural integrity of ligaments and other connective tissues.[40] SLRPs are also involved in the regulation of cell-matrix interactions by directly binding plasma membrane receptors and pericellular matrix molecules.[40] As a result, they can influence cellular differentiation, apoptosis, proliferation and migration through various signalling pathways and interactions. They therefore play an important role in maintaining tissue homeostasis, as well as having roles in tissue repair and regeneration/remodelling in pathological conditions.[192] SLRPs also have tissue-specific structural and signalling functions in various microenvironments that are defined by the differences in their protein cores and GAG side chains.[40]

Class I SLRPs include biglycan, decorin, asporin and extracellular matrix protein 2.[60] Typically, the N-termini of Class I SLRPs have a cysteine cluster that forms two disulphide bonds. Class I SLRPs all have a small exonic organisation, eight exons, with the intron/exon junctions being highly conserved.[163,193] Biglycan is expressed in skin, tendons and ligaments, as well as in bone and dentin matrices.[194] This proteoglycan is involved in the regulation of collagen fibrillogenesis and interacts with collagen fibrils.[195] Decorin is found in the tensile regions of tendons, ligaments and synovial capsule,[196] and is important in tissue development and assembly as well as regulating collagen fibrillogenesis.[197] Asporin, originally named periodontal ligament-associated protein 1,[60] has also been identified in human ligaments (Table 1.3) and is structurally similar to decorin.[176]

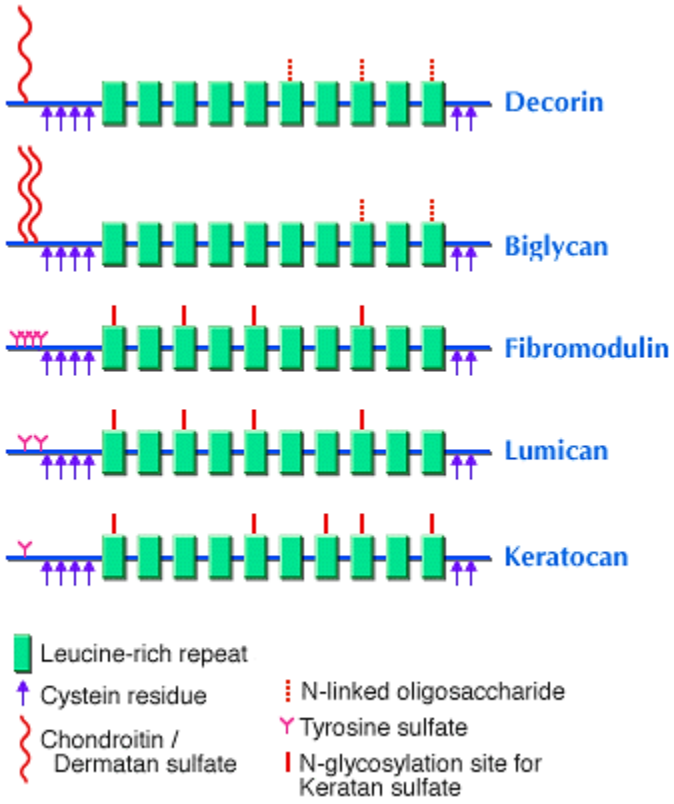


Figure 1.3: Structure of the small leucine-rich proteoglycans, displaying the important leucine-rich repeats, cysteine residues, and glycosaminoglycan side chain attachment sites.[198] Examples of Class I (decorin and biglycan) and Class II (fibromodulin, lumican and keratocan) small leucine-rich proteoglycans are displayed. Class III small leucine-rich proteoglycans are structurally similar.

Class II SLRPs include fibromodulin, lumican, proline/arginine-rich end leucine-rich repeat protein (PRELP), keratocan and osteomodulin.[60] These SLRPs have clusters of tyrosine sulfate residues at their N-termini (Figure 1.3), and bind primarily KS or an unsulfated form of keratan. The genes encoding Class II SLRPs have a highly conserved, very small exonic arrangement; only three exons, with the large central exon encoding most of the LRRs.[163,193] Fibromodulin helps define tissue integrity within ligaments and regulates collagen fibrillogenesis.[177,199,200] Lumican is expressed in various tissues, including ligaments, where it is important in fibrillogenesis and cellular proliferation.[178,199] Keratocan has also been identified in soft connective tissues such as tendons, but functions primarily in the cornea.[179]

Class III SLRPs include epiphycan, opticin and mimecan.[60] This class is classified by having a genomic arrangement comprising seven exons, and a low number of LRRs, only seven.[163,193] Epiphycan, also known as dermatan sulfate proteoglycan 3, has been identified in ligaments and has three GAG

binding sites, two N-glycosylation sites, a poly-glutamic acid stretch, and six cysteine residues.[180] This proteoglycan is important in cartilage development and remodelling.[201]

Class IV SLRPs consist of chondroadherin, nyctalopin and tsukushi, while class V SLRPs include podocan and podocan-like protein I.[60,163,193] Neither Class IV nor Class V SLRPs have yet been identified as being expressed in musculoskeletal soft tissue, such as tendons and ligaments, and are therefore beyond the scope of this review.

The heterogeneous modular group of large proteoglycans are classified as the assembly of several protein molecules in an elongated and often highly glycosylated structure.[202] The modular proteoglycans can be divided into two subfamilies: the hyalectans (HA- and lectin binding proteoglycans), and the non-HA-binding proteoglycans.[203] The non-HA-binding modular proteoglycans include perlecan, agrin, the testicans, phosphocan and several collagens. The non-HA-binding modular proteoglycans have not been identified in tendons or ligaments and are therefore beyond the scope of this review; they have recently been reviewed by Schaefer *et al.* (2010).[163]

Four proteoglycans make up the group of hyalectans: versican, aggrecan, neurocan and brevican. The hyalectan proteoglycans are described as having a tri-domain structure composed of a central domain with N- and C-terminal globular regions on either side (Figure 1.4). The central domain contains multiple attachment sites for GAGs, the N-terminal domain binds HA, and the C-terminal domain interacts with lectins (Figure 1.4).[163,202,203] Neurocan and brevican are CNS proteoglycans that inhibit neuronal attachment and neurite growth.[203]

The largest hyalectan is versican; it is present in a variety of soft tissues and is found within purely tensional regions of ligaments and tendons.[172,204] Versican is expressed by fibroblasts and functions in cell adhesion, migration and proliferation, as well as ECM assembly and fibrillogenesis of the elastic fibres.[205] Alternative RNA splicing occurs between the two large exons encoding the GAG attachment sites of versican to give rise to four different mRNA transcripts.[205,206] Because of its complex structure, versican is able to interact with various ECM molecules and cell surface proteins.[207] The variable expression of versican allows this proteoglycan to be a crucial regulator of cell-matrix interactions.[203]

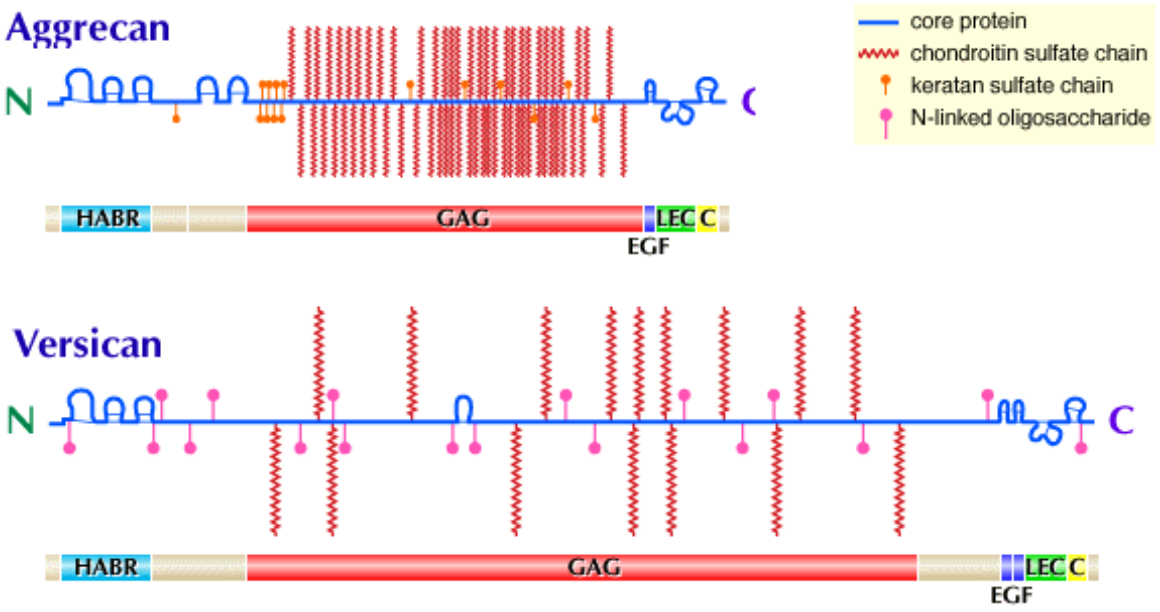


Figure 1.4: Structure of the large hyalectan proteoglycans, aggrecan and versican, displaying the central, N-terminal and C-terminal domains.[208] Abbreviations: C, complement regulatory protein-like module; EGF, epidermal growth factor-like repeat; GAG, glycosaminoglycan-attachment domain; HABR, hyaluronic acid-binding region; LEC, C-type lectin-like module.

Aggrecan exists in the form of aggregates stabilised by a link protein; it is composed of HA and up to one hundred aggrecan molecules (Figure 1.4).[209] This proteoglycan is expressed in the CNS, cartilage, ligaments and the tensional regions of tendons,[169,183,210] and regulates cartilage development, growth and homeostasis. Aggrecanases are responsible for the degradation of aggrecan in tendons and ligaments via proteolysis of specific Glu-Xaa bonds.[173] Cleavage of the aggrecan core protein by aggrecanases and other MMPs results in fragments of various sizes that can be retained and accumulate in the cartilage.[209] Chondrodysplasias, such as spondyloepiphyseal dysplasia, result from mutations in human, mice, and chicken aggrecan genes.[211-213]

The cell surface proteoglycans include those that are membrane spanning, and the glycosylphosphoinositide (GPI)-linked proteoglycans. The membrane spanning syndecans and the GPI-linked glypicans are the major HS proteoglycans found on the cell surface, where HS acts as a co-receptor that enables ligand encounters with signalling receptors.[214] Other membrane spanning proteoglycans include betaglycan, cell surface glycoprotein CD44 version 3, chondroitin sulfate proteoglycan 4, chondroitin sulfate proteoglycan 5 and receptor protein tyrosine phosphatase. The syndecans appear to be involved in complex signalling processes that regulate cellular proliferation, differentiation, adhesion and migration.[163,214,215] The GPI-linked cell surface proteoglycans include

the six members of the glypican family as well as a spliced form of brevican. Glypicans, like syndecans, are also involved in the regulation of several signalling pathways including fibroblast growth factor and transforming growth factor-β (TGF-β).[163,216]

Proteoglycan 4 (PG4), also known as lubrican, has an approximate molecular mass of 345 kDa,[217] and is suggested to prevent cell attachment to the articular surface of synovial joints, as well as lubricating articular cartilage at the synovial interface.[187,218] PG4/lubrican has been identified in both human [187] and bovine [188] tendons, as well as being observed to increase in the surrounding synovial fluid following an ACL injury.[186]

1.5 Genetic association studies and candidate genes

The aim of genetic association studies is to identify an association between one or more genetic variants and a trait.[219] Performing genetic association studies using common allelic variants is currently still cheaper than the complete sequencing of genes. Genetic association studies can be used as a powerful tool in mapping a region biologically relevant to a multifactorial condition, which can facilitate the development of specific treatments or personalised medicine. RA is a classic example of this approach being successfully applied.[220] Several variants and risk alleles within genes involved in the CD40 pathway have been implicated in the development of RA by genetic association studies. Following fine mapping and deep sequencing of the *CD40* risk locus, a potential causal SNP underlying the association of this locus with RA was identified. Functional studies implicated the risk allele in causing increased CD40 protein on the cell surface, and a reporter assay was used to examine the effect of several chemical compounds and drugs on a cell line with the risk allele. Two chemical compounds were identified to disrupt the amount of CD40 on the cell surface and are now being considered as possible treatments in the development of RA.[220] This is proof-of-concept that human genetics and association studies in particular, are an important first step in identifying potential novel loci to be targeted in therapies for complex traits such as RA. To date, case-control genetic association studies are one of the most successful methods in identifying genetic variants and regions that underlie complex multifactorial conditions, such as ACL injuries.[221] This approach has also been successfully used for musculoskeletal soft tissue injuries such as Achilles tendinopathy.[139]

Case-control genetic association studies examine polymorphisms between a group of people with a condition, such as an ACL injury, and a group without any history of the condition, which serves as a control. Statistical analyses are used to determine whether a particular genotypic variation is significantly over or under-represented in one of the groups. If so, this gene and its variation within, is said to be associated with either an increase or decrease in susceptibility to the injury. The advantage of case-control genetic association studies is that they exploit the differences between the genetic profile of affected and unaffected individuals. The disadvantage is that they do not prove that a variation is the cause of the injury, it is merely an association which provides evidence for the involvement of the investigated gene and warrants further investigation.[222]

Analysis of association can be informative even when based on non-functional genetic variants.[223] Association analyses of a functional genetic variant allows it to be directly associated with the cause

of the phenotype, and genetic variants in linkage disequilibrium (LD) may be indirectly associated.[219,224]

A candidate gene approach explores and selects genes and polymorphisms based on a prior hypothesis that the genes or polymorphisms have a potential biological role in the phenotype being researched.[225] It is a hypothesis driven approach. The genes selected are generally related to the biology of the condition or similar conditions and therefore have a higher probability of being associated with the phenotype of interest than any other random gene present in the genome.[226]

The proteoglycans act as both structural and signalling molecules in ligaments, regulating various biological pathways that include matrix remodelling.[158] Variations in the genes encoding proteoglycans may potentially alter the biomechanical properties of ligaments, thus contributing to ACL injury susceptibility. After investigating the proteoglycans present in ligaments and tendons (Table 1.3) aggrecan, biglycan, decorin, fibromodulin and lumican appeared to be biologically relevant to ACL pathology, and the genes encoding these proteins were therefore chosen as candidates to be investigated in this study for an association with ACL injury in a South African population.

1.5.1 Aggrecan

Aggrecan, a large proteoglycan, is one of the major components of cartilage, functioning in cartilage elasticity and compressibility.[227] The core protein of aggrecan has a molecular mass of approximately 230 kDa and can have multiple GAG chains attached to it. Three globular domains, G1, G2, and G3, and two GAG attachment domains (KS and CS domains) found between the G2 and G3 domains comprise the core protein of aggrecan. An interglobular domain lies between G1 and G2.[227] The coding sequence of the globular domains shows high homology among species, while the KS and CS GAG attachment domains are diverse.[228]

Aggrecan plays a major structural role in ligaments and the tensional regions of tendons, and regulates cartilage development, growth and homeostasis.[169,210] The gene encoding aggrecan (*ACAN*) (Figure 1.5) lies on chromosome 15q26.1. It is 70.89 Kb in length, contains 16 exons and has 9 possible transcripts.[229] A missense pathogenic mutation within the aggrecan gene (rs121913568) has been implicated in osteochondritis dissecans and spondyloepimetapheseal dysplasia.[159,160]

Several other polymorphisms within the aggrecan gene have previously been associated with (i) sickle cell anaemia,[230] (ii) height in individuals of European [231] and African ancestry,[232] (iii) lumbar disc degeneration [233] and (iv) high myopia in individuals from China.[234]

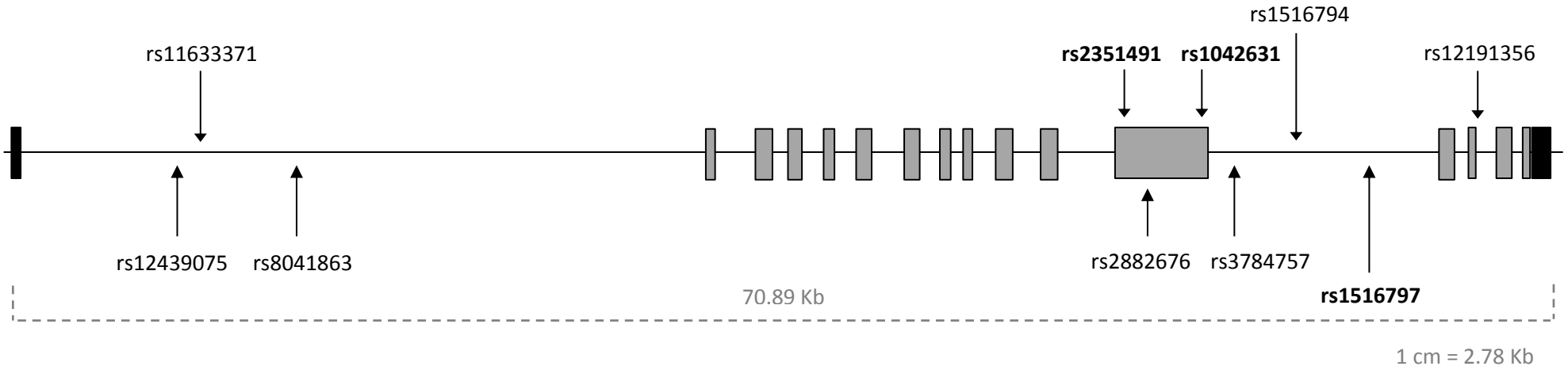


Figure 1.5: Schematic diagram of the aggrecan gene (*ACAN*) with some of the common polymorphisms shown. The grey boxes indicate the exons, and the horizontal lines that connect the exons represent the introns. The black boxes represent the untranslated regions. The polymorphisms chosen for investigation in this dissertation (rs2351491, rs1042631 and rs1516797) are indicated in bold. (Bioinformatic resources ncbi.nlm.nih.gov, ensembl.org and pga.gs.washington.edu (SeattleSNPs) were used to create this Figure).

1.5.2 Biglycan

Biglycan is a 38 kDa SLRP consisting of a LRR protein core which is covalently bound to two negatively charged CS or DS GAGs on serine 42 and 47.[163,235] This proteoglycan is expressed in skin, cartilage, bone and other connective tissues including ligaments.[41] When biglycan is localised in bone and dentin matrices, CS GAGs are predominantly bound, however when found in soft connective tissues such as skin, tendons and ligaments biglycan is bound by DS GAGs.[236] The function of biglycan is to regulate collagen fibrillogenesis and structure by interacting with collagen fibrils;[195] this allows biglycan to maintain the ECM structure of bone and connective tissues.[237] This proteoglycan is believed to influence the viscoelasticity of tendons,[238] as well as osteoblast function and positively regulate bio-mineralisation through processes that alter the activity of growth factors.[239]

The mechanism of action of the biglycan protein or its GAGs at the pathological level is not clear. Biglycan can however interact with growth factors such as TGF-β which suggests its involvement in modulating growth factor availability to cells.[189] Biglycan expression is increased by TGF-β [240] and biglycan itself, dose-dependently decreases the expression of TGF-β.[241] This proteoglycan also acts as an endogenous ligand for the toll-like receptors (TLRs) and is capable of producing a pro-inflammatory response.[242]

The gene encoding biglycan (*BGN*) lies on chromosome Xq28. It is 14.62 Kb in length, contains 8 exons and has 7 possible transcripts (Figure 1.6).[243] The biglycan promoter has several Sp1 transcription factor binding sites and several transcription start sites. There are also six interleukin (IL)-6 response elements within the biglycan promoter that generally induce biglycan expression when IL-6 is bound.[244] Biglycan expression is decreased by the binding of tumor necrosis factor-α (TNF-α) to the promoter region.[245] Biglycan mRNA levels and proteoglycan biosynthesis are invariably up-regulated by TGF-β.[246,247] Although biglycan is on the X chromosome, it is suggested to behave like a pseudo-autosomal gene that escapes X inactivation and has an active copy on the Y chromosome. There is however also evidence that biglycan does undergo X inactivation, and does not have a Y chromosome homologue.[248] The pseudo-autosomal expression of biglycan could be attributed to a gene or several genes adjacent to biglycan that escape X inactivation and control the transcriptional activity of the gene.[248] Polymorphisms within the biglycan gene have not yet been implicated in any multifactorial phenotypes.

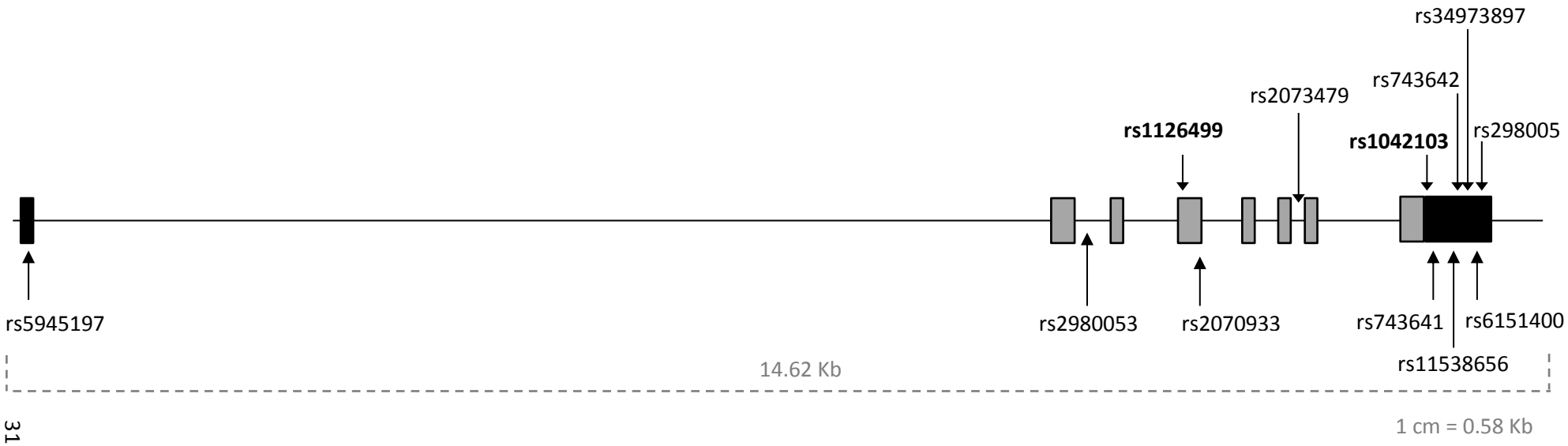


Figure 1.6: Schematic diagram of the biglycan gene (*BCN*) with some of the common polymorphisms shown. The grey boxes indicate the exons, and the horizontal lines that connect the exons represent the introns. The black boxes represent the untranslated regions. The polymorphisms chosen for investigation in this dissertation (rs1126499 and rs1042103) are indicated in bold. (Bioinformatic resources ncbi.nlm.nih.gov, ensembl.org and pga.gs.washington.edu (SeattleSNPs) were used to create this Figure).

1.5.3 Decorin

The major proteoglycan found in tensile regions of fibrous connective tissues such as tendons, ligaments and synovial capsule is decorin.[196] Decorin is a SLRP with a molecular mass of approximately 40 kDa. This proteoglycan has important roles in tissue development and assembly, and binds to collagen to regulate collagen fibrillogenesis.[197] Decorin functions in maintaining tissue integrity by binding fibronectin and thrombospondin. This proteoglycan also regulates collagen degradation,[249] cell growth [250] and extracellular signalling.[251] The LRRs of the core protein of decorin facilitate its non-covalent binding to the surface of fibrillar collagens, predominantly type I collagen, where it can retard the rate and degree of collagen fibrillogenesis.[252-255] The binding of decorin to collagen V has either a direct effect on the collagen, or mediates the interaction of collagen V with other molecules.[256] Fibril-collagen interactions are highly important with regards to the ligaments response to loading.[257] Decorin can also bind collagens II, III, VI and XIV and can thus affect a variety of ECM components.[258] The GAG chains of decorin function in maintaining the inter-fibrillar spacing of the collagen fibrils,[252] and the coordinated expression of decorin and associated collagen fibrils may therefore regulate ordered matrix assembly.[43]

Decorin can bind all three isoforms of TGF-β,[259] sequestering the molecule in the ECM.[260] This proteoglycan is also involved in cellular proliferation and the cell cycle,[261] angiogenesis [262] and apoptosis.[263] Decorin can also act as a damage-associated molecular pattern (DAMP), interacting with the immunity receptors TLR 2 and 4, therefore regulating a sterile inflammatory response within connective tissue.[264]

The gene encoding decorin (*DCN*) (Figure 1.7) lies on chromosome 12q21.33. It is 37.88 Kb in length, contains 8 exons and has 22 possible transcripts.[265] It is interesting to note that *DCN* gene expression can be regulated via several cell-signalling responses such as TNF-α inducible nuclear proteins, IL-1β,[266] IL-I, IL-4 [267,268] and TGF-β.[269] The collective evidence is suggesting that decorin plays a critical role in matrix assembly regulation through its myriad of potential interactions with ECM components. It is therefore not surprising that several signalling mechanisms are in place to tightly regulate the synthesis of this proteoglycan in response to the ever changing cellular environment. Interestingly, there have been no published reports describing polymorphisms within the *DCN* gene to be implicated in any multifactorial phenotypes although various studies have investigated this gene.[270-272]

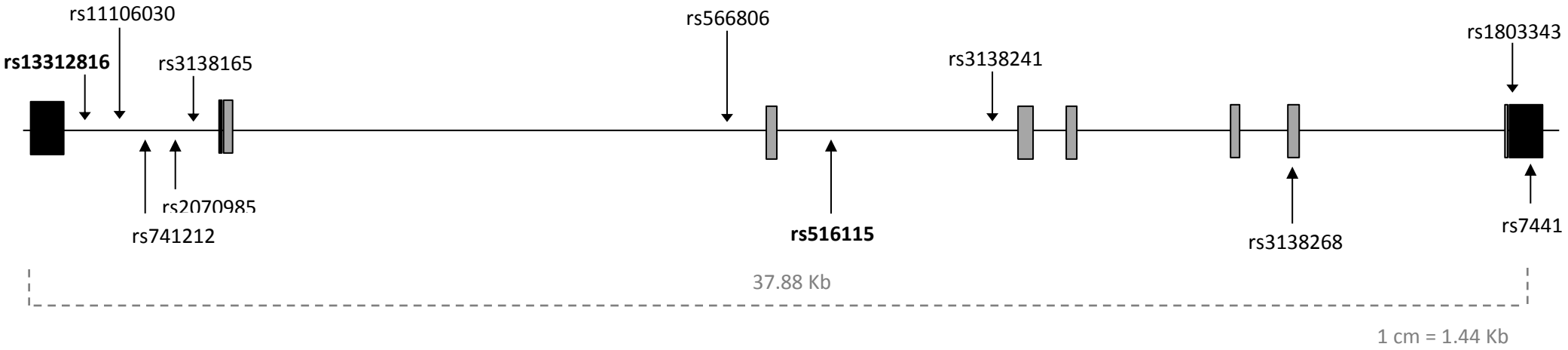


Figure 1.7: Schematic diagram of the decorin gene (*DCN*) with some of the common polymorphisms shown. The grey boxes indicate the exons, and the horizontal lines that connect the exons represent the introns. The black boxes represent the untranslated regions. The polymorphisms chosen for investigation in this dissertation (rs13312816 and rs516115) are indicated in bold. (Bioinformatic resources ncbi.nlm.nih.gov, ensembl.org and pga.gs.washington.edu (SeattleSNPs) were used to create this Figure).

1.5.4 Fibromodulin

Fibromodulin is expressed in collagenous connective tissues, including ligaments, and plays a role in defining tissue integrity [199] and regulating collagen fibrillogenesis.[200] As a SLRP, fibromodulin is important for maintaining tissue homeostasis, as well as for tissue repair and regeneration in normal and pathological conditions.[192]

The gene encoding fibromodulin (*FMOD*) lies on chromosome 1q32. It is 10.86 Kb in length, contains 3 exons and has 6 possible transcripts (Figure 1.8).[273] Genetic association studies have not yet implicated the fibromodulin gene in any multifactorial phenotypes.

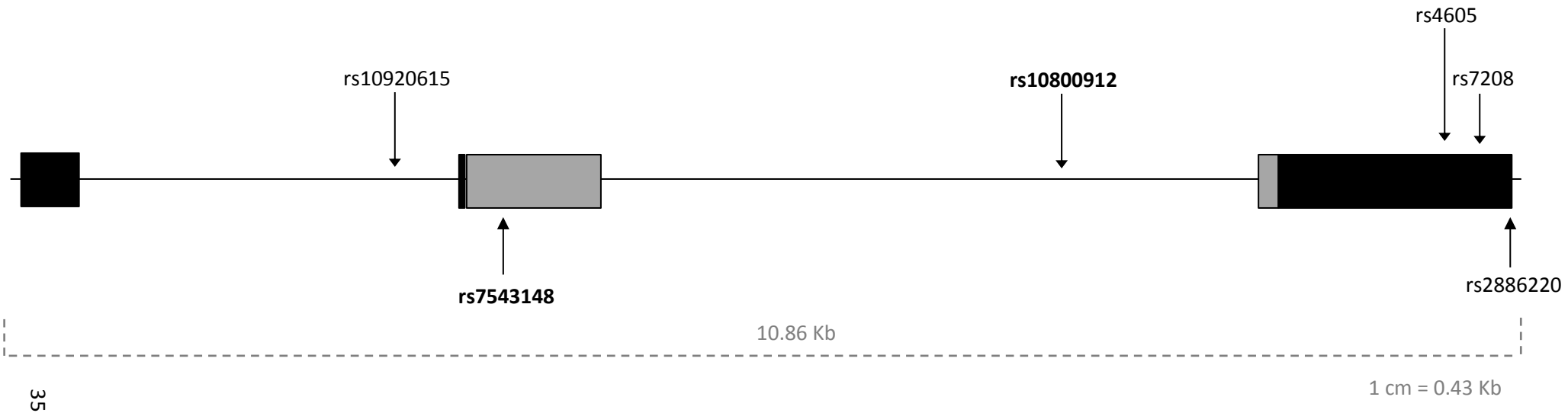


Figure 1.8: Schematic diagram of the fibromodulin gene (*FMOD*) with some of the common polymorphisms shown. The grey boxes indicate the exons, and the horizontal lines that connect the exons represent the introns. The black boxes represent the untranslated regions. The polymorphisms chosen for investigation in this dissertation (rs7543148 and rs10800912) are indicated in bold. (Bioinformatic resources ncbi.nlm.nih.gov, ensembl.org and pga.gs.washington.edu (SeattleSNPs) were used to create this Figure).

1.5.5 Lumican

Lumican is a SLRP,[274] which is expressed in various different tissues including skin,[275] arteries,[276] intervertebral discs,[277] bone,[278] aorta,[279] articular cartilage,[280] tendons and ligaments.[172,178] Lumican has an important role in collagen fibrillogenesis [199,275] as well as functioning in tissue homeostasis by regulating cellular proliferation,[281] migration [281,282] and adhesion.[282] The core protein of lumican has a molecular weight of approximately 40 kDa and like most other SLRPs, has three major domains: a negatively charged N-terminal domain; tandem LRRs adapted for protein–protein interactions (where lumican can interact with fibrillar collagens and modulate fibril formation); and a C-terminal domain containing two conserved cysteine residues.[283]

The gene encoding lumican (*LUM*) lies on chromosome 12q21.3-q22. It is 9.2 Kb in length, contains 3 exons and has 3 possible transcripts (Figure 1.9).[284] This gene has previously been associated with high myopia susceptibility in a Han Chinese population,[285] as well as being associated with susceptibility to pathological myopia in the Northern Han Ethnic Chinese.[286] A particular polymorphism (rs2268578) has been associated with breast cancer [271] as well as the risk of pancreatic cancer in a Caucasian population.[287]

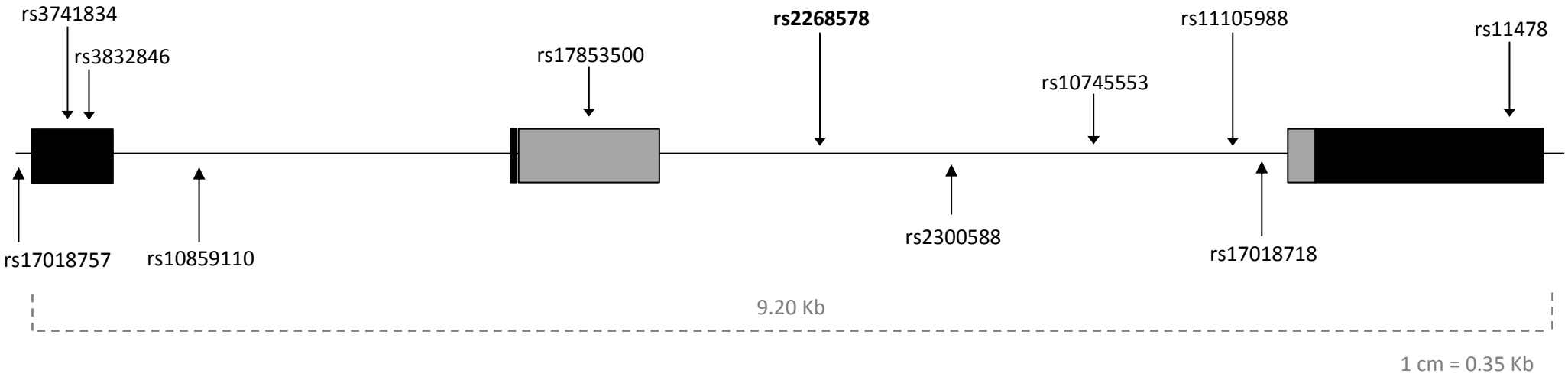


Figure 1.9: Schematic diagram of the lumican gene (*LUM*) with some of the common polymorphisms shown. The grey boxes indicate the exons, and the horizontal lines that connect the exons represent the introns. The black boxes represent the untranslated regions. The polymorphism chosen for investigation in this dissertation (rs2268578) is indicated in bold. (Bioinformatic resources ncbi.nlm.nih.gov, ensembl.org and pga.gs.washington.edu (SeattleSNPs) were used to create this Figure).

1.6 The structural role of proteoglycans

Proteoglycans are a diverse group of molecules that function in both cell signalling pathways and in maintaining the structural integrity of tissues.[162] The proteoglycans form part of the ECM, which is an integral part of any connective tissues ability to withstand mechanical loads. The multivalent binding capabilities of proteoglycans assist the ECM by attaching to and stabilising collagen fibrils in an organised manner.[255] The interaction of DS proteoglycans with collagen fibrils leads to increased ECM stability,[288] and the negative charge of SLRPs attract water preventing fusion of the fibrils and segregation of the extra-fibrillar molecules.[189] Fibril-proteoglycan interactions are one of the important interactions that take place during a connective tissues response to loading (Figure 1.10).[289]

One of the important functions shared by several SLRPs is the regulation of collagen fibrillogenesis. The SLRPs are key regulators of collagen fibril assembly as they can delay fibril formation.[44,290] Type I and type V collagens are closely regulated by SLRPs during fibrillogenesis; the proteoglycans interact with the collagen fibres, forming the basis of the structural organisation of the tissue.[254,291] The SLRPs regulate collagen fibrillogenesis by directly interacting with the collagens via their central domains,[44] and aggrecan functions in maintaining the structural integrity of ligaments by associating with the collagen network, although the interaction of aggrecan with the collagen network is not fully understood (Figure 1.10).[195,197,256] Decorin influences collagen fibril diameter [292] and shape by restricting lateral growth,[44] and biglycan inhibits collagen fibrillogenesis in vitro, although the interaction of biglycan with collagens is not yet well characterised. During fibrillogenesis, both lumican and fibromodulin inhibit lateral growth, fibril thickness, and may also influence fibril number.[290] The GAGs of SLRPs are involved in the regulation of fibrillogenesis and also impact inter-fibrillar spacing and organisation during matrix assembly.[40]

Knock-out mice models have demonstrated the important instructive roles of proteoglycans in fibrillogenesis and matrix assembly. Unfavourable structural alterations of the collagen fibrils are observed in mice deficient in one or two of the predominant SLRPs (biglycan, decorin, fibromodulin or lumican); this frequently results in impaired connective tissue function.[42,43,199,239,290]

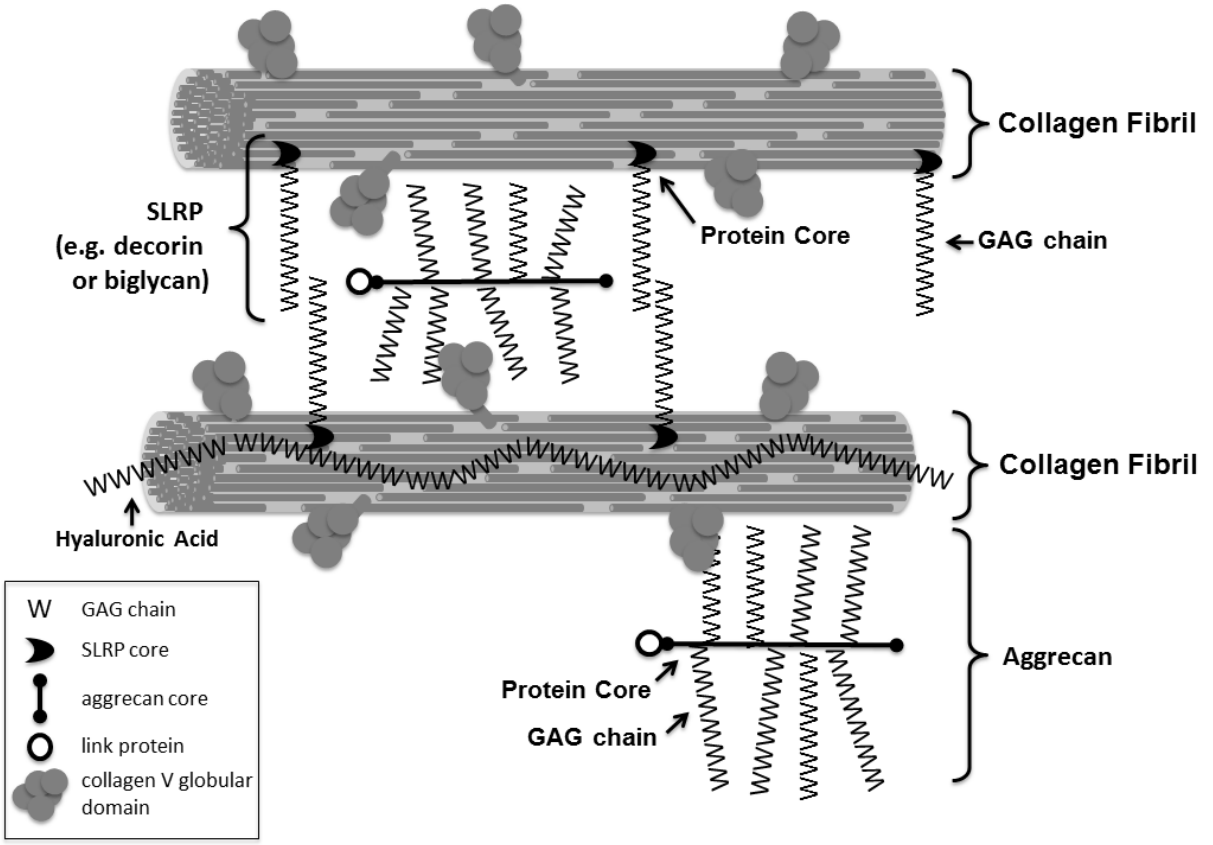


Figure 1.10: A schematic diagram representing the association of aggrecan and the small leucine-rich proteoglycans (SLRPs) biglycan, decorin, fibromodulin and lumican, as well as the glycosaminoglycans (GAGs), with the collagen network.

Murine tissues and mice that lack SLRPs such as biglycan, decorin, fibromodulin, and lumican, have similar physical phenotypes to humans with classic EDS.[41-43] EDS is a genetic disorder found in humans, the classic form of which is caused by the disruption of the production of type V collagen. Similarities between humans with EDS and SLRP knockout tissues include abnormal morphology of the collagen fibrils which results in loss of mechanical strength, joint laxity and skin fragility.[41,43,199] Mice deficient in the biglycan gene have connective tissue disorders including collagen fibril abnormalities in skin, tendon and bone, OA, low bone mass and ectopic tendon ossification.[41,42,293] Decorin knockout mice (decorin deficient) display collagen fibrils with highly irregular shapes/diameters and abnormal fibrillar organization that is loosely packed due to abnormal lateral fusion.[43] This disruption can lead to pathology. The decorin knockout mice also show decreased skin and tendon tensile strength due to the abnormal morphology of the fibrils, which results in loss of mechanical strength and the reduction of collagen bound proteoglycans.[43,197,294] This demonstrates the vital role of decorin in regulating collagen fibre formation. In vivo, periodontal ligaments from decorin deficient mice have been found to have a 2 fold increase in ligament fibroblasts in

comparison to tissues from wild type mice.[295] Decorin deficient mice also show an increased ligation of TGF-β to its receptors.[296] Following tendon injury, biglycan deficient mice have impaired initial healing while decorin deficient mice show diminished healing later on. Mice deficient in these two proteoglycans also have modifications to collagen alignment and fibril structure during injury response.[297]

Knockout mice that are fibromodulin deficient have tendons with collagen fibrils of small diameters.[200,290] They also have an increase in age dependent OA and degeneracy of the articular cartilage.[298] Fibromodulin knock-out mice highlight the importance of this proteoglycan in maintaining tissue integrity as these mice have abnormal morphology of the collagen fibrils, which results in loss of mechanical strength, joint laxity and skin fragility.[41,43,199] Irregular fibril organisation observed in fibromodulin knock-out mice highlights that this proteoglycan is responsible for regulating fibril thickness.[290]

Knockout mice that do not have a functioning lumican gene show abnormal connective tissue phenotypes that include skin laxity and fragility, highlighting the crucial role of lumican in the regulation of collagen fibrillogenesis.[199,275]

The cornea, like ligaments and tendons, is composed of a closely regulated collagen network. Eye abnormalities have been genetically related to mutations in SLRPs. Mutations in the human decorin gene have been associated with congenital stromal corneal dystrophy [299] due to abnormal organisation of the corneal collagen fibrils.[300] Loss of function mutations in either the lumican, fibromodulin, PRELP or optican proteoglycan genes have been implicated in myopia.[163,301] Further studies have shown that lumican knock-out mice have cloudy corneas due to the disruption of the collagen organisation,[275] whereas keratocan deficient mice develop corneal collagen fibrils with larger diameters.[302]

An important concept must however be noted, that is overlapping functions between the SLRPs, such that one SLRP can compensate for the loss of function of another SLRP. For example, in the absence of fibromodulin, lumican is up-regulated,[290] and decorin accumulates in the absence of biglycan.[42] Gene target studies have shown that re-expression of lumican in the cornea can rescue ocular abnormalities in lumican knock-out, or fibromodulin-lumican double knockout mice.[275,303]

One can therefore speculate that this observed biological redundancy in the functions of the SLRPs may be reflecting the significant biological role they play in regulating fibrillogenesis and maintaining the structural properties of tissues. The SLPRs, like the collagen genes, can therefore be considered as candidate genes for genetic association studies.

1.7 Proteoglycans in the injured state

Proteoglycans have a critical role in the pathophysiology of basement membrane related diseases including diabetes, arteriosclerosis and metastasis.[304] Changes in proteoglycan expression and the accumulation of GAGs have also been associated with tendon pathology.[11,14] Increased levels of proteoglycans in the ECM of tendinopathic tendons are thought to influence the increased hydration and swelling of the tissue observed.[305] Mechanical loading to the deep flexor tendons from the hind legs of near-term bovine foetuses leads to an increase in the mRNA expression of aggrecan, versican, biglycan and decorin, as well as an increase in TGF-β1 mRNA levels.[306] In these same tendons, the addition of TGF-β producing cells leads to an increase in aggrecan and biglycan expression as well as a slight increase in α1 collagen production. Expression of decorin was decreased slightly.[306] The addition of TGF-β to the tendons induced similar changes to that of mechanical loading. This once again highlights the role of proteoglycans, with TGF-β, in regulating ECM remodelling following mechanical loading. Another study by Attia *et al.* (2012) detected increased levels of the proteoglycans aggrecan, versican, biglycan and decorin, as well as GAG content in rat supraspinatus tendons that were subjected to overuse.[307] This same group also recently associated tendon aggrecan and GAG content with the degree of tendon pathology. This suggests that proteoglycans and GAGs are potential markers of injury progression.[308] Parkinson *et al.* (2010) have also shown an increased synthesis and loss of proteoglycans from the ECM of abnormal (tendinopathic) tendons in comparison to normal controls, highlighting the differential proteoglycan turnover between the tissues.[309]

In 2011 Young *et al.* investigated the ECM content of ruptured ACL tissue.[169] They observed lower levels of both small and large proteoglycans as well as GAG content in ruptured human ACL tissue in comparison to the non-ruptured controls. It is however unknown whether the low levels of proteoglycans altered the composition and structure of the ACL ECM, reducing the ability to withstand a load and thus causing the injury, or whether the changes in proteoglycans occurred after and are a result of the injury.[169] Turunen *et al.* (2013) also recently observed reduced proteoglycan content in the cartilage of ACL transected rabbit knee joints.[310]

Aggrecan has been found to increase in the joint fluid of canines following an ACL injury by unilateral surgical transection.[311] It thus appears as though aggrecan is lost from the ACL tissue [169] to the joint fluid following an injury. A significant increase in aggrecan was seen in the joint fluid of canines three and twelve weeks after transection surgery.[311] Nelson *et al.* (2006) have also shown that one year

after ACL injury, GAG content, of which the majority is assumed to be aggrecan, is significantly increased in the femoral articular cartilage of humans. After one year, proteoglycan content is still increased but not significantly.[182] This is in line with previous studies by Lohmander *et al.* (1991, 1999) which showed an increase of aggrecan in the joint fluid following injury.[312,313] The same study by Nelson *et al.* (2006) identified a direct correlation between proteoglycan/GAG content and total collagen content of cartilage tissue.[182]

Increased synthesis and accumulation of biglycan in the ECM is phenotypic of fibrotic conditions,[314-316] including hypertrophic scar [317] and atherosclerotic plaque formation.[318] There are contrasting results regarding biglycans expression within the ACL following an injury. Lo *et al.* (1998, 2003) revealed that injured/disrupted ACLs express higher quantities of biglycan, and levels remain elevated for longer than one year after injury,[319,320] whereas Young *et al.* (2011) reported lower levels of biglycan in ruptured ACL tissue in comparison to control tissue.[169]

Provenzano *et al.* (2005) have previously shown an increase in mRNA expression of decorin and fibromodulin seven days after a ligament sprain injury.[321] Unchanged levels of decorin have however also been observed in the comparison of ruptured and non-ruptured ACL samples.[169] Decorin does however seem to be more complicated because the core protein bound by GAGs increases inflammatory signalling,[264] whereas the core protein alone is anti-inflammatory.[322]

Although the knowledge of the role of proteoglycans in ACL ruptures is limited, the above literature review provides reasonable evidence to choose the genes encoding proteoglycans as candidates to be explored in a genetic association study using an ACL injury risk model.

Chapter 2: Genes encoding proteoglycans are associated with the risk of anterior cruciate ligament ruptures

2.1 Introduction

As discussed in the previous chapter, although the aetiology of the molecular mechanisms is poorly understood, multiple intrinsic and extrinsic risk factors, including genetics, have been associated with ACL ruptures (Chapter 1, Section 1.3).[9,34-39] Polymorphisms within several genes encoding collagens, implicated in the regulation of fibrillogenesis, have been associated with ACL ruptures (Section 1.3.2).[35-38] Similar to the collagens, the proteoglycans aggrecan, biglycan, decorin, fibromodulin and lumican have important structural roles in ligaments and also play an essential role in regulating fibrillogenesis (Section 1.6).[40]

Mutations within the large *ACAN* gene cause either dominant familial osteochondritis dissecans, or a recessive skeletal dysplasia in humans.[159,160] Murine tissues and mice deficient in the SLRPs biglycan, decorin, fibromodulin, or lumican, have similar physical phenotypes to humans with classic EDS: fibrillogenesis is compromised resulting in collagen fibrils of highly irregular diameters and abnormal fibrillar organisation.[41-43] Young *et al*. (2011) recently investigated the ECM content of ruptured ACL tissue. Lower proteoglycan and GAG levels were observed in ruptured human ACL tissue in comparison to the non-ruptured controls.[169] Genes encoding proteoglycans are therefore plausible candidate genes to be investigated for an association with ACL injury risk.

Variants within the *ACAN* and *LUM* genes, on chromosomes 15q26.1 and 12q21.3 respectively, have previously been associated with several multifactorial conditions.[231-234,285,286] The genes encoding *BGN* on chromosome Xq28, *DCN* on chromosome 12q21.33, and *FMOD* on chromosome 1q32, have however not been associated with any multifactorial conditions to date.

This chapter aimed to investigate the association of polymorphisms in the *ACAN*, *BGN*, *DCN*, *FMOD* and *LUM* candidate genes with ACL ruptures, and specifically non-contact ACL ruptures, based on the important biological functions of these five proteoglycan encoding genes in maintaining the structural integrity of tissues and regulation of fibrillogenesis. More importantly, the study aimed to identify genomic regions encompassing these five genes which may be harbouring DNA sequence signatures relevant to our understanding of ACL injury susceptibility. In addition, we aimed to investigate if there was any contribution of these variants to sex-linked susceptibility, as has been previously noted with ACL ruptures.[37,38]

2.2 Materials and Methods

The reporting of this case-control genetic association study is in alignment with the recommendations outlined by the STREGA (STrengthening the REporting of Genetic Association studies) initiative, which is an extension of the STROBE (STrengthening the Reporting of OBservational Studies in Epidemiology) statement.[323]

2.2.1 Participants

A total of 461 physically active, unrelated, self-reported Caucasian participants were recruited for this case-control genetic association study. These participants consisted of 227 (166 male) participants with surgically diagnosed ACL ruptures (ACL group) recruited from the Sports Science Orthopaedic Clinic in Cape Town, South Africa, and 234 (144 male) apparently healthy participants, without any history of ACL injuries (CON group) that were recruited as controls from sports clubs and wellness centres within Cape Town, South Africa. Recruitment of participants took place between the years of 2006 and 2013. Previously described inclusion and exclusion criteria were used.[37]

All participants were given information about the study (Appendix A) and were required to complete a written informed consent form according to the declaration of Helsinki (Appendix A). Participants were also requested to complete a questionnaire regarding personal details, medical history, personal and family ligament and tendon injury history, as well as sports participation (Appendix A). All participants were of self-reported Caucasian ancestry.

Participants within the ACL group could be classified into subgroups by mechanism of injury according to the American Orthopaedic Society for Medicine classification system.[324] These subgroups were direct contact, indirect contact, non-contact, skiing accident or trauma (for example motor vehicle accident). There was insufficient evidence to identify the exact mechanism of injury of 22 participants (9.69%, 16 male). The most common mechanism of ACL injury was non-contact (NON). This subgroup consisted of 126 participants (61.46%, 94 male) that were analysed as part of the ACL group and underwent additional statistical evaluation separately.

Sports participation of the CON and ACL groups was characterised into contact sports, non-contact jumping sports, non-contact non-jumping sports and skiing sports as previously defined,[9] with slight

modification.[37] The most common sports played by the ACL group males were rugby (45.2%) and soccer (8.4%), whilst the ACL group females played predominantly hockey (26.2%) and netball (14.8%).

This study was approved by the Research Ethics Committee of the Faculty of Health Sciences within the University of Cape Town, South Africa (reference number 164/2006) (Appendix B).

2.2.2 DNA extraction

Approximately 5ml of venous blood was obtained from each participant by venepuncture of a forearm vein and collected into an ethylenediaminetetraacetic acid (EDTA) vacuum container tube. Blood samples were stored at -20˚C until total DNA extraction was performed as previously described by Lahiri and Nurnberger (1991),[325] and modified by Mokone *et al*. (2005).[326] (Appendix C)

2.2.3 Single nucleotide polymorphism selection and Genotyping

SNPs within each of the five candidate genes were identified using the genome database hosted by the National Centre for Biotechnology Information (NCBI) (http://ncbi.nlm.nih.gov/) as well as SeattleSNPs (http://pga.gs. washington. edu/). SNPs were selected based on previous associations, whether they were Tag SNPs and should thus provide moderate coverage of the genetic interval,[270,271] or if the SNP had a heterozygosity score greater than 30%.

Three SNPs were investigated in *ACAN* (Figure 1.5) and included: (i) rs2351491 23 C>T, previously associated with height in individuals of African ancestry,[232] (ii) rs1042631 4157 T>C and (iii) rs1516797 -1133 T>G, with the latter two SNPs being previously associated with lumbar disc degeneration.[233] The two SNPs investigated within *BGN* (Figure 1.6) included rs1126499 189 C>T and rs1042103 359 G>A, both of which were previously investigated for an association with congenital muscular dystrophy but excluded as a cause of the disorder.[327] *DCN* rs13312816 IVS1 A>T and rs516115 IVS3 A>G, are both Tag SNPs [270,271] and should therefore provide a large coverage of the genomic region spanning approximately 18 Kb of the *DCN* gene (Figure 1.7). Rs7543148 244 G>A, with a heterozygosity score greater than 30%, and rs10800912 -1338 C>T, a Tag SNP, were chosen for investigation within the *FMOD* gene. *LUM* rs2268578 697 T>C, a Tag SNP,[271,272] has previously been

associated with multifactorial phenotypes.[271,287] Schematic diagrams of the *FMOD* and *LUM* genes are given in Figure 1.8 and Figure 1.9 respectively. Nomenclature used to describe the SNPs investigated is in accordance with the NCBI (http://ncbi.nlm.nih.gov/).

TaqMan® allele-discrimination assays (Applied Biosystems, Foster City, California, USA) were used to genotype participants for the ten SNPs. Previously inventoried TaqMan® primer sets and allele specific MGB-labelled probes were used together with the PCR master mix, containing ampliTaq® DNA polymerase Gold (Applied Biosystems, Foster City, California, USA), as per manufacturers recommendations in a final reaction volume of 8µl. The PCR reactions were performed on an Applied Biosystems StepOnePlus™ Real-Time PCR system (Applied Biosystems, Foster City, California, USA) using the Applied Biosystems Step-OnePlus™ Real-Time PCR software Version 2.2.2 (Applied Biosystems, Foster City, California, USA). The PCR parameters comprised a 30 second hold step at 60˚C followed by a 10 minute heat activation step at 95˚C, 40 cycles of 95˚C for 15 seconds and 60˚C for 1 minute, ending with a 30 second hold step at 60˚C. Genotypes were determined by endpoint fluorescence. For PCR and genotype quality control purposes, a number of positive (known genotypes) and DNA-free controls were randomly included on every 96-welled PCR plate. The DNA free controls were successful 100% of the time and the positive controls were successful 99.3% of the time. The genotypes of each of the samples were checked by an independent person to guarantee accuracy.

All laboratory work, DNA extraction and sample genotyping, took place at the UCT/MRC Research Unit for Exercise Science & Sports Medicine Laboratory, Faculty of Health Sciences, University of Cape Town. Asanda Mtintsilana is acknowledged for the genotyping, in part, of the SNPs within the *FMOD* and *LUM* genes.

2.2.4 Statistics

Quanto Version 1.2 (http://hydra.usc.edu/gxe) was used to determine the statistical power of the sample size. Assuming allele frequencies between 0.1 and 0.9 for the “risk” allele of each SNP investigated, our sample size of 227 cases would be adequate to detect an allelic odds ratio of 1.8 and greater, at a power of 80% and a significance level of 5%.

Genotype and allele frequencies were analysed using Statistica Version 11 (StatSoft Inc, Tulsa, Oklahoma, USA) and GraphPad InStat Version 5 (GraphPad software, San Diego, California, USA). The *BGN* gene is on the X chromosome and therefore genotype and allele frequencies of SNPs investigated in this gene were compared separately between male and female participants. One-way analysis of variance (ANOVA) was used to compare continuous biological characteristics between the CON and ACL groups and between the CON group and NON subgroup. Chi-squared and Fisher exact tests were used to compare categorical variables (sex and country of birth) between the CON group, ACL group and NON subgroup, as well as to analyse any differences in genotype and allele frequencies between the groups. A Fisher Exact test was used to compare variables between groups when $n<10$ for one of the groups, and a Chi-squared test was used when $n>10$ for both of the groups being examined. Inferred haplotypes were constructed for the *ACAN*, *BGN*, *DCN* and *FMOD* genes using the specific SNPs investigated within each gene. A haplotype was also constructed to overlap the *LUM-DCN* genetic interval (12q21.3-12q21.33) using the SNPs investigated within these genes. The Chaplin case-control haplotype inference software program Version 1.2.2 (http://www.genetics.emory.edu/labs/epstein/software/chaplin/index.html) was used to compare allele frequencies of the variants within each haplotype between cases and controls. CubeX: cubic exact solution (http://www.oege.org/software/cubex/)[328] was used to determine which of the SNPs investigated are in LD and thus likely to be inherited together. Significance was accepted at $p<0.05$. In order to determine whether the genotypes obtained for each of the SNPs investigated were in Hardy-Weinberg equilibrium (HWE), the data was analysed using Genepop Version 4.2 (http://genepop.curtin.edu.au/).

2.3 Results

2.3.1 Participant characteristics

There were significantly more males in the ACL group (p=0.008) and NON (p=0.013) subgroup in comparison to the CON group (Table 2.1). Participants within the CON group, ACL group, and NON subgroup, were similarly matched for critical BMI (BMI at recruitment for CON group) (p=0.107 and p=0.141 respectively) and country of birth (COB) (p=0.769 and p=0.814 respectively). The ACL group (p=0.019) and NON subgroup (p=0.028) were significantly younger than the CON group. When co-varied for the differences in sex and age, the ACL group (p=0.002) and NON subgroup (p=0.011) still weighed significantly more than the CON group. When co-varied for sex, height did not differ significantly between groups (p=0.231 and p=0.153 respectively).

Table 2.1: Characteristics of the asymptomatic control group (CON), the ACL rupture group (ACL), and the ACL subgroup with a non-contact (NON) mechanism of injury.

	CON (n = 234)	ACL (n = 227)	p Value †	NON (n = 126)	p Value ‡
Sex, % males	61.5 (234)	73.1 (227)	**0.008**	74.6 (126)	**0.013**
Age, years	29.3 ± 11.3 (228)	26.8 ± 11.0 (198)	**0.019**	26.6 ± 10.5 (120)	**0.028**
Weight, kg	74.0 ± 14.7 (229)	80.2 ± 16.9 (206)	**<0.001 (0.002)**	79.8 ± 15.8 (123)	**<0.001 (0.011)**
Height, cm	175.0 ± 9.4 (228)	177.5 ± 9.3 (206)	0.006 (0.231)	178.0 ± 9.1 (122)	0.004 (0.153)
BMI, kg/cm^2	23.8 ± 4.1 (228)	24.6 ± 5.8 (206)	0.107	24.6 ± 5.4 (122)	0.141
COB, % SA	86.2 (225)	82.7 (208)	0.769	82.9 (123)	0.814

Data reported as mean ± standard deviation, except for sex and COB which are presented as frequency (%).
The number of subjects with available data for each variable is reported in parentheses.
† CON vs ACL
‡ CON vs NON
Age, weight and BMI are self-reported values at the time of first ACL rupture for the ACL group and the NON subgroup, and at time of recruitment for the CON group.
Un-adjusted p values are shown, with the exception of weight and height which have adjusted p values in parentheses. Weight has been adjusted for age and sex, and height has been adjusted for sex only.
Significant p values are noted in bold.
Abbreviations: BMI, body mass index; COB, country of birth; SA, South African.

Age and weight at the time of recruitment of the ACL group was 4.6 ± 8.9 years (n=225) older and 2.0 ± 12.1 kg (n=221) heavier than at the time of the first ACL rupture. In the NON subgroup, the age and weight at time of recruitment was 4.1 ± 7.3 years (n=126) older and 2.5 ± 11.0 kg (n=124) heavier than at the time of first ACL rupture.

The female ACL and CON groups were matched for participation in non-contact non-jumping sports (p=0.764) and skiing sports (p=0.526) (data not shown). The male ACL and CON groups were matched for non-contact non-jumping sports (p=0.138). Significantly more participants (male and female) within the ACL group, participated in contact sports (p=0.009 and p=0.001 respectively) and non-contact jumping sports (p=0.003 and p<0.001 respectively) in comparison to the control participants. In addition, significantly more participants within the ACL group participated in skiing sports in comparison to the male controls (p<0.001).

With the exception of a significant *ACAN* rs1042631 genotype effect on sex (p=0.033), there were no other genotype effects on participant characteristics (Appendix D, Supplementary Table 1).

2.3.2 *ACAN* gene

There were no significant differences in the genotype (Figures 2.1A and 2.1B) and allele (Appendix D, Supplementary Table 2) frequency distributions between the CON and ACL groups for the *ACAN* rs2351491 (p=0.547 and p=0.415) and rs1042631 (p=0.168 and p=0.064) SNPs. Similar results were noted when stratified by mechanism of injury (NON subgroup) and sex (Appendix D, Supplementary Table 2). There was a trend (p=0.059) for the TT genotype to be over-represented in the CON group (51.1%, n=119) when compared to the ACL group (42.3%, n=96) for rs1516797. Interestingly, the G allele of rs1516797 was significantly under-represented in the CON group (27.5%, n=128) (p=0.024; OR=0.72; 95% CI:0.55-0.96) in comparison to the ACL group (34.4%, n=156) (Appendix D, Supplementary Table 2). No significant differences (p=0.325) in the allele frequencies for rs1516797 were however noted between the CON and NON groups. The genotype and allele frequency distributions for all three SNPs were similar between the male and female participants for all groups (CON, ACL and NON). All the groups were in HWE for all three *ACAN* SNPs (Appendix D, Supplementary Table 2).

Only six of the possible eight haplotypes constructed from the three *ACAN* variants (rs2351491 C>T - rs1042631 T>C - rs1516797 T>G) had a frequency greater than 2%. The haplotype containing alleles T-C-T was significantly over-represented (p=0.001; LR=10.30) in the CON group (N=233, 43.55%) in comparison to the ACL group (N=225, 32.74%), while T-C-G was significantly under-represented (p=0.005; LR=7.79) in the CON group (N=233, 20.83%) in comparison to the ACL group (N=225, 29.08%) (Figure 2.1D). *ACAN* rs2351491 and rs1042631 were found to be in complete LD (D' = 1.0), while rs1042631 and rs1516797 were not in LD (D'=-0.784).

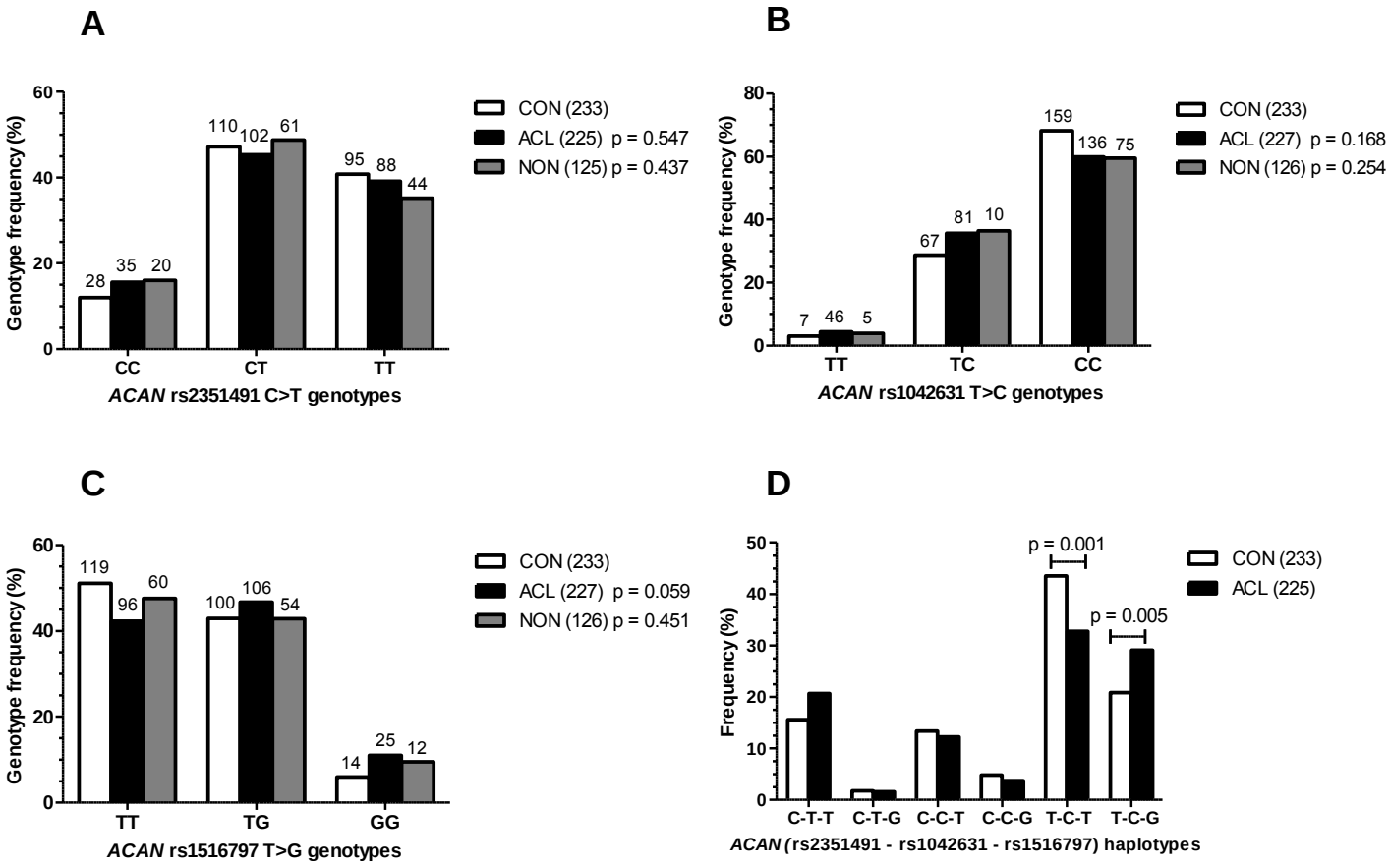


Figure 2.1: A-C: Genotype frequency distribution of *ACAN* rs2351491, rs1042631 and rs1516797 for the control (CON, white bars), anterior cruciate ligament rupture (ACL, black bars) and non-contact anterior cruciate ligament rupture (NON, grey bars) groups. **D:** Frequency distribution of the *ACAN* (rs2351491 C>T - rs1042631 T>C - rs1516797 T>G) inferred haplotypes amongst the control (CON, white bars) and anterior cruciate ligament rupture (ACL, black bars) groups. The p values of significantly different distributions are noted. The total number of participants with available genotype data within each group is indicated in parenthesis on the graph.

2.3.3 *BGN* gene

There were no significant differences in the genotype frequency distributions between the CON and ACL groups, or between the CON group and NON subgroup at the *BGN* rs1126499 or rs1042103 loci for either males (rs1126499: CON vs ACL, p=0.533; CON vs NON, p=0.672; rs1042103: CON vs. ACL, p=0.383; CON vs. NON, p=0.580) or females (rs1126499: CON vs ACL, p=0.105; CON vs NON, p=0.278; rs1042103: CON vs. ACL, p=0.226; CON vs. NON, p=0.584) (Figure 2.2A-D). Similarly, no significant differences in allele frequencies were noted (Appendix D, Supplementary Table 2). However, there was a trend for the *BGN* rs1126499 T allele to be under-represented (p=0.068) in the female CON group (48.3%, n=85) when compared to the female ACL group (59.0%, n=72). All the female groups were in HWE for both *BGN* SNPs.

There were no significant differences in the distribution of the inferred haplotypes constructed from the *BGN* variants (rs1126499 C>T - rs1042103 G>A) when only the male participants were compared between the CON and ACL groups (Figure 2.2E). However, when the female participants were compared, the *BGN* C-G inferred haplotype was significantly over-represented (p=0.027) in the CON group (N=88, 39.64%) in comparison to the ACL group (N=61, 27.15%) (Figure 2.2F). *BGN* rs1126499 and rs1042103 were found to be in low LD (D'=0.321).

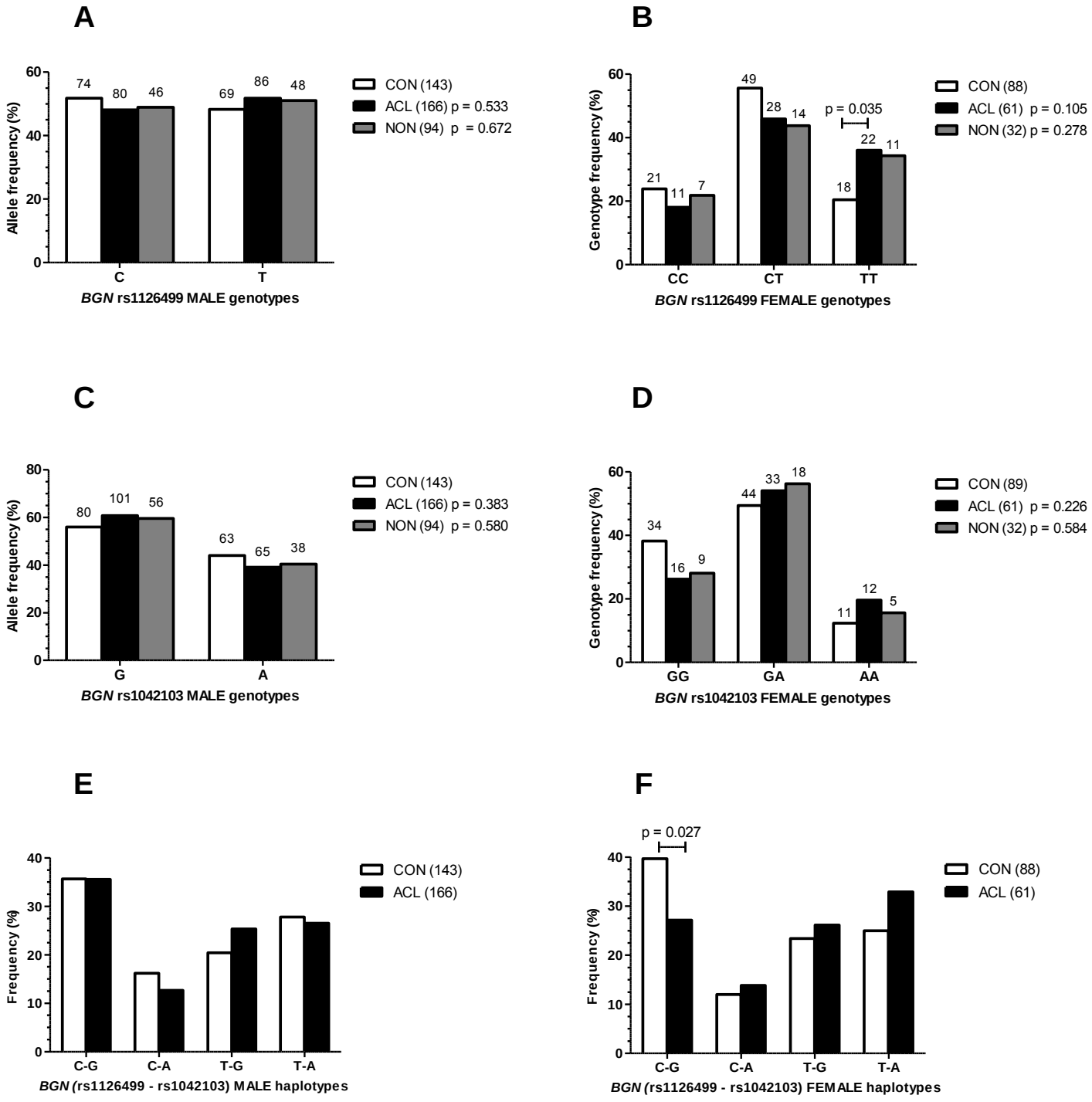


Figure 2.2: A-D: Genotype frequency distribution of *BGN* rs1126499 and rs1042103 for the control (CON, white bars), anterior cruciate ligament rupture (ACL, black bars) and non-contact anterior cruciate ligament rupture (NON, grey bars) groups for male and female participants. **E-F:** Frequency distribution of the *BGN* (rs1126499 C>T - rs1042103 G>A) inferred haplotypes amongst the control (CON, white bars) and anterior cruciate ligament rupture (ACL, black bars) groups for male and female participants. The p values of significantly different distributions are noted. The total number of participants with available genotype data within each group is indicated in parenthesis on the graph.

2.3.4 *DCN* gene

No significant differences in genotype frequencies were noted for the *DCN* rs13312816 (p=0.221) or rs516115 (p=0.926) SNPs when all participants (male and female) were analysed, or when only the male participants (rs13312816: p=0.256; rs516115: p=0.334) were analysed between the groups (Figure 2.3A and 2.3C); similarly, no significant differences in allele frequencies were noted (Appendix D, Supplementary Table 2). Furthermore, no significant difference in the genotype distribution of *DCN* rs13312816 was noted in females (p=0.214) (Figure 2.3B). However, the GG genotype of rs516115 was significantly over-represented in the CON group (13.3%, p=0.015; OR=9.23; 95% CI:1.17–73.01) when compared to the ACL group (1.6%), as well as being significantly over-represented in the CON group (13.3%, p=0.035; OR=10.35; 95% CI:0.60–180.20) in comparison to the NON subgroup (0.0%), where the GG genotype was absent when only female participants were compared. In contrast, the AA genotype was under-represented in the CON group (38.9%, p=0.065) in comparison to the ACL group (54.2%) and significantly under-represented in the CON group (38.9%, p=0.013; OR=0.33; 95% CI:0.14-0.78) compared to the NON subgroup (65.6%) when only female participants were analysed (Figure 2.3D). In addition, the G allele of rs516115 was significantly over-represented in the CON group (37.2%) when compared to the ACL group (23.8%, p=0.014; OR=1.90; 95% CI:1.14–3.18) and the NON subgroup (17.2%, p=0.003; OR=2.86; 95% CI:1.40–5.85) when only female participants were analysed (Appendix D, Supplementary Table 2). There were no significant differences in the allele frequency distributions of rs13312816 between the three groups for the female participants. All the *DCN* variants were in HWE for all groups.

Only three of the possible four inferred haplotypes constructed for the two *DCN* variants (rs13312816 A>T - rs516115 A>G) had a frequency greater than 0%. There were no significant differences in the distribution of the *DCN* inferred haplotypes when only male participants were compared between the CON and ACL groups (Figure 2.3E). When the female participants were analysed, there was a trend for the T-A haplotype to be under-represented in the CON group (N=90, 62.8%) in comparison to the ACL group (N=61, 76.2%) (Figure 2.3F). *DCN* rs13312816 and rs516115 were found to be in compete LD (D'=1.0).

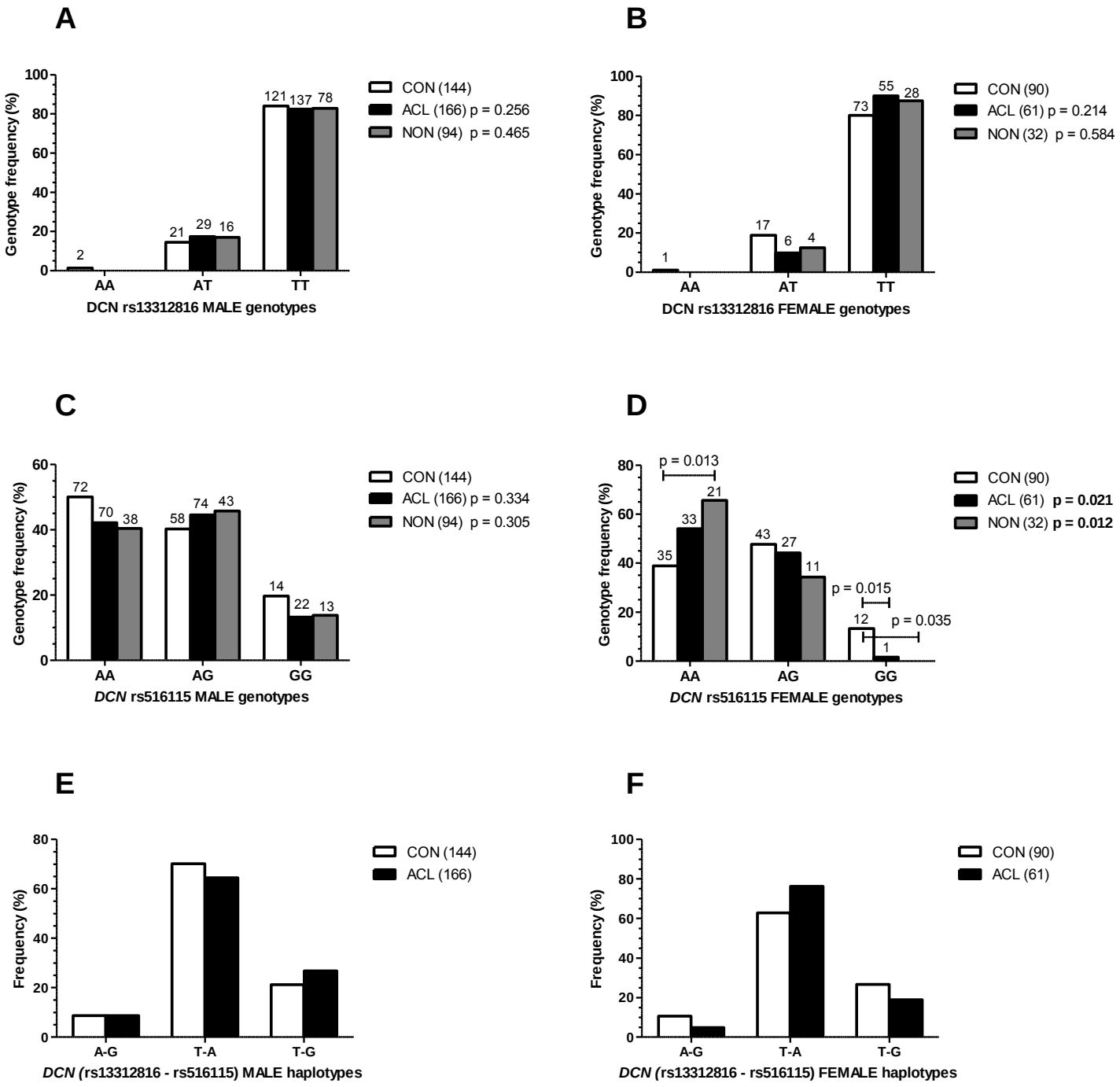


Figure 2.3: A-D: Genotype frequency distribution of *DCN* rs13312816 and rs516115 for the control (CON, white bars), anterior cruciate ligament rupture (ACL, black bars) and non-contact anterior cruciate ligament rupture (NON, grey bars) groups for male and female participants. **E-F:** The allele frequency distribution of the *DCN* (rs13312816 A>T - rs516115 A>G) inferred haplotypes amongst the control (CON, grey bars) and anterior cruciate ligament rupture (ACL, black bars) groups for male and female participants. The p values of the significantly different haplotypes are indicated. The total number of participants with available genotype data within each group is indicated in parenthesis on the graph.

2.3.5 *FMOD* and *LUM* genes

No significant differences in genotype frequencies were noted between the CON and ACL groups when the variants in the *FMOD* (rs7543148: p=0.458; rs10800912: p=0.616) (Figures 2.4A and 2.4B) and *LUM* (rs2268578: p=0.598) (Figure 2.4C) genes were analysed. Similarly, no significant differences in allele frequencies were noted between these groups (Appendix D, Supplementary Table 2). Likewise, no significant differences in the genotype and allele frequency distributions were noted when data was stratified by mechanism of injury or sex.

There were no significant differences in the distribution of the inferred haplotypes for *FMOD* (rs7543148 G>A – rs10800912 C>T) between CON and ACL groups (data not shown). Only six of the possible eight haplotypes constructed for the 56 Kb genetic interval overlapping the *LUM* and *DCN* genes (rs2268578 T>C - rs13312816 A>T - rs516115 A>G) had a frequency greater than 2%. The T-A-G inferred haplotype was significantly over-represented (p=0.038) in the CON group (N=234, 9.16%) in comparison to the ACL group (N=227, 7.26%) (Figure 2.4D). *LUM* rs2268578 and *DCN* rs13312816 SNPs were in high LD (D'=0.927), and *DCN* rs13312816 and rs516115 were in complete LD (D'=1.000).

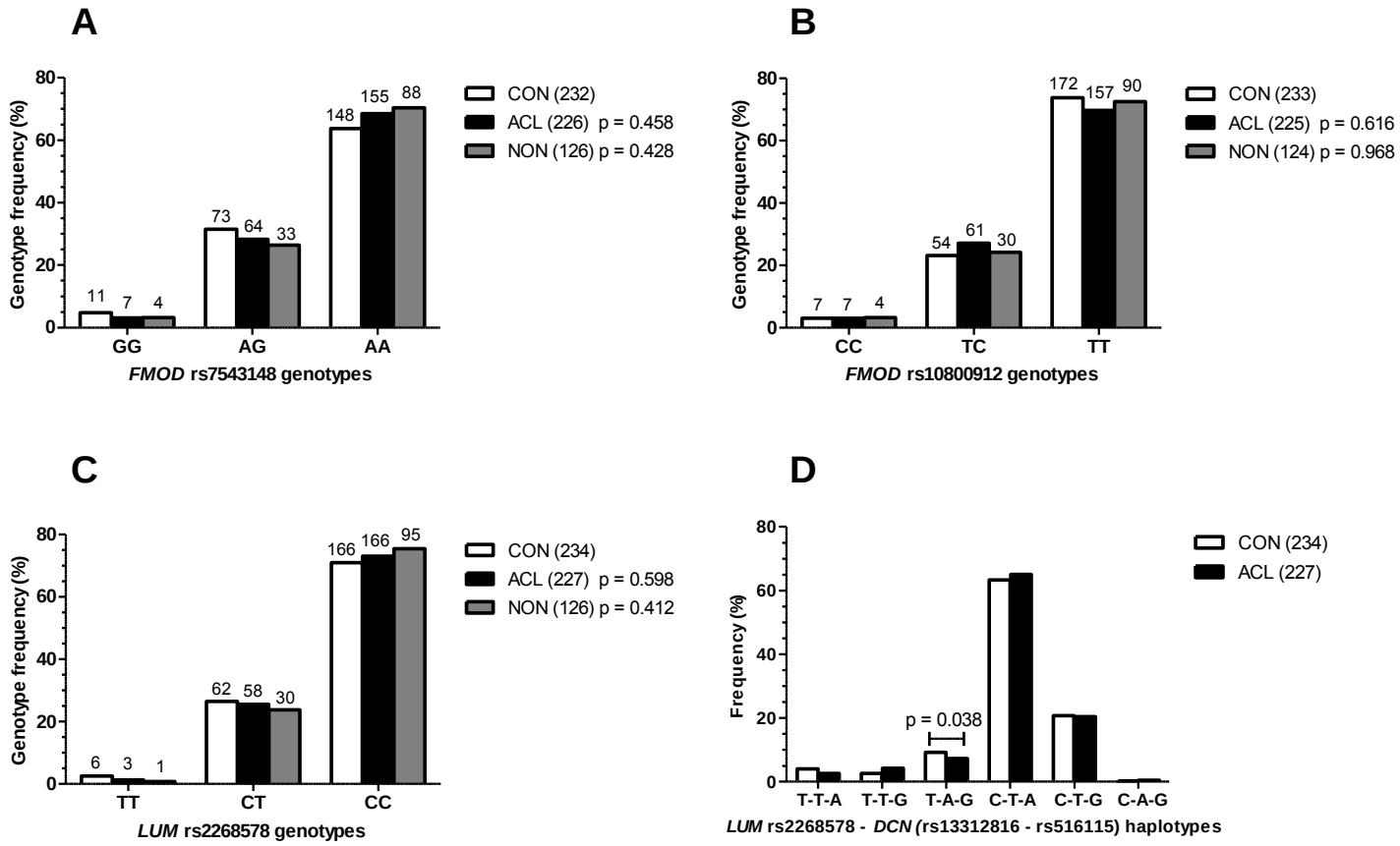


Figure 2.4: A-B: Genotype frequency distribution of *FMOD* rs7543148 and rs10800912 for the control (CON, white bars), anterior cruciate ligament rupture (ACL, black bars) and non-contact anterior cruciate ligament rupture (NON, grey bars) groups. **C:** Genotype frequency distribution of *LUM* rs2268578 for the control (CON, white bars), anterior cruciate ligament rupture (ACL, black bars) and non-contact anterior cruciate ligament rupture (NON, grey bars) groups. **D:** The frequency distribution of the inferred haplotypes constructed for the genetic interval spanning the *LUM* and *DCN* genes (rs2268578 T>C - rs13312816 A>T - rs516115 A>G) for the control (CON, grey bars) and anterior cruciate ligament rupture (ACL, black bars) groups. The p values of the significantly different distributions are indicated. The total number of participants with available genotype data within each group is indicated in parenthesis on the graph.

2.4 Discussion

Proteoglycans, such as aggrecan, have major structural roles in ligaments, and the SLRPs biglycan, decorin, fibromodulin and lumican are critical in regulating both ECM remodelling and collagen fibrillogenesis through their interactions with the collagen fibril, the major building block of ligaments and tendons.[195,252,258] In light of the essential role of proteoglycans in fibrillogenesis, and the previous associations of sequence variants within genes implicated in fibrillogenesis (*COL1A1*, *COL5A1* and *COL12A1*) with ACL injury risk,[35-38] this study aimed to investigate ten variants within five genes encoding proteoglycans (*ACAN* rs2351491, rs1042631, rs1516797, *BGN* rs1126499, rs1042103, *DCN* rs13312816, rs516115, *FMOD* rs7543148, rs10800912, and *LUM* rs2268578) for an association with the risk of ACL injuries. The main findings of this study include: (i) *ACAN* rs1516797 was independently associated with the risk of ACL injury in all participants; (ii) *DCN* rs516115 was independently associated with the risk of injury in female participants (sex-linked association); and (iii) inferred haplotype analyses further implicated regions overlapping four of the proteoglycan encoding genes (*ACAN*, *BGN*, and *LUM-DCN*) with ACL injury susceptibility. This study is the first report of genetic associations between the genes encoding proteoglycans and ACL injury susceptibility.

Aggrecan is a large, structural proteoglycan that associates with the collagen network, maintaining the structure of ligaments by attracting water (Figure 1.10). It is composed of a protein core (approximately 230kDa) comprising three globular domains.[329-331] The present study found that participants with the rs1516797 G allele had an increased risk of rupturing their ACL (p=0.024; OR=0.72; 95% CI:0.55-0.96). Although the genotype frequency distribution was not significantly different between the ACL and CON groups, the trend suggests that the G allele may be associated with increased ACL injury susceptibility, while the TT genotype appears to be protective. The biological function of this T>G substitution in intron 12 of *ACAN* is unknown. The significant genotype effect of *ACAN* rs1042631 on sex (p=0.033) (Appendix D, Supplementary Table 1) noted in this study may be explained by the observation that the TT genotype of rs1042631 is rare among the participants (2.39%); only 11 males (3.55%) and 6 females (4.00%) in the study group have the TT genotype. This is in agreement with the reported TT genotype frequency of 1.8% within the general Caucasian population according to the HapMap-CEU (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=1042631). The authors are not aware of any plausible biological explanation for this observed genotype effect association. The significance of these observed frequencies needs to be further explored in larger sample sizes.

Analysis of *ACAN* inferred haplotypes (rs2351491 C>T - rs1042631 T>C - rs1516797 T>G) further highlights the potential role of aggrecan in the pathobiology of ACL injury susceptibility. The T-C-G haplotype was associated with an increased risk of ACL ruptures (p=0.005) while the T-C-T haplotype was associated with a decreased risk of ACL ruptures (p=0.001). These inferred haplotypes, which overlap the G3 domain, are therefore suggesting that this genomic interval, including *ACAN* rs1516797 T>G, may be harbouring DNA sequence signatures which may possibly modify aggrecans interaction with the collagen fibril, and this DNA region should thus be further explored. Interestingly, disease-causing mutations in proximity to rs1516797 T>G have been associated with inherited forms of skeletal dysplasia, a connective tissue disorder.[159,160]

This study also provided preliminary evidence suggesting that SLRPs such as biglycan and decorin may play a role in the pathobiology of ACL injuries. Disease-causing mutations within the *DCN* gene have previously been associated with connective tissue disorders.[300,332] The SLRPs function predominantly in collagen fibrillogenesis, but have also been implicated in regulating cell growth and matrix remodelling.[163,189,195,197,199,252]

BGN was chosen as a candidate for investigation because it is on the X chromosome; sex is an intrinsic risk factor for ACL ruptures [107] and previous ACL injury genetic association studies have observed sex-genetic interactions.[37,38] Although neither of the two SNPs investigated within the *BGN* gene were independently associated with ACL injury risk, the C-G *BGN* inferred haplotype (rs1126499 C>T – rs1042103 G>A) was associated with a decreased risk of ACL injury in females (p=0.027). This haplotype analysis suggests that the region overlapping *BGN* may be modulating ACL injury susceptibility and more specifically, this genomic interval in the *BGN* gene should be further explored to identify the causal regulatory DNA sequence signatures.

Biglycan regulates collagen fibrillogenesis and structure by interacting with collagen fibrils (Figure 1.10).[195] In addition, biglycan also interacts with growth factors such as TGF-β, indicating this proteoglycans involvement in modulating growth factor availability to cells and its role in regulating matrix turnover.[189] It is interesting to note that previous genetic association studies have also observed a similar sex-specific selective advantage, as identified in this study, with variants localised to the *COL5A1* 3'-UTR which is also implicated in fibrillogenesis.[37]

The core protein of *DCN*, another SLRP investigated, binds to collagen to regulate collagen fibrillogenesis (Figure 1.10).[197] This study noted that *DCN* rs516115 A>G was implicated in ACL injury risk in female participants, with the GG genotype specifically associated with a 10.4 fold decreased risk of injury (p=0.035; OR=10.35; 95% CI:0.60–180.20) and the AA genotype associated with an increased risk of injury (p=0.013; OR=0.33; 95% CI:0.14-0.78) when the CON and NON groups were compared. The associations were further illustrated when the A and G alleles were significantly over-represented in the ACL and CON groups respectively (p=0.014), and mirrored when data was stratified by mechanism of injury (p=0.003). The functional significance of this variant is unknown but one can hypothesise that variations in the LRR protein core may modify the interaction of decorin with TGF-β or the collagen fibril,[252,260] thereby possibly modifying fibrillogenesis and affecting the mechanical properties of the ligament.

Although biglycan is on the X chromosome, it is suggested to behave like a pseudo-autosomal gene that escapes X-inactivation and has an active copy on the Y chromosome. There is however also conflicting evidence illustrating that biglycan does undergo X-inactivation and does not have a Y chromosome homologue.[248] The pseudo-autosomal expression of biglycan could be attributed to a gene or several genes adjacent to *BGN* that escape X-inactivation and control the transcriptional activity of biglycan.[248] The difference, or lack thereof, in dosage and expression of *BGN* between males and females has not been fully explored, and may play a role in the altered risk of ACL injury in females. Ovarian hormone levels, and particularly oestrogen, modulate the synthesis and degradation of SLRPs such as biglycan and decorin.[333,334] Receptor sites for oestrogen and progesterone have been found in the ACL,[52,119] and sex hormones are suggested to affect ACL structure and composition.[119] Therefore it is reasonable to propose that the regulation of SLRPs by oestrogen [335-337] may account for the sex-specific association of these proteoglycan genes with the risk of ACL injury in females.

Haplotype analysis, the investigation of a set of variants inherited together, is often more informative in detecting an association compared to analysing individual variants alone.[225] No independent associations were noted for *LUM* rs2268578, however analysis of the inferred haplotype encompassing the *LUM* and *DCN* (rs2268578 T>C – *DCN* rs13312816 A>T - rs516115 A>G) genes implicated the T-A-G allele combination with reduced ACL injury risk (p=0.038). This region, from rs2268578 to rs516115, spans approximately 56 Kb (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs) encompassing the *LUM* and *DCN* genes. One can speculate that this genomic interval overlapping *LUM* and *DCN* influences ACL injury

susceptibility by effecting fibrillogenesis. Therefore it is critical that this genomic region is further interrogated to identify functional DNA sequences. Both rs2268578 and rs13312816 (D'=0.927) and rs13312816 and rs516115 (D'=1.0) are in LD.

Proteoglycans such as aggrecan, biglycan, decorin, fibromodulin and lumican play an important role in fibrillogenesis; possibly through their myriad of interactions with various proteins, including the collagens (Figure 1.10), and cell-signalling molecules within the ECM. Altering the properties of the collagen fibril will most likely alter both the biomechanical and functional properties of the ligament and one can therefore hypothesise that this modulation will impact on ACL injury risk.[140] It was thus not surprising that our novel results are implicating sequence variants within four proteoglycan genes with ACL injury susceptibility. Although there is no immediate clinical translation from this study, the results are suggesting that inter-individual variations in the collagen network and fibril assembly might be an important molecular mechanism contributing to the aetiology of ACL ruptures, similar to Achilles tendinopathy.[338] To improve our understanding of ACL pathogenesis and susceptibility, it is imperative that we start elucidating the net effect of the intricate interactions of proteoglycans with their receptors in regulating the structural properties of the ECM, including the collagen fibril. The elucidation of these interactions is vital for the development of possible therapeutic interventions targeting proteoglycans specifically.[242]

The cases and controls investigated in this study were not matched for the confounding variables weight and height, and this is therefore a limitation of the study. Heavier individuals are more likely to sustain an ACL rupture as adiposity is a contributing risk factor to the development of musculoskeletal soft tissue injuries.[21,33] Although care was taken to recruit controls that participated in the same sports as those who had sustained an ACL rupture, a limitation is that the cases and controls were not matched exactly for exposure to contact and non-contact jumping sports, which are high risk for ACL injuries.

The pathway-based approach followed in this study provides evidence that highlights the potentially important biological role that proteoglycans *ACAN*, *BGN*, *DCN* and/or *LUM* may have in modulating ACL injury susceptibility. These findings should however be repeated in independent populations to confirm the associations described. These results suggest a need to further interrogate the genomic intervals encompassing the proteoglycan genes with regards to ACL injury risk, as well as to identify the reasons for the multiple sex-specific associations observed.

Chapter 3: The interaction of genes encoding proteins involved in fibrillogenesis, in the modulation of anterior cruciate ligament injury risk

3.1 Introduction

The novel study presented in Chapter 2 highlights the potential role of proteoglycan genes in the modulation of ACL injury risk. *ACAN* rs1516797 was independently associated with ACL injury risk and *DCN* rs516115 was associated with the risk of ACL ruptures in females only. Furthermore, inferred haplotype analyses suggested that genomic regions overlapping *ACAN*, *BGN*, *DCN* and *LUM*, may be biologically important in understanding the mechanisms underlying ACL injury susceptibility. The *BGN* and *DCN* genes were implicated in the risk of ACL injury in female participants only.

Biglycan and decorin have been implicated in the process of collagen fibrillogenesis through their interactions with collagens (Figure 1.10).[195,197] Decorin is known to directly interact with collagen V, having either a direct effect on the collagen fibril or affecting the interaction of collagen V with other molecules.[256] As described in Chapter 1, (Section 1.3.2.2) rs12722 C>T, the *Bst*UI RFLP within the *COL5A1* gene, was previously associated with ACL injury risk in a female Caucasian population.[37] The CC genotype of *COL5A1* rs12722 was found to be significantly over-represented in the control group, while the T allele was significantly over-represented in the group with ACL ruptures, when only females were investigated in the Caucasian cohort.[37] This Caucasian cohort is the same cohort that was analysed for associations between the proteoglycan genes and ACL injury risk in Chapter 2 of this dissertation, with additions.

The aim of this study was therefore to investigate if there are any gene-gene interactions between genes involved in regulating fibrillogenesis, and the modulation of ACL injury susceptibility. The objectives were to investigate interactions between the proteoglycan genes implicated in the risk of ACL injury in Chapter 2, and the *COL5A1* gene previously associated with ACL injury risk.[37] Inferred allele constructs were used to investigate the collective interactions between the *BGN*, *DCN* and *COL5A1* risk-associated SNPs, for their possible role in modulating sex-specific ACL injury susceptibility. Taking into account the previous findings presented in Chapter 2 and the data presented by Posthumus *et al.* (2009),[37] the hypothesis was that the C allele of rs1126499 C>T and the G allele of rs1042103 T>C within *BGN,* as well as the C allele of *COL5A1* rs12722, would be associated with a decreased risk of ACL injury. With regards to *DCN,* the hypothesis was that the A allele of rs516115 A>G, together with the T allele of *COL5A1* rs12722, would be associated with an increased risk of injury, and the alternate alleles (G and C respectively) would be implicated in a decreased risk of injury.

3.2 Materials and Methods

A total of 461 participants in the previously defined ACL (227 individuals with surgically diagnosed ACL ruptures), and CON (234 healthy active individuals without any history of ACL injury) groups, were genotyped as part of a previous study (Chapter 2) for SNPs within the *BGN* (rs1126499 C>T and rs1042103 G>A) and *DCN* (rs516115 A>G) genes. These participants were previously genotyped for the rs12722 C>T SNP within the *COL5A1* gene by Posthumus *et al.* (2009).[37] Refer to Chapter 2, Section 2.3.1 for details on participant characteristics.

In order to determine whether these genes interact in the modulation of ACL injury risk, inferred allele combinations were constructed using the Chaplin case-control haplotype inference software program Version 1.2.2 (http://www.genetics.emory.edu/labs/epstein/software/chaplin/index.html). The allele constructs included (i) *BGN* (rs1126499 C>T – rs1042103 G>A) – *COL5A1* rs12722 C>T, (ii) *DCN* rs516115 A>G – *COL5A1* rs12722 C>T (*DCN* rs13312816 and rs516115 are in complete LD and therefore only one SNP was included), (iii) *BGN* (rs1126499 C>T – rs1042103 G>A) – *DCN* rs516115 A>G and (vi) a combination of all four SNPs within the three genes: *BGN* (rs1126499 C>T – rs1042103 G>A) - *DCN* (rs516115 A>G) - *COL5A1* (rs12722 C>T). Allele frequencies of SNPs within each allele construct were compared between the ACL group and CON group using the Chaplin software Chi-squared tests. The NON subgroup defined in Chapter 2 was not analysed for inferred allele combinations in this chapter because the sample size became too small to identify valid statistical associations. Due to the sex-specific associations identified in Chapter 2, and by Posthumus *et al.* (2009),[37] all analyses were performed on male and female participants separately.

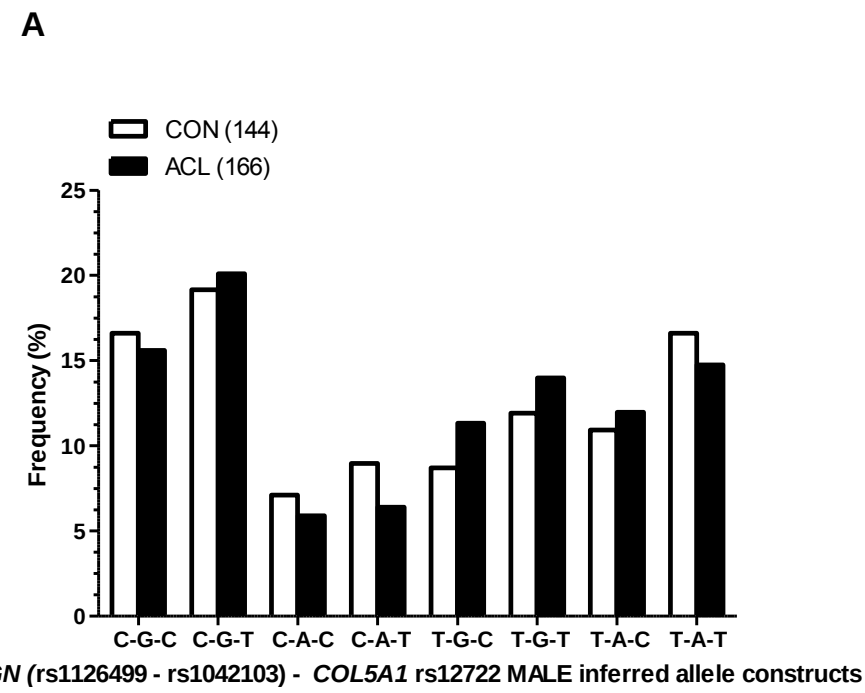
3.3 Results

3.3.1 *BGN – COL5A1*

All of the allele combinations constructed for *BGN* (rs1126499 C>T – rs1042103 G>A) and *COL5A1* rs12722 C>T had a frequency greater than 2%. There were no significant frequency differences between the ACL and CON groups for the inferred allele constructs for *BGN* (rs1126499 C>T – rs1042103 G>A) and *COL5A1* rs12722 C>T when only male participants were analysed (Figure 3.1A). However, when only the females participants were analysed, the C-G-C allele construct was significantly over-represented (p=0.001) in the CON group (N=90, 23.0%) in comparison to the ACL group (N=61, 9.7%), and the T-A-T construct was significantly under-represented (p=0.001) in the CON group (N=90, 11.3%) in comparison to the ACL group (N=61, 28.4%) (Figure 3.1B).

3.3.2 *DCN – COL5A1*

All of the allele combinations constructed for *DCN* rs516115 A>G and *COL5A1* rs12722 C>T had a frequency greater than 5%. The inferred allele constructs for *DCN* rs516115 A>G and *COL5A1* rs12722 C>T for male and female participants are displayed in Figures 3.2A and 3.2B respectively. The A-T allele construct was significantly over-represented (p=0.027) in the male CON group (N=144, 41.6%) in comparison to the ACL group (N=166, 36.9%). For the female participants, the A-T construct was significantly under-represented (p<0.001) in the CON group (N=90, 30.9%) in comparison to the ACL group (N=61, 51.1%), and G-C was significantly over-represented (p<0.001) in the CON group (N=90, 19.3%) in comparison to the ACL group (N=61, 6.2%).



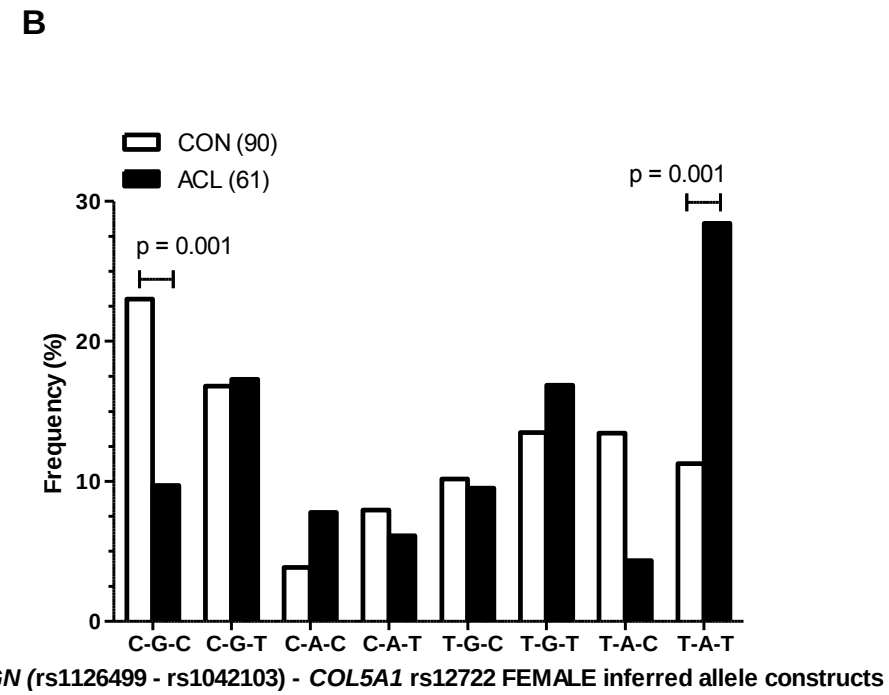


Figure 3.1: Frequency distribution of the inferred allele constructs for *BGN* (rs1126499 C>T – rs1042103 G>A) and *COL5A1* rs12722 C>T amongst the control (CON, white bars) and anterior cruciate ligament rupture (ACL, black bars) groups for male (**A**) and female (**B**) participants. The p values of significantly different distributions are noted. The total number of participants with available genotype data within each group is indicated in parenthesis on the graph.

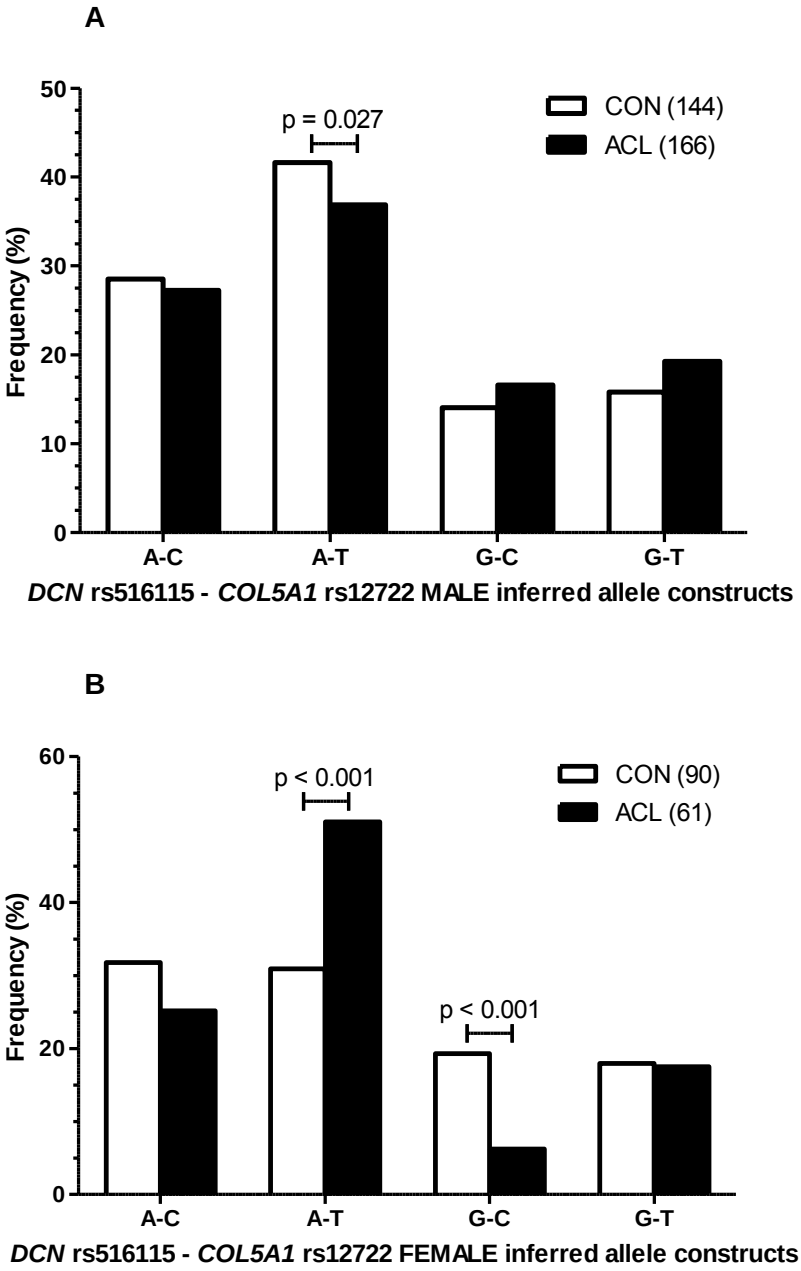


Figure 3.2: Frequency distribution of the inferred allele constructs for *DCN* rs516115 A>G and *COL5A1* rs12722 C>T amongst the control (CON, white bars) and anterior cruciate ligament rupture (ACL, black bars) groups for male (**A**) and female (**B**) participants. The p values of significantly different distributions are noted. The total number of participants with available genotype data within each group is indicated in parenthesis on the graph.

3.3.3 *BGN – DCN*

All of the allele combinations constructed for *BGN* (rs1126499 C>T – rs1042103 G>A) and *DCN* rs516115 A>G had a frequency greater than 2%. There were no significant differences between the ACL and CON groups for the inferred allele constructs for *BGN* (rs1126499 C>T – rs1042103 G>A) and *DCN* rs516115 A>G when only male participants were analysed (Figure 3.3A). However, when the females were analysed, the C-G-G allele construct was significantly over-represented (p=0.011) in the CON group (N=87, 13.9%) in comparison to the ACL group (N=57, 7.2%). The T-G-G allele construct was also significantly over-represented (p=0.026) in the CON group (N=87, 13.9%) in comparison to the ACL group (N=57, 5.7%). In contrast, the T-A-A construct was significantly under-represented (p=0.021) in the CON group (N=87, 20.2%) in comparison to the ACL group (N=57, 24.8%) (Figure 3.3B).

3.3.4 *BGN – DCN – COL5A1*

In order to investigate the potential gene-gene interactions between these three genes (*BGN*, *DCN* and *COL5A1*), inferred allele constructs were generated from the genotype data of SNPs within the *BGN* (rs1126499 C>T – rs1042103 G>A), *DCN* (rs516115 A>G) and *COL5A1* (rs12722 C>T) genes. No significant differences in the inferred allele constructs were noted between the CON and ACL groups when only the male participants were compared (Figure 3.3A).

For the females participants however, the C-G-G-C inferred allele construct was significantly over-represented (p=0.001) in the CON group (N=90, 10.7%) in comparison to the ACL group (N=61, 4.2%). C-G-A-C was also significantly over-represented (p=0.044) in the CON group (N=90, 13.3%) in comparison to the ACL group (N=61, 3.7%). T-G-A-T on the other hand was significantly under-represented (p=0.038) in the CON group (N=90, 7.6%) in comparison to the ACL group (N=61, 10.1%), along with T-A-A-T being significantly under-represented (p<0.001) in the CON group (N=90, 7.4%) in comparison to the ACL group (N=61, 22.1%) (Figure 3.3B). Inferred allele combinations which had frequencies less than 5% in both the ACL and CON groups were not displayed in the graphs for Figure 3.3, thus only ten of the possible sixteen allele combinations are presented.

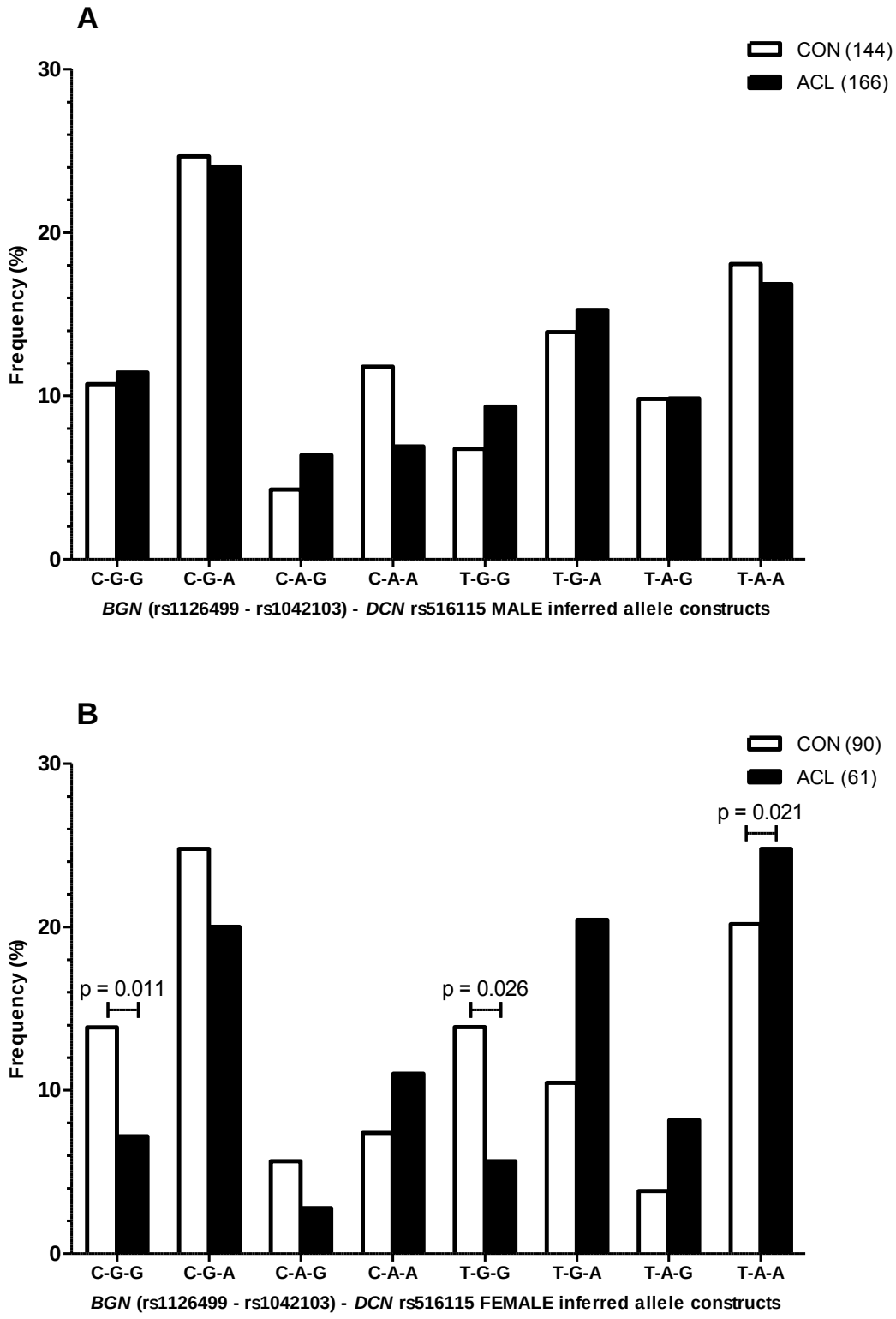


Figure 3.3: Frequency distribution of the inferred allele constructs for *BGN* (rs1126499 C>T – rs1042103 G>A) and *DCN* rs516115 A>G amongst the control (CON, white bars) and anterior cruciate ligament rupture (ACL, black bars) groups for male (**A**) and female (**B**) participants. The p values of significantly different distributions are noted. The total number of participants with available genotype data within each group is indicated in parenthesis on the graph.

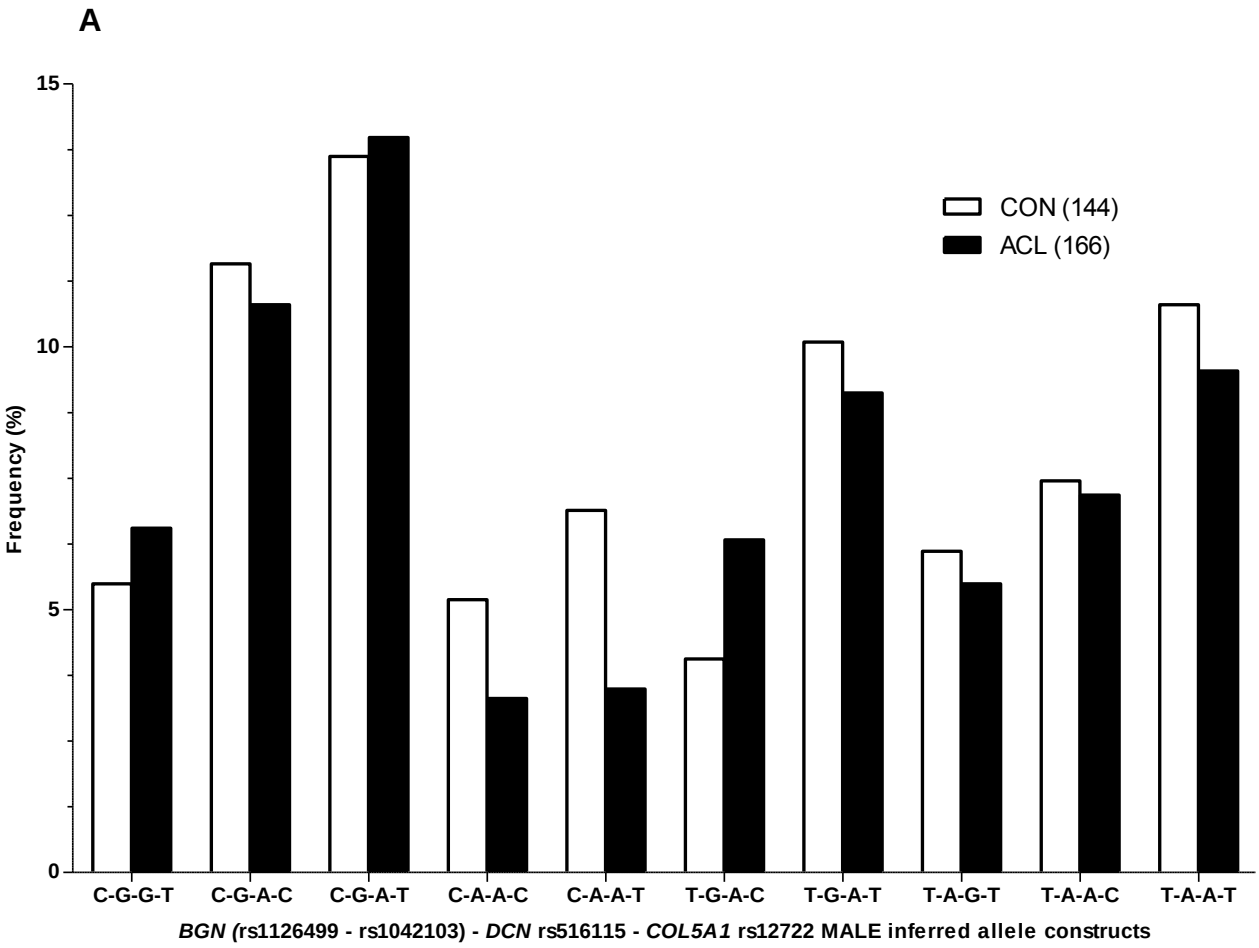


Figure 3.4A: The frequency distribution of the inferred alleles constructs for *BGN*, *DCN* and *COL5A1* (rs1126499 C>T – rs1042103 G>A – rs516115 A>G – rs12722 C>T) amongst the control (CON, grey bars) and anterior cruciate ligament rupture (ACL, black bars) groups for male participants. The p values of the significantly different distributions are indicated. The total number of participants with available genotype data within the CON and ACL groups are indicated in parenthesis on the graph.

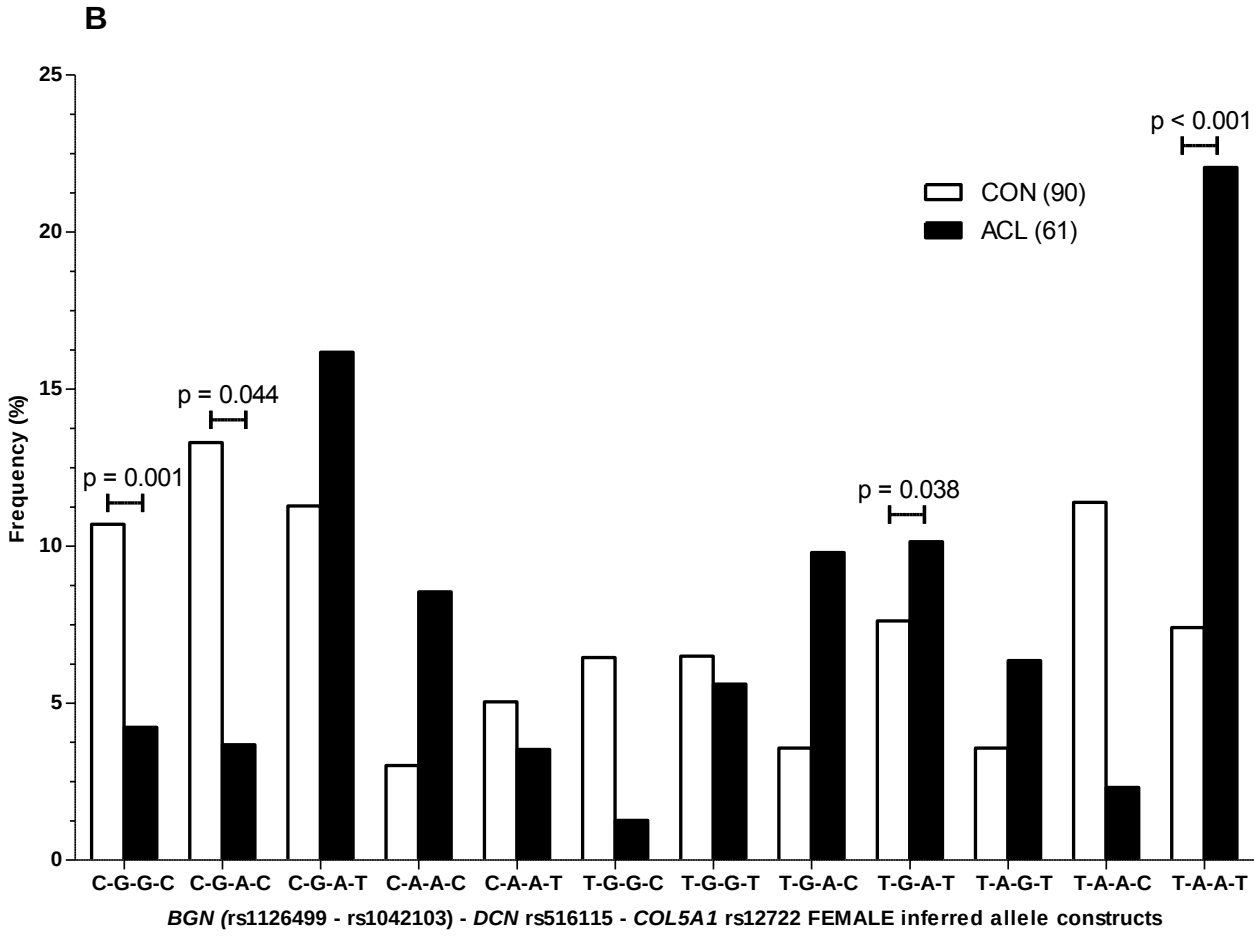


Figure 3.4B: The frequency distribution of the inferred allele constructs for *BGN*, *DCN* and *COL5A1* (rs1126499 C>T – rs1042103 G>A – rs516115 A>G – rs12722 C>T) amongst the control (CON, grey bars) and anterior cruciate ligament rupture (ACL, black bars) groups for female participants. The p values of the significantly different distributions are indicated. The total number of participants with available genotype data within the CON and ACL groups are indicated in parenthesis on the graph.

3.4 Discussion

The *LUM-DCN* haplotype presented in Chapter 2 suggested that gene-gene interactions between proteoglycan encoding genes, may be taking place in the modulation of ACL injury risk. The investigation of inferred haplotypes or inferred allele combinations is potentially more powerful in effectively capturing a risk profile than looking at one polymorphism independently.[339] Independent associations with ACL injury risk were noted for sequence variants within (i) proteoglycan genes encoding proteins that regulate fibrillogenesis (Chapter 2), and (ii) collagen encoding genes (*COL5A1*)[37] also implicated in fibrillogenesis. Therefore it is reasonable to predict that a model comprising both the proteoglycan and collagen associated genes involved in fibrillogenesis, may have a greater impact on modulating ACL injury susceptibility.

Biglycan regulates collagen fibrillogenesis and structure by interacting with collagen fibrils.[195] The LRRs of the core protein of decorin facilitate its binding to the surface of fibrillar collagens where it can retard the rate and degree of collagen fibrillogenesis.[252-255] Type V collagen intercalates with type I collagen and is believed to regulate the collagen fibril diameter during fibrillogenesis.[138] The binding of decorin to collagen V can have either a direct effect on the collagen itself or may mediate the interaction of collagen V with other molecules.[256]

Fibril-collagen interactions are highly important with regards to the ligaments response to loading and matrix remodelling (Figure 1.10). Recently it has been hypothesised that sequence variations in the collagen encoding genes may alter the structure and assembly of the collagen fibril and therefore affect the mechanical properties of the fibril and its response to loading.[140,338] The *BGN* and *COL5A1* (rs1126499 C>T – rs1042103 G>A - rs12722 C>T) inferred allele construct revealed two alternate allele combinations associated with either protection from or risk of injury in female participants. The C-G-C allele combination was associated with protection from ACL injury ($p=0.001$), while T-A-T was associated with risk of ACL injury ($p=0.001$). The inferred allele constructs mimic the observations from Chapter 2, where the C-G *BGN* (rs1126499 – rs1042103) haplotype was associated with a decreased risk of ACL injury. This also aligns with the hypothesis and previous findings that the C allele of *COL5A1* rs12722 is associated with protection from ACL injury and the T allele is associated with risk of injury.[37]

The A-T combination of the *DCN* and *COL5A1* (rs516115 A>G - rs12722 C>T) inferred allele construct was associated with risk of injury ($p<0.001$) in females, and the alternate G-C allele construct was

associated with protection from injury (p<0.001) in females. Once again, this finding is in line with the results of the current and previous studies,[37] highlighting that the A and G alleles of *DCN* rs516115 are associated with risk of and protection from injury respectively (Chapter 2).

When analysing the *BGN* (rs1126499 C>T – rs1042103 G>A) and *DCN* rs516115 A>G inferred allele construct, two alternate allele combinations were associated with protection from (C-G-G, p=0.011) and risk of (T-A-A, p=0.021) ACL injury in female participants only. These combinations are reflective of the results observed in Chapter 2 and agree with the hypotheses implicating the respective alleles in protection from and risk of ACL injury. A third allele construct (T-G-G) was also associated with decreased risk of ACL injury (p=0.026), indicating that the *BGN* rs1126499 SNP may not have as strong of a role in the modulation of ACL injury risk as the other two SNPs investigated in this gene-gene interaction.

When the SNPs from *BGN*, *DCN* and *COL5A1* were combined in an inferred allele construct [*BGN* (rs1126499 C>T – rs1042103 G>A), *DCN* rs516115 A>G and *COL5A1* rs12722 C>T], four strong associations in line with the above mentioned results were identified. This again echoes the potential biological consequence the interactions between the proteoglycan and collagen genes may have in collectively modulating the risk of sustaining an ACL rupture. These results emphasise that ACL injuries are not only multifactorial, but also that multiple genes may be interacting and contributing to the aetiology of this injury.

Similar to the results observed in Chapter 2, female-specific associations were identified in this study. The reason for this remains unknown although suggestions have been made. One of these is based on the difference, or lack thereof, in the expression of *BGN* between males and females, because this gene is on the X chromosome.[248] Sex hormones are also speculated to affect the structure and composition of the ACL,[119] as well as modulate the synthesis and degradation of SLRPs such as biglycan and decorin.[333,334] The underlying mechanisms for the observed sex-specific associations need to be further explored.

One limitation of the analyses presented in this Chapter was that when participants were stratified by sex, and by their specific inferred allele combinations, the sample size was invariably diminished. It would be therefore be valuable if the sample size of this study is increased. More importantly, it would be beneficial to repeat the study in an independent population.

In conclusion, this study revealed evidence suggesting that gene-gene interactions are taking place between the proteoglycan genes and the *COL5A1* gene in the modulation of ACL injury risk. It is likely that multiple genetic variants within genes encoding proteins involved in the regulation of fibrillogenesis, ECM structure and remodelling, and multiple gene-gene interactions may affect ones susceptibility to sustaining an ACL injury. The interaction of genes within a pathway, such as fibrillogenesis (Figure 1.10) and ECM remodelling, is more informative than the independent effect of a single gene alone. These results add to the current risk model for ACL injuries. Identifying genes implicated in complex multifactorial phenotypes however only partially explains the biological pathway that leads to the disease. Only once gene expression, protein function and protein interactions are understood, will the physiological mechanisms become visible.[223]

Chapter 4: Final remarks

4.1 Summary

ACL ruptures, as highlighted in Chapter 1, are considered one of the most severe and detrimental injuries sustained in sports.[30] High risk sports for ACL ruptures, as identified in this study and others,[340,341] include rugby and soccer for males, and hockey and netball for females. This is due to the fact these are fast moving sports that require rapid stops and changes in direction during play. However, ACL injuries are not only contributed to by the sports and activities an individual takes part in, it is considered a multifactorial injury caused by multiple risk factors,[21] (Figure 4.1) which can be divided into five interacting groups: extrinsic risk factors, and the four intrinsic risk factors, anatomy, hormones, neuromuscular activity and genetics.[4,33] Although this dissertation focused on the investigation of one of these risks, it is important to understand that the five groups comprise a network of factors that act together in the modulation of ACL injury susceptibility.

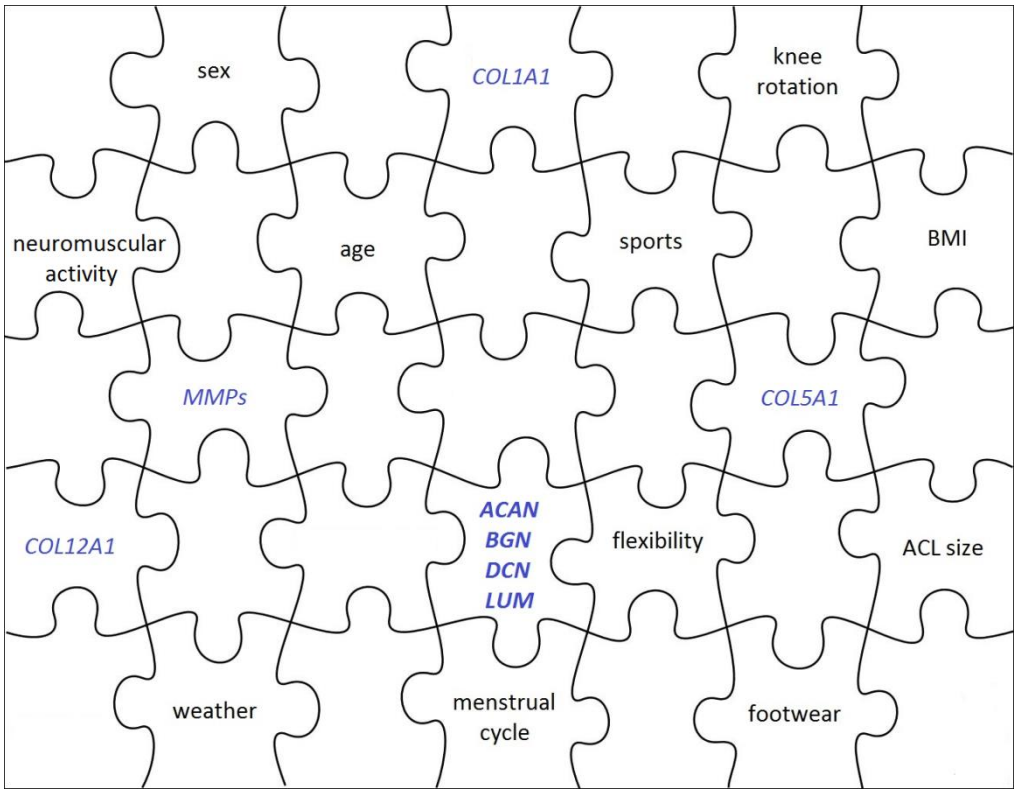


Figure 4.1: The puzzle pieces of risk factors that contribute to ACL injury susceptibility. Genetic risk factors are noted in blue, with the proteoglycan genes (identified in this dissertation) noted in bold.

The genetic contribution to ACL injuries, as well as other multifactorial soft tissue injuries, has become well-established.[5,21,85,342] Sequence variants within several genes have been implicated in the modulation of ACL injury risk, highlighting that this is a polygenic condition which is likely contributed to by multiple genes in various pathways. Although proteoglycans are important structural components of ligaments, the association of variants within genes encoding proteoglycans have not been investigated to date. This dissertation therefore investigated a group of genes, encoding proteoglycans, for an association with ACL injury risk.

The aims of this dissertation were:

- to identify novel genetic loci within proteoglycan encoding genes, which may predispose an individual to sustaining an ACL rupture,
- to use inferred haplotype analyses to identify genomic regions within each of the proteoglycan genes that may be contributing to ACL injury susceptibility and thus require further investigation,
- to determine whether gene-gene interactions between the proteoglycan genes and the previously associated *COL5A1* gene may collectively contribute to the modulation of ACL injury risk, by exploiting inferred allele constructs.

Proteoglycans form a very important component of the ground substance that comprises the ECM of ligaments.[40] They have been implicated in maintaining both the structural and functional integrity of tissues by facilitating matrix-matrix interactions, cell-matrix organisation and cell-matrix signalling interactions.[158] The genes specifically investigated in this dissertation (*ACAN*, *BGN*, *DCN*, *FMOD* and *LUM*), encode proteins that interact with the collagen network and regulate fibrillogenesis (Figure 4.2) and are thus vitally important in ECM homeostasis in the normal state and ECM remodelling following mechanical loading.[162] The identification of genes involved in the aetiology of ACL injuries will elucidate the significant biological pathways that may be contributing to the mechanisms that underlie this injury. Linking genetic variants to injury susceptibility is a tool used to understand gene functions [60] and injury development. Only once the mechanisms of injury are understood, can we improve prevention strategies to decrease the incidence of injury, as well as to refine treatment and rehabilitation techniques post-injury.

4.1.1 Novel findings

This dissertation identified the following novel associations:

- *ACAN* rs1516797 T>G was independently associated with the risk of ACL injury in all participants. More specifically, the G allele was associated with a 1.4 fold increased risk of ACL rupture (p=0.024).

- *DCN* rs516115 A>G was independently associated with the risk of injury in female participants only (sex-specific association). The GG genotype was associated with a 10.4 fold decreased risk of injury (p=0.015) and the AA genotype was associated with a 3 fold increased risk of injury (p=0.013) when the CON and NON groups were compared. Furthermore, the A and G alleles were significantly over-represented in the ACL and CON groups respectively (p=0.014), and mirrored when data was stratified by non-contact mechanism of injury (p=0.003).

- Inferred haplotype analyses implicated regions overlapping four of the five investigated proteoglycan encoding genes with ACL injury susceptibility.
    - Two *ACAN* (rs2351491 C>T - rs1042631 T>C - rs1516797 T>G) haplotypes, on chromosome 15q26, were associated with risk of injury: the T-C-G haplotype was associated with an increased risk of ACL ruptures (p=0.005) while the T-C-T haplotype was associated with a decreased risk of ACL ruptures (p=0.001).
    - The C-G *BGN* (rs1126499 C>T – rs1042103 G>A) haplotype, on chromosome Xq28, was associated with a decreased risk of ACL injury in female participants only (p=0.027).
    - The haplotype overlapping a 56 Kb region on chromosome 12q21, encompassing the *LUM* and *DCN* genes [*LUM* rs2268578 T>C – *DCN* (rs13312816 A>T - rs516115 A>G)] was also implicated in ACL injury risk; specifically the T-A-G allele combination was associated with a decreased risk of ACL injury (p=0.038).

- Allele constructs revealed that the proteoglycan genes (i) collectively interact with each other, and (ii) their interactions with *COL5A1* potentially contribute to the modulation of ACL injury susceptibility.

- The *BGN-DCN* (rs1126499 C>T - rs1042103 G>A - rs516115 A>G) allele construct revealed two alternate combinations that were associated with protection from (C-G-G, p=0.011) and risk of (T-A-A, p=0.021) ACL injury in female participants only.
- Two alternate combinations of the *BGN* - *COL5A1* (rs1126499 C>T – rs1042103 G>A - rs12722 C>T) inferred allele construct, were associated with protection from (C-G-C, p=0.001) and risk of (T-A-T, p=0.001) injury in female participants specifically.
- The G-C combination of the *DCN* - *COL5A1* (rs516115 A>G - rs12722 C>T) allele construct was associated with protection from injury (p<0.001), while the alternate A-T haplotype was associated with risk of injury (p<0.001), both in female participants.
- Four female-specific associations were identified when *BGN-DCN-COL5A1* (rs1126499 C>T - rs1042103 G>A - rs516115 A>G - rs12722 C>T) allele constructs were investigated. The C-G-G-C and C-G-A-C constructs were associated with a decreased risk (p=0.001 and p=0.044 respectively) of ACL injury in females. While T-G-A-T and T-A-A-T were associated with an increased risk of ACL injury (p=0.038 and p<0.001 respectively) in female participants only.

In summary, two of the ten investigated polymorphisms were independently associated with the risk of ACL injury (*ACAN* rs1516797 and *DCN* rs516115). Haplotype analyses further implicated regions overlapping *ACAN*, *BGN*, and *LUM-DCN* in ACL injury susceptibility. Inferred allele constructs revealed that gene-gene interactions between the proteoglycan genes and *COL5A1* may collectively modulate ACL injury risk. It is important to note that the results of the associated haplotypes and allele constructs, were in line with the independently associated polymorphisms and the specific alleles were consistently associated with either an increased or decreased risk of ACL injury. This dissertation highlights the polygenic nature of ACL injuries and the probability that multiple genetic variants within genes involved in the regulation of fibrillogenesis, ECM structure and remodelling, and the complex array of gene-gene interactions in response to environmental triggers may together contribute to the biological mechanisms underpinning ACL injury risk.

4.1.2 The biologically significant pathways highlighted in this dissertation

Genetic association studies are important tools used to identify genes and biological pathways which may contribute to an individual having an inherited predisposition to a multifactorial condition, such as an ACL injury. Although one genotype or allele may impact an individual's susceptibility to injury, as was noted with *ACAN* rs1516797 and *DCN* rs516115, the combination of several SNPs or genes together is likely to have a stronger effect.[343] This dissertation investigated the interaction of genes associated with the collagen network, and involved in fibrillogenesis and ECM remodelling; the exploration of a pathway is more informative of the mechanisms that may be contributing to injury susceptibility than merely examining the independent effect of a single gene.[339] The association of sequence variants in the proteoglycan encoding genes can be used as part of a polygenic profile, including previously associated SNPs, to identify which individuals have an inherited predisposition to ACL injuries. However, it must once again be highlighted that genetics is one piece of the multiple risk factors that contribute to the puzzle that is injury (Figure 4.1), and should be used together with the other factors in a risk model to determine susceptibility.

The observed net effect of multiple genes and multiple interactions within the collagen network (Figure 4.2), as occurs in the regulation of fibrillogenesis, highlights the importance of this process in ligament homeostasis and thus provides an additional level of evidence implicating this pathway as part of the mechanisms that contribute to injury susceptibility. In the normal, uninjured state, ligaments are subjected to small amounts of mechanical loading due to everyday activities.[4] The ECM of ligaments is thus continuously undergoing degradation and remodelling in order to maintain homeostasis, and fibrillogenesis is an important part of the remodelling process. During this process, particular proteins involved, such as the SLRPs and collagens, will have altered expression. It is thus reasonable to deduce that if protein expression or function is modified due to variants within the respective genes, homeostasis and more importantly fibrillogenesis could be compromised. An individual may consequently be at an increased risk of sustaining an ACL injury because the biomechanical properties of the ligament will be altered and perhaps even compromised. The exact biological implications of each of the proteoglycan genes on the mechanism of injury will however remain unknown until the functional role of the polymorphisms and genetic interactions within the cells of the ECM are determined.

A recent study investigated the binding of decorin to collagen. The highly conserved concave region of SLRPs is believed to interact with the collagen triple helices. If mutation sites are engineered at the collagen-SLRP interface, it reduces the ability of the SLRP to bind collagen and decorin is consequently unable to inhibit collagen fibrillogenesis.[344] Different mutations result in varying degrees of binding ability, and it is thus plausible that various genetic polymorphisms, as investigated in the studies presented, may affect the binding of SLRPs to collagen and therefore alter the process of fibrillogenesis and compromise ligament homeostasis and the ability to effectively resist a load. It is therefore reasonable to propose that these genes may harbour potential biologically significant polymorphisms which if inherited may modulate an individual's predisposition to ACL injury. The results of this dissertation highlight that the collagen network and fibrillogenesis pathway, and possibly the rate and timing of fibrillogenesis as the *COL5A1* gene is the rate limiting gene of fibrillogenesis,[345] needs to be further explored in order to properly characterise the contribution of this pathway to the mechanism of ACL injury and subsequent injury susceptibility.

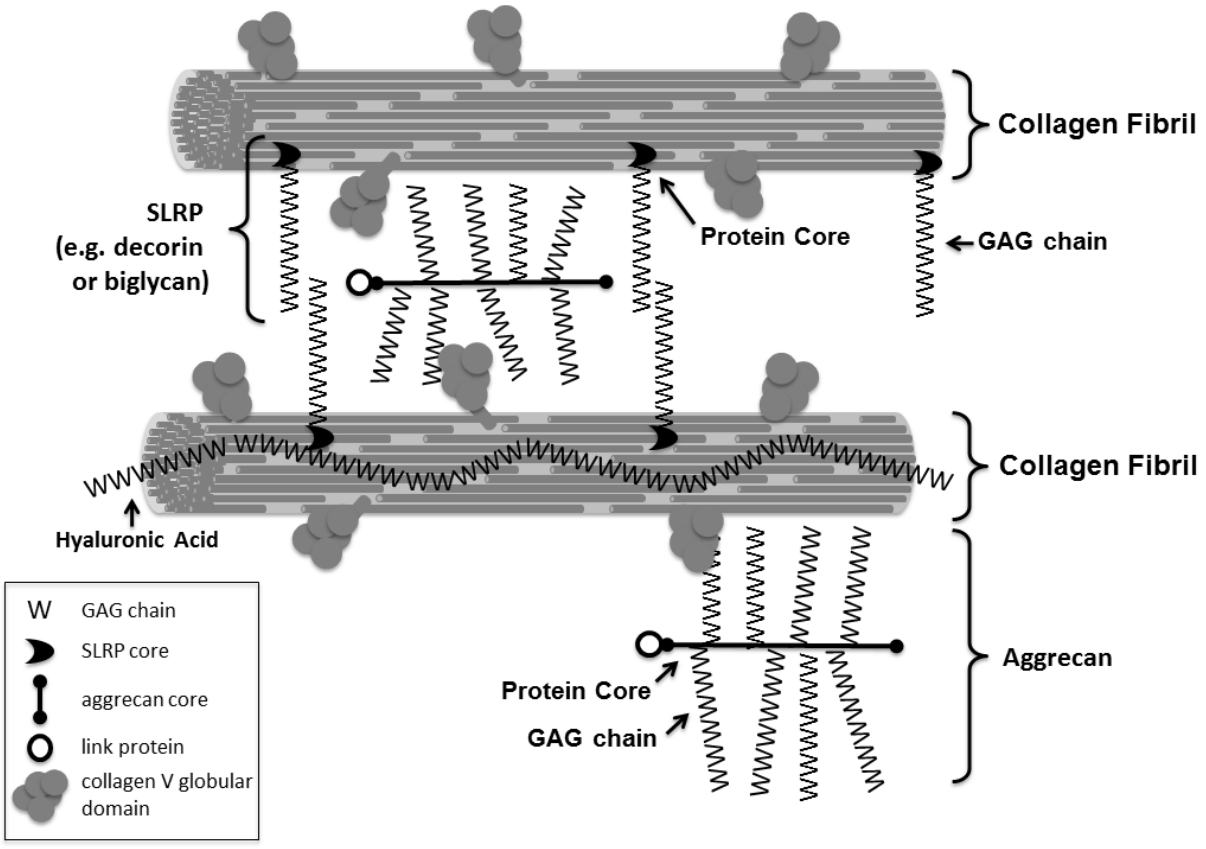


Figure 4.2: A schematic diagram representing the association of aggrecan and the small leucine-rich proteoglycans (SLRPs) biglycan, decorin, fibromodulin and lumican, as well as the glycosaminoglycans (GAGs), with the collagen network.

4.1.3 Sex-specific associations

Within this dissertation sex-specific associations were identified for *BGN*, *DCN* and *COL5A1*. The reasons for this remain unknown but several suggestions were made. Various intrinsic risk factors described in this dissertation are female-specific, such as BMI, ACL size and anatomy, and joint laxity. Hormones are believed to greatly influence ACL injury susceptibility (Chapter 1, Section 1.3.1.2), as sex hormones modulate the synthesis and degradation of the SLRPs, such as *BGN* and *DCN*.[333,334] Several other genetic studies investigating ACL injury risk have also observed sex-specific associations.[37,38] Oestrogen and progesterone receptor sites have been located in the ACL,[52,119] and the sex hormones are suggested to affect ACL structure and composition.[119] The regulation of SLRPs by oestrogen [335-337] may be one of the reasons for the sex-specific associations identified in this dissertation. For example, the sex hormones may alter the availability of the SLRPs and collagens within the ECM. This would therefore effect the ability of these proteins to assist with fibrillogenesis during ECM homeostasis and remodelling, thus compromising the integrity of the ligament structure, resulting in greater injury susceptibility. The role of sex and sex hormones in ACL injury susceptibility is thus another factor that needs to be further investigated in order to determine the exact mechanisms that result in females being at a greater risk for ACL injuries.[28,86]

In summary, this dissertation successfully used molecular genetics as a tool to identify genes within a pathway that may be contributing to the mechanisms that underlie ACL injury risk. It is now apparent that the proteoglycan genes and their role in the regulation of fibrillogenesis, may be one of the biological mechanisms that if disturbed could result in an individual having a higher probability of sustaining an ACL injury. Identifying genes implicated in complex multifactorial phenotypes only partially explains the biological pathway that leads to the condition. Only once gene expression, protein function and protein interactions are understood will the physiological mechanisms become more apparent.[223] Understanding the complex anatomy, function, biomechanics and biology of the ACL is critical to elucidate the mechanisms of injury and improve prevention and treatment techniques.

4.2 Strengths and limitations

This dissertation, as with any research, has strengths and limitations to the studies presented. Although the total sample size of the study was adequate, as determined by Quanto, to detect associations with 80% power, it has become clear that because sex is a risk factor for ACL ruptures and females are at an increased risk,[28,86] males and females need to be analysed together and independently for associations with ACL injuries, to enable the identification of any sex-specific genetic loci. This study was not optimally designed to explore sex-specific risk factors and it was therefore not surprising that when participants were stratified by sex the sample size, especially for the female participants, became underpowered. The percentage of males differed significantly between groups and an equal number of males and females therefore need to be recruited for the ACL and CON groups in order to more effectively compare the similarities and differences between groups.

While effort was made to equally match the CON and ACL participants for the extrinsic risk factors they are exposed to, there were differences in sports participation between the groups. Sports participation and level of play were well documented, but these were all self-reported and in some cases many years after play; it is thus likely that some accounts of sports participation are inaccurate. Both the male and female ACL groups had a greater percentage of participants that engaged in contact and non-contact jumping sports in comparison to their respective CON groups. These types of sports are high risk for the occurrence of ACL ruptures, and it is thus important in the future to recruit a greater percentage of controls that participate in these types of sports as well. Recruitment therefore needs to be focused at rugby and soccer clubs for males and females, as well as netball clubs for females.

Another limitation of this study is in the fact that the weight of participants differed between groups, as those with an ACL rupture weighed significantly more than controls, even when adjusted for sex and age. Weight is an intrinsic risk for ACL ruptures and a confounding variable because increased weight often results in greater knee extension when landing from a jump.[21,33] The weight of participants in the ACL group is however recorded as their weight at the time of first ACL rupture and this can be inaccurate as some participants are recruited many months or years after the occurrence. However this is somewhat out of the control of the investigators, unless the participant had seen a doctor soon after the injury and their weight recorded in a patient file which could be accessed. Weight and height, and thus BMI, were self-reported values and perhaps exact

measurements, including waist circumference, should be taken by investigators to ensure accuracy in the future. There were however no genotype effects on weight and BMI detected (Appendix D, Supplementary Table 1).

It is acknowledged that this dissertation only investigated a small group of genes and a limited number of SNPs that may potentially play a role in ligament homeostasis and the mechanical properties of the ligament. It is therefore important that the need to investigate additional loci within these genes as well as investigating different genes that may be implicated in the risk of sustaining an ACL rupture, is understood. Versican, for example, is another hyalectan proteoglycan that should be investigated with regards to ACL injury risk. Alternative splicing of the gene encoding versican however yields four different mRNA transcripts,[188,189] which adds a level of complication, but our group is currently exploring this. Fibrillogenesis is not the only pathway that may contribute to the mechanisms that underlie ACL injuries and various other pathways also need to be explored (Section 4.3) in order to fully understand the injury mechanisms, as well as to be able to identify individual predisposition to ACL injury.

The great strength to this dissertation is the hypothesis driven approach that was followed. Chapter 1 presented convincing biological evidence of the significant functions of the candidate genes explored in this study with relevance to ligament biology, ECM homeostasis and their potential impact on the pathobiological mechanisms underpinning ACL ruptures.

The value of case-control genetic association studies is that they exploit the differences in the genetic profile of two extreme phenotypes, being the uninjured (CON) and injured (ACL) participants. This case-control genetic association study was also strengthened by the fact that it was conducted, and reported in line with the guidelines provided by the STREGA statement.[323] This statement provides a checklist to be followed during genetic association studies, adapted from the STROBE statement on reporting observational studies. The checklist explicitly describes all rationale that should be considered in deciding on study design and provides guidelines for ethical approval, recruiting participants and carrying out laboratory work. Finally the statement advises on the accurate and unbiased reporting of the results of the research with emphasis on transparency in reporting evidence of genetic association studies. This statement was developed by a multidisciplinary group of scientists and thus reflects a collaborative approach. Consistent with the STREGA laboratory guidelines, multiple control samples were included in the genotyping of participants for the various SNPs investigated. On each 96 well PCR plate run, three repeat samples

and three DNA free controls were included. The DNA free controls were successful 100% of the time and the repeat samples were successful 99.3% of the time, where in one instance a repeat sample was uncalled. Lastly, the genotypes of each of the samples were further checked by an independent person to guarantee quality control. With regards to genotyping, 99.8% of all samples were successfully genotyped for all ten SNPs. All attempts were made to adhere to the strict principles and guidelines of the STREGA report for quality control purposes.

However, as with any novel findings, the associations noted in this study should be repeated in larger independent populations in order to confirm all the observed results.[346]

4.3 Future prospects

The results of this dissertation provide the initial evidence implicating the genes encoding proteoglycans in the risk of ACL injuries. *ACAN* and *DCN* were independently associated with ACL injury susceptibility, and regions overlapping *ACAN*, *BGN*, *LUM* and *DCN* were also implicated in the risk of injury. Variants within *BGN* and *DCN* were found to interact with a variant in *COL5A1* in the modulation of ACL injury risk. This is however only a first look at polymorphisms, within the proteoglycan encoding genes that may be contributing to the mechanisms that underlie individual susceptibility to ACL injuries, and the regions that have been implicated in this study need to be further explored. We hypothesise that variants within the proteoglycan encoding genes may influence protein expression and potentially protein interactions, specifically interactions with the collagen network in the modulation of fibrillogenesis. Proteoglycan genes also function as cell signalling molecules and therefore variants within the genes may also affect ECM homeostasis and turnover, and thus comprise the ability of the ligament to respond to loading, which could lead to injury.[162,305] However, in order to understand the mechanisms that lead to injury, and the exact role of proteoglycans, further investigation needs to take place.

The first step in further exploring the mechanistic role of proteoglycans in ACL ruptures is to confirm the findings observed in an independent population. Associating a variant with ACL injury risk in a single population does not prove that the polymorphism is actually associated with the pathophysiology of the injury. The confirmation of genetic associations in independent populations is required to evaluate the validity of the association or the contribution of a particular gene to the risk of the multifactorial condition. Investigating a different population would also potentially assist in narrowing the genetic interval of interest that may underlie predisposition to ACL ruptures, because LD relationships differ across populations.[346] If the association of a particular variant persists in a different population, where the frequency of the variant differs and genomic variation is greater, it suggests that the variant may have, or be tightly linked to a variant with a functional role in the mechanism of injury. It is thus crucial to repeat this study in an independent Caucasian population to provide the findings with credibility,[346] as well as to repeat this investigation in a different population of dissimilar ancestry, in order to determine whether the association persists or not. The genetic risks for ACL ruptures may differ between population groups and therefore this needs to be examined.

There are very limited studies investigating genetic variants in non-Caucasian populations, but our research group has already begun recruiting individuals of other South African populations, including Black and Mixed Ancestry, in order to repeat the genetic associations identified in this study and previous.[36-39] African populations have higher levels of genetic variation and diversity in comparison to non-African populations, who contain a subset of the genetic diversity seen in Africa.[347,348] The high genetic variation observed in African populations could assist in defining the particular genetic intervals associated with predisposition to ACL ruptures, and it is thus important to evaluate genetic associations in the broader populations of South Africa. The similarities or differences between variants in different populations will assist in identifying the relative contribution of each variant to the risk of ACL ruptures.

This dissertation implicated several regions overlapping proteoglycan genes in ACL injury susceptibility which now need to be narrowed down. Sequencing of the proteoglycan genes can be used to identify the important genetic regions responsible for the modulation of ACL injury risk. Targeted sequencing using next generation sequencing technology, and several bioinformatics tools can be employed to narrow the genetic interval of interest, as well as to identify additional novel polymorphisms which may be contributing to ACL injury risk.[349] Genotyping variants identified in the target region, in a large cohort, enables you to fine-map the association.[349] Next generation sequencing is however an expensive process and thus the confirmation of the genetic associations observed and the investigation of the associations in different populations is the first critical step in the discovery process.

Once sequencing has been used to verify the genetic interval of interest and subsequent associated polymorphisms confirmed, the region should be further scrutinised. Associated variants should be subjected to functional studies in order to determine the functional significance of particular polymorphic alleles. Reporter assays can be used to determine gene activity and mRNA stability when alternate alleles are present.[139,220] Bioinformatic tools can be applied in order to determine whether any binding sites are present in the genetic interval that was narrowed down by sequencing.[139] Polymorphisms within binding target regions could compromise the interaction of the gene with various other molecules, and polymorphisms within proteoglycans in particular could alter interactions with the collagen network and negatively affect ligament homeostasis. Determining the functional consequences of associated polymorphisms will aid in understanding the molecular mechanisms that are responsible for the pathway that leads to an ACL rupture. Exploring the interactions and functional relationships between multiple genes will help us to comprehend the

biological dynamics of the ligament. This will enable us to become closer to understanding a cause-effect relationship that could possibly be intervened.[220]

This dissertation investigated the interaction of the proteoglycan genes with *COL5A1* (Chapter 3), and the possibility that variation in these interactions could modulate ligament fibrillogenesis and lead to injury. Proteoglycans are unique, multi-functional molecules with diverse structural and signalling roles within ligaments, and so far we have only touched on the role of proteoglycans in fibrillogenesis. Proteoglycans bind a variety of proteins and cell-surface receptors with high specificity, activating various cell signalling pathways within the matrix (Recently reviewed by Chen and Birk, 2013).[40] They are able to interact with fibrillar collagens as well as growth factors and cytokines, to regulate cellular proliferation and ECM remodelling. The important cell signalling roles of proteoglycans, in various pathways responsible for ECM turnover in the normal and injured state, should also be investigated in the future.

TGF-β is suggested to have an important role in switching the healing process from normal to pathological,[350,351] and during tissue repair TGF-β is essential for the synthesis of a new matrix.[352] Both biglycan and decorin interact with all three isoforms of this regulatory cytokine, with several biological implications on cellular proliferation and differentiation.[260] TGF-β modulates inflammatory responses, angiogenesis, proteoglycan deposition, and stimulates collagen production.[353] Biglycan and decorin can significantly and dose-dependently decrease TGF-β induced gene expression in vitro,[241,259] but there is also conflicting reports that these SLRPs can in some cases increase the expression of TGF-β.[189] We can hypothesise that variants in the *BGN* and *DCN* genes (Chapter 2) may potentially alter the interactions of these proteins with TGF-β in the ECM. The involvement of TGF-β in the regulation of proteoglycan production, collagen synthesis, MMP synthesis and angiogenesis [206,240,353-355] (Figure 4.3) suggests that an alteration in the amount of TGF-β present may modulate any one of these pathways and compromise ECM homeostasis and remodelling. TGF-β is evidently an important cytokine in the regulation of ECM remodelling and genetic variations within the genes encoding the three isoforms of TGF-β should be investigated for an association with ACL injury risk. The interaction between TGF-β and the proteoglycans should also be explored as was done with *COL5A1* in Chapter 3. The functional consequences of such variants or interactions should then be examined in order to understand the dynamic between these molecules, disturbance of which may lead to injury susceptibly.

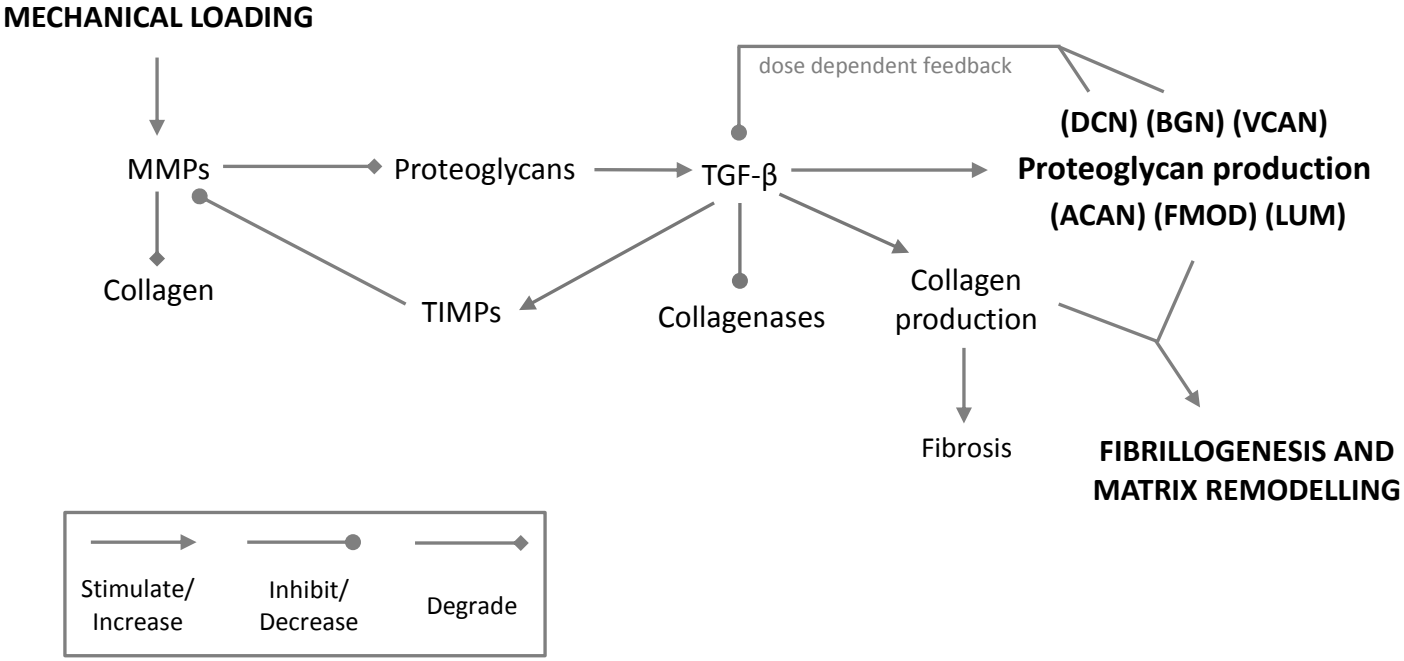


Figure 4.3: A schematic representation of the interaction of proteoglycans with TGF-β. Abbreviations: ACAN, aggrecan; BCN, biglycan; DCN, decorin; FMOD, fibromodulin; LUM, lumican; MMPs, matrix metalloproteinases; TGF-β, transforming growth factor-β; TIMPs, tissue inhibitor metalloproteinases; VCAN, versican.

During the regulation of innate immunity functions, DAMPs, or 'danger signals', are recognised by the TLRs 2 and 4.[356,357] Biglycan and decorin,[264,358] versican,[359] and HA [360] are able to act as DAMPS, endogenous ligands of the TLRs. Under normal physiological conditions the proteoglycans are sequestered, but during tissue stress or damage, the ECM undergoes proteolytic digestion during remodelling, and the proteoglycans become soluble potent danger signals.[361] Although the proteoglycans all signal through the same TLRs, they activate different responses with different biological consequences.[242]

Biglycan is an early response gene under inflammatory conditions and generates a rapid reaction.[361] Various pro-inflammatory cytokines and chemokines such as TNF-α and IL-1β are downstream effector molecules of biglycan-TLR2/4 signalling.[242] Biglycan is therefore capable of prompting autonomous sterile (non-pathogen-mediated) inflammation. Decorin is also an endogenous ligand for TLR2/4, which when bound triggers a cascade of cell signalling molecules such as the synthesis of cytokines, TNF-α and IL-12.[264] Decorin also inhibits TGF-β1 (independent of TLR2/4) and IL-10, immunosuppressive and anti-inflammatory molecules respectively, thereby creating a pro-inflammatory environment.[264] Lumican has also been implicated as an important mediator of the

cell signalling pathway by inducing the release of pro-inflammatory cytokines involved in the recruitment of macrophages and neutrophils,[362] however, lumican does not directly interact with TLR2/4 but rather with their adaptor molecule CD14.[363] Versican is also implicated in cell signalling and is suggested to act as a ligand to the TLR2/4 heterodimer and the CD14 adaptor in order to activate the production of pro-inflammatory cytokines.[359]

The interaction of various proteoglycans with the TLR2/4 receptors highlights the versatile roles of these proteins that have signals embedded in their core and GAG chains. The role of proteoglycans in these pathways needs to be explored and the genes encoding molecules within the cell signalling pathways, involved in ECM remodelling following mechanical loading, should be investigated for associations with ACL injury susceptibility. This adds further complexity to the functions of proteoglycans, whose interactions with cell signalling genes may pose as possible therapeutic targets for inflammatory conditions.[242]

The interaction of proteoglycans with the group of ADAMTS (a disintegrin and metalloproteinase with thrombospondin motifs) proteinases can also be explored. The ADAMTS are a group of enzymes that, among other functions, are involved in regulating ECM homeostasis. Deficiency in ADAMTS5, a proteinase that interacts with aggrecan, results in abnormal tendon collagen organisation, affecting the biomechanical properties of the tendon.[364] ADAMTS14 and tissue inhibitor metalloproteinase 2 (TIMP2), have previously been associated with the age of onset, and risk of Achilles tendon pathology respectively.[365] The TIMPS are involved in the compositional regulation of the ECM, and thus the interactions of the proteoglycans, ADAMTS and TIMPS in ECM regulation is another process that could be investigated to add to the polygenic risk model for ACL injuries.

4.4 Clinical significance

Understanding the biological dynamic interaction between proteins, such as proteoglycans with their receptors, and protein-protein interactions, will not only aid to elucidate the mechanisms that bring about an ACL injury, but these interactions could become therapeutic targets in the prevention and treatment of injury. An endothelial cell specific DS proteoglycan has previously been observed to be overexpressed in cancerous tumours and directly linked to the tumour progression. This proteoglycan has thus been identified as target in cancer therapies as well as a biomarker for inflammatory disorders and tumour progression.[366] Lubrican, a CS/KS proteoglycan, has too been identified as a target for immunotherapy in the treatment of various melanomas.[367,368] It is therefore plausible to consider proteoglycans as both biomarkers and targets in the prevention and treatment of multifactorial conditions such as ACL injuries.

Proteoglycans will serve as an important part of a polygenic, multifactorial risk model that can be used to identify individuals that have an increased predisposition to sustaining an ACL injury. Those individuals may benefit from injury preventions strategies in the form of training techniques, as well as possible therapeutic interventions including personalised medicine. A recent study managed to inhibit the loss of proteoglycans from the matrix of bovine tendons using highly sulfated polysaccharides, thus highlighting the opportunity of targeting proteoglycans and interacting proteins in the prevention of injuries.[369] Interventions that are able to sustain ligament homeostasis and biomechanical composition in an otherwise compromised tissue, could prevent injuries.

The potential of using proteoglycans and GAGs as injury biomarkers is already being fulfilled. KS levels in the synovial fluid of ACL injured knees are associated with the number of high-grade cartilage lesions,[370] and aggrecan and GAG content has been associated with the degree of tendon pathology.[308] Although proteoglycan biomarkers may be used to detect the severity of injury, magnetic resonance imaging is the best way to accurately diagnose an ACL injury without surgical incisions.

With regards to using proteoglycans as a part of treatment following an ACL injury: intra-articular expression of proteoglycan 4 has recently been identified to protect against the development of OA in mice [371] by increasing lubrication within the joint. Individuals that have sustained an ACL rupture are at an increased risk of developing OA of the knee joint [32] and such treatments may be possible for humans in the future. TGF-β has been highlighted as an important cytokine in the regulation of

ECM remodelling in the injured and uninjured state; this molecule and its interactions with the proteoglycans may also serve as a therapeutic target in both prevention and treatment strategies. Following further research regarding the binding mechanisms of proteoglycans, the possibility also arises of pharmacologically influencing the inflammation experienced with injury,[163] and the interaction of proteoglycans with the TLRs could possibly be intervened. All these possibilities are however only plausible once the effects of the interactions of both proteoglycans and other effector molecules, with each other and with their receptors, are fully understood.

A recent study investigated the potential benefit of platelet-rich plasma (PRP) on the graft healing process following ACL reconstruction in canines. PRP was observed to increase mRNA expression levels of *COL1A1*, *COL3A1*, *TGF-β1*, *MMP1*, *BGN* and *TIMP-1*, while *DCN* and *MMP13* levels were decreased in ACL autographs, thereby positively affecting ECM remodelling following surgery.[372] PRP has thus been suggested as a potential treatment in improving ACL injury recovery following surgery by affecting the expression of various molecules that include the proteoglycans. Exogenous expression of proteoglycans may also aid the healing process following an ACL injury.

Gene therapy is also a prospective treatment in the prevention and management of ACL injuries. This allows extra synthesis of certain molecules by introducing coding sequences that prolong the synthesis of bioactive anabolic agents.[373] Delivery mechanisms (vectors) are used to incorporate the coding sequence into the cell or matrix. The intra-articular injection of retroviral vectors carrying IL-1 receptor agonist and soluble TNF-α receptor into rabbit knee joints emitted an anti-inflammatory and chondro-protective response in the injected and contralateral joint, improving the induced OA.[374] IL-1 receptor agonist has also been effectively used via gene therapy in the treatment of RA in a cohort of 472 patients.[375] Gene therapy thus appears to be an effective therapeutic intervention and similar methodology could be used in the treatment of ACL injuries to improve healing. A review by Nixon *et al.* (2007) highlights the positive results observed with gene therapy that combines cell supplementation and cells overexpressing certain growth factors, on musculoskeletal repair.[373] Some growth factors and cytokines used in gene therapy stimulated repair include, as mentioned, TGF-β, IL-1 receptor agonist and TNF-α, all of which are factors involved in ECM remodelling following mechanical loading.

Increased knowledge of the molecular mechanisms of ACL injuries is of course required before gene therapies become possible. Once appropriate structural or signalling molecules with potential

therapeutic effects are identified, gene transfer vectors encoding such molecules can be transferred to the knee joint in trials, and the consequences observed.[376]

In summary, the interactions of proteoglycans with their receptors as well as with other molecules, such as TGF-β, may present as possible drug targets in order to disrupt pathological mechanisms that lead to injury. The proteoglycans however interact with an array of molecules in the regulation of ECM remodelling and other complex cell signalling pathways such as apoptosis and angiogenesis. Only a few interactions have been mentioned in this dissertation as directions for future research and the possible clinical implications of these interactions as drug targets, but there is much more knowledge to be gained regarding the dynamic role of proteoglycans in the ligament tissue. It is particularly important to identify pharmacological targets that may aid in the treatment of ACL injuries by influencing ECM remodelling, fibrillogenesis and ligament healing as this ligament typically has a poor healing capacity.[73]

Only once we advance our knowledge regarding the contribution of each molecule to ligament homeostasis and the mechanisms that result in ACL injury may potential treatments become a reality for human patients. Genetics and the results of this dissertation have however taught us that individual genetic variation in sequence signatures could result in predisposition to injury due to disruption of ligament properties via improper regulation of various protein expression or protein interactions. Different individuals will thus have varying factors that are either lacking within their ligaments or factors that are being over-expressed. With greater understanding, these differences could be targeted in the prevention or treatment of injury, but inter-individual genetic variation is likely to require personalised therapeutic intervention.

4.5 Concluding remarks

In conclusion, this dissertation identified independent associations between the *ACAN* and *DCN* genes, and ACL injury susceptibility. Regions overlapping *ACAN*, *BGN*, *LUM* and *DCN*, and interactions between the proteoglycan genes and *COL5A1* were also implicated in the risk of injury. This implies that the proteoglycans, and their role in the regulation of ligament structure organisation and fibrillogenesis, may be a pathway involved in the mechanism of ACL injuries. The results of this dissertation should be repeated in an independent population, and the regions within the proteoglycan genes implicated in this dissertation should be further interrogated.

With further research, the results presented may have a potential impact on the prevention and treatment of ACL injuries in the future. Polymorphisms within the proteoglycan genes can potentially be used together with previously associated polymorphisms in a polygenic risk profile to determine which individuals are at a greater risk of sustaining an ACL injury, and prevention strategies put in place. It may also be possible to target proteoglycans and their interactions, in the treatment of injuries, once the mechanisms are fully understood.

References

1 McGinnis JM. The public health burden of a sedentary lifestyle. Med Sci Sports Exerc 1992;24:196-200.

2 Matheson GO, Klugl M, Engebretsen L, Bendiksen F, Blair SN, Borjesson M, Budgett R, Derman W, Erdener U, Ioannidis JP, Khan KM, Martinez R, van Mechelen W, Mountjoy M, Sallis RE, Schwellnus M, Shultz R, Soligard T, Steffen K, Sundberg CJ, Weiler R, Ljungqvist A. Prevention and Management of Non-Communicable Disease: The IOC Consensus Statement, Lausanne 2013. Sports Med 2013;43:1075-1088.

3 Jones BH, Cowan DN, Knapik JJ. Exercise, training and injuries. Sports Med 1994;18:202-214.

4 Meeuwisse WH. Assessing causation in sport injury:a multifactorial model. Clin J Sport Med;4:166–170.

5 Collins M, Raleigh SM. Genetic risk factors for musculoskeletal soft tissue injuries. Med Sport Sci 2009;54:136-149.

6 Suchak AA, Bostick G, Reid D, Blitz S, Jomha N. The incidence of Achilles tendon ruptures in Edmonton, Canada. Foot Ankle Int 2005;26:932-936.

7 Scott A, Ashe MC. Common tendinopathies in the upper and lower extremities. Curr Sports Med Rep 2006;5:233-241.

8 Harvie P, Ostlere SJ, Teh J, McNally EG, Clipsham K, Burston BJ, Pollard TC, Carr AJ. Genetic influences in the aetiology of tears of the rotator cuff. Sibling risk of a full-thickness tear. J Bone Joint Surg Br 2004;86:696-700.

9 Flynn RK, Pedersen CL, Birmingham TB, Kirkley A, Jackowski D, Fowler PJ. The familial predisposition toward tearing the anterior cruciate ligament: a case control study. Am J Sports Med 2005;33:23-28.

10 Abbassian A, Thomas R. Ankle ligament injuries. Br J Hosp Med (Lond) 2008;69:339-343.

11 Rees JD, Wilson AM, Wolman RL. Current concepts in the management of tendon disorders. Rheumatology 2006;45:508-521.

12 Jarvinen TA, Kannus P, Maffulli N, Khan KM. Achilles tendon disorders: etiology and epidemiology. Foot Ankle Clin 2005;10:255-266.

13 Kannus P, Jarvinen TL, Jarvinen TA, Jarvinen M, Parkkari J, Khan K. Painful Achilles tendon and its treatment. Scand J Med Sci Sports 2004;14:69-71.

14 Riley G. The pathogenesis of tendinopathy. A molecular perspective. Rheumatology 2004;43:131-142.

15 Childs SG. Pathogenesis of anterior cruciate ligament injury. Orthop Nurs 2002;21:35-40.

16 Bahr R, Krosshaug T. Understanding injury mechanisms: a key component of preventing injuries in sport. Br J Sports Med 2005;39:324-329.

17 Barr KP. Rotator cuff disease. Phys Med Rehabil Clin N Am 2004;15:475-491.

18 Gomoll AH, Katz JN, Warner JJ, Millett PJ. Rotator cuff disorders: recognition and management among patients with shoulder pain. Arthritis Rheum 2004;50:3751-3761.

19 Hewett TE, Myer GD, Ford KR. Anterior cruciate ligament injuries in female athletes: Part 1, mechanisms and risk factors. Am J Sports Med 2006;34:299-311.

20 Matava MJ, Purcell DB, Rudzki JR. Partial-thickness rotator cuff tears. Am J Sports Med 2005;33:1405-1417.

21 Posthumus M, Collins M, September AV, Schwellnus MP. The intrinsic risk factors for ACL ruptures: an evidence-based review. Phys Sportsmed 2011;39:62-73.

22 Murakami H, Shinomiya N, Kikuchi T, Yoshihara Y, Nemoto K. Upregulated expression of inducible nitric oxide synthase plays a key role in early apoptosis after anterior cruciate ligament injury. J Orthop Res 2006;24:1521-1534.

23 Beynnon BD, Johnson RJ, Abate JA, Fleming BC, Nichols CE. Treatment of anterior cruciate ligament injuries, part I. Am J Sports Med 2005;33:1579-1602.

24 Woo SL, Abramowitch SD, Kilger R, Liang R. Biomechanics of knee ligaments: injury, healing, and repair. J Biomech 2006;39:1-20.

25 Arendt E, Dick R. Knee injury patterns among men and women in collegiate basketball and soccer. NCAA data and review of literature. Am J Sports Med 1995;23:694-701.

26 Arendt EA, Agel J, Dick R. Anterior cruciate ligament injury patterns among collegiate men and women. J Athl Train 1999;34:86-92.

27 Ireland ML. Anterior cruciate ligament injury in female athletes: epidemiology. J Athl Train 1999;34:150-154.

28 Sutton KM, Bullock JM. Anterior cruciate ligament rupture: differences between males and females. J Am Acad Orthop Surg 2013;21:41-50.

29 Claes S, Verdonk P, Forsyth R, Bellemans J. The "ligamentization" process in anterior cruciate ligament reconstruction: what happens to the human graft? A systematic review of the literature. Am J Sports Med 2011;39:2476-2483.

30 Brooks JH, Fuller CW, Kemp SP, Reddin DB. Epidemiology of injuries in English professional rugby union: part 1 match injuries. Br J Sports Med 2005;39:757-766.

31 Deacon A, Bennell K, Kiss ZS, Crossley K, Brukner P. Osteoarthritis of the knee in retired, elite Australian Rules footballers. Med J Aust 1997;166:187-190.

32 Higuchi H, Shirakura K, Kimura M, Terauchi M, Shinozaki T, Watanabe H, Takagishi K. Changes in biochemical parameters after anterior cruciate ligament injury. Int Orthop 2006;30:43-47.

33 Griffin LY, Albohm MJ, Arendt EA, Bahr R, Beynnon BD, Demaio M, Dick RW, Engebretsen L, Garrett WE,Jr, Hannafin JA, Hewett TE, Huston LJ, Ireland ML, Johnson RJ, Lephart S, Mandelbaum

BR, Mann BJ, Marks PH, Marshall SW, Myklebust G, Noyes FR, Powers C, Shields C,Jr, Shultz SJ, Silvers H, Slauterbeck J, Taylor DC, Teitz CC, Wojtys EM, Yu B. Understanding and preventing noncontact anterior cruciate ligament injuries: a review of the Hunt Valley II meeting, January 2005. Am J Sports Med 2006;34:1512-1532.

34 Harner CD, Paulos LE, Greenwald AE, Rosenberg TD, Cooley VC. Detailed analysis of patients with bilateral anterior cruciate ligament injuries. Am J Sports Med 1994;22:37-43.

35 Khoschnau S, Melhus H, Jacobson A, Rahme H, Bengtsson H, Ribom E, Grundberg E, Mallmin H, Michaelsson K. Type I collagen alpha1 Sp1 polymorphism and the risk of cruciate ligament ruptures or shoulder dislocations. Am J Sports Med 2008;36:2432-2436.

36 Posthumus M, September AV, Keegan M, O'Cuinneagain D, Van der Merwe W, Schwellnus MP, Collins M. Genetic risk factors for anterior cruciate ligament ruptures: COL1A1 gene variant. Br J Sports Med 2009;43:352-356.

37 Posthumus M, September AV, O'Cuinneagain D, van der Merwe W, Schwellnus MP, Collins M. The COL5A1 gene is associated with increased risk of anterior cruciate ligament ruptures in female participants. Am J Sports Med 2009;37:2234-2240.

38 Posthumus M, September AV, O'Cuinneagain D, van der Merwe W, Schwellnus MP, Collins M. The association between the COL12A1 gene and anterior cruciate ligament ruptures. Br J Sports Med 2010;44:1160-1165.

39 Posthumus M, Collins M, van der Merwe L, O'Cuinneagain D, van der Merwe W, Ribbans WJ, Schwellnus MP, Raleigh SM. Matrix metalloproteinase genes on chromosome 11q22 and the risk of anterior cruciate ligament (ACL) rupture. Scand J Med Sci Sports 2011;22:522-533.

40 Chen S, Birk DE. The regulatory roles of small leucine-rich proteoglycans in extracellular matrix assembly. FEBS J 2013;280:2120-2137.

41 Corsi A, Xu T, Chen XD, Boyde A, Liang J, Mankani M, Sommer B, Iozzo RV, Eichstetter I, Robey PG, Bianco P, Young MF. Phenotypic effects of biglycan deficiency are linked to collagen fibril abnormalities, are synergized by decorin deficiency, and mimic Ehlers-Danlos-like changes in bone and other connective tissues. J Bone Miner Res 2002;17:1180-1189.

42 Ameye L, Young MF. Mice deficient in small leucine-rich proteoglycans: novel in vivo models for osteoporosis, osteoarthritis, Ehlers-Danlos syndrome, muscular dystrophy, and corneal diseases. Glycobiology 2002;12:107-16.

43 Danielson KG, Baribault H, Holmes DF, Graham H, Kadler KE, Iozzo RV. Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. J Cell Biol 1997;136:729-743.

44 Kalamajski S, Oldberg A. The role of small leucine-rich proteoglycans in collagen fibrillogenesis. Matrix Biol 2010;29:248-253.

45 Frank CB. Ligament structure, physiology and function. J Musculoskelet Neuronal Interact 2004;4:199-201.

46 Viidik A. Simultaneous mechanical and light microscopic studies of collagen fibers. Z Anat Entwicklungsgesch 1972;136:204-212.

47 Benjamin M, Ralphs JR. The cell and developmental biology of tendons and ligaments. Int Rev Cytol 2000;196:85-130.

48 Sakane M, Fox RJ, Woo SL, Livesay GA, Li G, Fu FH. In situ forces in the anterior cruciate ligament and its bundles in response to anterior tibial loads. J Orthop Res 1997;15:285-293.

49 Moses B, Orchard J, Orchard J. Systematic review: Annual incidence of ACL injury and surgery in various populations. Res Sports Med 2012;20:157-179.

50 Mysid. Right knee. http://en.wikipedia org/wiki/File:Knee_diagram svg April 2011;October 2013.

51 Bernard M, Hertel P, Hornung H, Cierpinski T. Femoral insertion of the ACL. Radiographic quadrant method. Am J Knee Surg 1997;10:14-21.

52 Silvers HJ, Mandelbaum BR. Prevention of anterior cruciate ligament injury in the female athlete. Br J Sports Med 2007;41:52-59.

53 Girgis FG, Marshall JL, Monajem A. The cruciate ligaments of the knee joint. Anatomical, functional and experimental analysis. Clin Orthop Relat Res 1975;106:216-231.

54 Harner CD, Baek GH, Vogrin TM, Carlin GJ, Kashiwaguchi S, Woo SL. Quantitative analysis of human cruciate ligament insertions. Arthroscopy 1999;15:741-749.

55 Amis AA, Dawkins GP. Functional anatomy of the anterior cruciate ligament. Fibre bundle actions related to ligament replacements and injuries. J Bone Joint Surg Br 1991;73:260-267.

56 Remer EM, Fitzgerald SW, Friedman H, Rogers LF, Hendrix RW, Schafer MF. Anterior cruciate ligament injury: MR imaging diagnosis and patterns of injury. Radiographics 1992;12:901-915.

57 Hollis JM, Takai S, Adams DJ, Horibe S, Woo SL. The effects of knee motion and external loading on the length of the anterior cruciate ligament (ACL): a kinematic study. J Biomech Eng 1991;113:208-214.

58 Bach JM, Hull ML, Patterson HA. Direct measurement of strain in the posterolateral bundle of the anterior cruciate ligament. J Biomech 1997;30:281-283.

59 Petersen W, Tillmann B. Anatomy and function of the anterior cruciate ligament. Orthopade 2002;31:710-718.

60 Juneja S, and Veillette C. Defects in Tendon, Ligament, and Enthesis in Response to Genetic Alterations in Key Proteoglycans and Glycoproteins: A Review. Arthritis 2013;2013:1-30.

61 Duthon VB, Barea C, Abrassart S, Fasel JH, Fritschy D, Menetrey J. Anatomy of the anterior cruciate ligament. Knee Surg Sports Traumatol Arthrosc 2006;14:204-213.

62 Arnoczky SP. Anatomy of the anterior cruciate ligament. Clin Orthop Relat Res 1983;172:19-25.

63 Zantop T, Petersen W, Sekiya JK, Musahl V, Fu FH. Anterior cruciate ligament anatomy and function relating to anatomical reconstruction. Knee Surg Sports Traumatol Arthrosc 2006;14:982-992.

64 Norwood LA, Cross MJ. Anterior cruciate ligament: functional anatomy of its bundles in rotatory instabilities. Am J Sports Med 1979;7:23-26.

65 Otsubo H, Shino K, Suzuki D, Kamiya T, Suzuki T, Watanabe K, Fujimiya M, Iwahashi T, Yamashita T. The arrangement and the attachment areas of three ACL bundles. Knee Surg Sports Traumatol Arthrosc 2012;20:127-134.

66 Tantisricharoenkul G, Linde-Rosen M, Araujo P, Zhou J, Smolinski P, Fu FH. Anterior cruciate ligament: an anatomical exploration in humans and in a selection of animal species. Knee Surg Sports Traumatol Arthrosc 2013.

67 Strocchi R, de Pasquale V, Gubellini P, Facchini A, Marcacci M, Buda R, Zaffagnini S, Ruggeri A. The human anterior cruciate ligament: histological and ultrastructural observations. J Anat 1992;180:515-519.

68 Spalazzi JP, Boskey AL, Pleshko N, Lu HH. Quantitative mapping of matrix content and distribution across the ligament-to-bone insertion. PLoS One 2013.

69 Hashemi J, Chandrashekar N, Gill B, Beynnon BD, Slauterbeck JR, Schutt RC,Jr, Mansouri H, Dabezies E. The geometry of the tibial plateau and its influence on the biomechanics of the tibiofemoral joint. J Bone Joint Surg Am 2008;90:2724-2734.

70 Beynnon BD, Johnson RJ, Fleming BC, Peura GD, Renstrom PA, Nichols CE, Pope MH. The effect of functional knee bracing on the anterior cruciate ligament in the weightbearing and nonweightbearing knee. Am J Sports Med 1997;25:353-359.

71 Yoon KH, Yoo JH, Kim KI. Bone contusion and associated meniscal and medial collateral ligament injury in patients with anterior cruciate ligament rupture. J Bone Joint Surg Am 2011;93:1510-1518.

72 Collins JE, Katz JN, Donnell-Fink LA, Martin SD, Losina E. Cumulative incidence of ACL reconstruction after ACL injury in adults: role of age, sex, and race. Am J Sports Med 2013;41:544-549.

73 Bach BR,Jr, Levy ME, Bojchuk J, Tradonsky S, Bush-Joseph CA, Khan NH. Single-incision endoscopic anterior cruciate ligament reconstruction using patellar tendon autograft. Minimum two-year follow-up evaluation. Am J Sports Med 1998;26:30-40.

74 Dallalana RJ, Brooks JH, Kemp SP, Williams AM. The epidemiology of knee injuries in English professional rugby union. Am J Sports Med 2007;35:818-830.

75 Kim S, Bosque J, Meehan JP, Jamali A, Marder R. Increase in outpatient knee arthroscopy in the United States: a comparison of National Surveys of Ambulatory Surgery, 1996 and 2006. J Bone Joint Surg Am 2011;93:994-1000.

76 Ahlden M, Samuelsson K, Sernert N, Forssblad M, Karlsson J, Kartus J. The Swedish National Anterior Cruciate Ligament Register: a report on baseline variables and outcomes of surgery for almost 18,000 patients. Am J Sports Med 2012;40:2230-2235.

77 Lind M, Menhert F, Pedersen AB. Incidence and outcome after revision anterior cruciate ligament reconstruction: results from the Danish registry for knee ligament reconstructions. Am J Sports Med 2012;40:1551-1557.

78 Nagineni CN, Amiel D, Green MH, Berchuck M, Akeson WH. Characterization of the intrinsic properties of the anterior cruciate and medial collateral ligament cells: an in vitro cell culture study. J Orthop Res 1992;10:465-475.

79 Schmidt CC, Georgescu HI, Kwoh CK, Blomstrom GL, Engle CP, Larkin LA, Evans CH, Woo SL. Effect of growth factors on the proliferation of fibroblasts from the medial collateral and anterior cruciate ligaments. J Orthop Res 1995;13:184-190.

80 Wang Y, Tang Z, Xue R, Singh GK, Shi K, Lv Y, Yang L. Combined effects of TNF-alpha, IL-1beta, and HIF-1alpha on MMP-2 production in ACL fibroblasts under mechanical stretch: An in vitro study. J Orthop Res 2011;29:1008-1014.

81 Xie J, Wang C, Huang DY, Zhang Y, Xu J, Kolesnikov SS, Sung KL, Zhao H. TGF-beta1 induces the different expressions of lysyl oxidases and matrix metalloproteinases in anterior cruciate ligament and medial collateral ligament fibroblasts after mechanical injury. J Biomech 2013;46:890-898.

82 Fetto JF, Marshall JL. The natural history and diagnosis of anterior cruciate ligament insufficiency. Clin Orthop Relat Res 1980;147:29-38.

83 Lohmander LS, Ostenberg A, Englund M, Roos H. High prevalence of knee osteoarthritis, pain, and functional limitations in female soccer players twelve years after anterior cruciate ligament injury. Arthritis Rheum 2004;50:3145-3152.

84 Gianotti SM, Marshall SW, Hume PA, Bunt L. Incidence of anterior cruciate ligament injury and other knee ligament injuries: a national population-based study. J Sci Med Sport 2009;12:622-627.

85 September AV, Schwellnus MP, Collins M. Tendon and ligament injuries: the genetic component. Br J Sports Med 2007;41:241-246.

86 Renstrom P, Ljungqvist A, Arendt E, Beynnon B, Fukubayashi T, Garrett W, Georgoulis T, Hewett TE, Johnson R, Krosshaug T, Mandelbaum B, Micheli L, Myklebust G, Roos E, Roos H, Schamasch P, Shultz S, Werner S, Wojtys E, Engebretsen L. Non-contact ACL injuries in female athletes: an International Olympic Committee current concepts statement. Br J Sports Med 2008;42:394-412.

87 Bahr R, Holme I. Risk factors for sports injuries--a methodological approach. Br J Sports Med 2003;37:384-392.

88 Orchard JW, Powell JW. Risk of knee and ankle sprains under various weather conditions in American football. Med Sci Sports Exerc 2003;35:1118-1123.

89 Pasanen K, Parkkari J, Rossi L, Kannus P. Artificial playing surface increases the injury risk in pivoting indoor sports: a prospective one-season follow-up study in Finnish female floorball. Br J Sports Med 2008;42:194-197.

90 Hershman EB, Anderson R, Bergfeld JA, Bradley JP, Coughlin MJ, Johnson RJ, Spindler KP, Wojtys E, Powell JW, National Football League Injury and Safety Panel. An analysis of specific lower extremity injury rates on grass and FieldTurf playing surfaces in National Football League Games: 2000-2009 seasons. Am J Sports Med 2012;40:2200-2205.

91 Ekstrand J, Nigg BM. Surface-related injuries in soccer. Sports Med 1989;8:56-62.

92 Rishiraj N, Taunton JE, Lloyd-Smith R, Woollard R, Regan W, Clement DB. The potential role of prophylactic/functional knee bracing in preventing knee ligament injury. Sports Med 2009;39:937-960.

93 Rovere GD, Haupt HA, Yates CS. Prophylactic knee bracing in college football. Am J Sports Med 1987;15:111-116.

94 Deppen RJ, Landfried MJ. Efficacy of prophylactic knee bracing in high school football players. J Orthop Sports Phys Ther 1994;20:243-246.

95 Clarke KS, Buckley WE. Women's injuries in collegiate sports. A preliminary comparative overview of three seasons. Am J Sports Med 1980;8:187-191.

96 Engstrom B, Johansson C, Tornkvist H. Soccer injuries among elite female players. Am J Sports Med 1991;19:372-375.

97 Ferretti A, Papandrea P, Conteduca F, Mariani PP. Knee ligament injuries in volleyball players. Am J Sports Med 1992;20:203-207.

98 Gray J, Taunton JE, McKenzie DC, Clement DB, McConkey JP, Davidson RG. A survey of injuries to the anterior cruciate ligament of the knee in female basketball players. Int J Sports Med 1985;6:314-316.

99 Gwinn DE, Wilckens JH, McDevitt ER, Ross G, Kao TC. The relative incidence of anterior cruciate ligament injury in men and women at the United States Naval Academy. Am J Sports Med 2000;28:98-102.

100 Agel J, Arendt EA, Bershadsky B. Anterior cruciate ligament injury in national collegiate athletic association basketball and soccer: a 13-year review. Am J Sports Med 2005;33:524-530.

101 Parkkari J, Pasanen K, Mattila VM, Kannus P, Rimpela A. The risk for a cruciate ligament injury of the knee in adolescents and young adults: a population-based cohort study of 46 500 people with a 9 year follow-up. Br J Sports Med 2008;42:422-426.

102 Chandrashekar N, Slauterbeck J, Hashemi J. Sex-based differences in the anthropometric characteristics of the anterior cruciate ligament and its relation to intercondylar notch geometry: a cadaveric study. Am J Sports Med 2005;33:1492-1498.

103 Chandrashekar N, Mansouri H, Slauterbeck J, Hashemi J. Sex-based differences in the tensile properties of the human anterior cruciate ligament. J Biomech 2006;39:2943-2950.

104 Hashemi J, Chandrashekar N, Mansouri H, Slauterbeck JR, Hardy DM. The human anterior cruciate ligament: sex differences in ultrastructure and correlation with biomechanical properties. J Orthop Res 2008;26:945-950.

105 Uhorchak JM, Scoville CR, Williams GN, Arciero RA, St Pierre P, Taylor DC. Risk factors associated with noncontact injury of the anterior cruciate ligament: a prospective four-year evaluation of 859 West Point cadets. Am J Sports Med 2003;31:831-842.

106 Teitz CC, Lind BK, Sacks BM. Symmetry of the femoral notch width index. Am J Sports Med 1997;25:687-690.

107 Anderson AF, Dome DC, Gautam S, Awh MH, Rennirt GW. Correlation of anthropometric measurements, strength, anterior cruciate ligament size, and intercondylar notch characteristics to sex differences in anterior cruciate ligament tear rates. Am J Sports Med 2001;29:58-66.

108 Horton MG, Hall TL. Quadriceps femoris muscle angle: normal values and relationships with gender and selected skeletal measures. Phys Ther 1989;69:897-901.

109 Souryal TO, Freeman TR. Intercondylar notch size and anterior cruciate ligament injuries in athletes. A prospective study. Am J Sports Med 1993;21:535-539.

110 LaPrade RF, Burnett QM,2nd. Femoral intercondylar notch stenosis and correlation to anterior cruciate ligament injuries. A prospective study. Am J Sports Med 1994;22:198-203.

111 Houseworth SW, Mauro VJ, Mellon BA, Kieffer DA. The intercondylar notch in acute tears of the anterior cruciate ligament: a computer graphics study. Am J Sports Med 1987;15:221-224.

112 Lund-Hanssen H, Gannon J, Engebretsen L, Holen KJ, Anda S, Vatten L. Intercondylar notch width and the risk for anterior cruciate ligament rupture. A case-control study in 46 female handball players. Acta Orthop Scand 1994;65:529-532.

113 Miljko M, Grle M, Kozul S, Kolobaric M, Djak I. Intercondylar notch width and inner angle of lateral femoral condyle as the risk factors for anterior cruciate ligament injury in female handball players in Herzegovina. Coll Antropol 2012;36:195-200.

114 Dejour H, Bonnin M. Tibial translation after anterior cruciate ligament rupture. Two radiological tests compared. J Bone Joint Surg Br 1994;76:745-749.

115 Bonci CM. Assessment and evaluation of predisposing factors to anterior cruciate ligament injury. J Athl Train 1999;34:155-164.

116 Coplan JA. Rotational motion of the knee: a comparison of normal and pronating subjects. J Orthop Sports Phys Ther 1989;10:366-369.

117 Loudon JK, Jenkins W, Loudon KL. The relationship between static posture and ACL injury in female athletes. J Orthop Sports Phys Ther 1996;24:91-97.

118 Jansson A, Saartok T, Werner S, Renstrom P. General joint laxity in 1845 Swedish school children of different ages: age- and gender-specific distributions. Acta Paediatr 2004;93:1202-1206.

119 Liu SH, al-Shaikh R, Panossian V, Yang RS, Nelson SD, Soleiman N, Finerman GA, Lane JM. Primary immunolocalization of estrogen and progesterone target cells in the human anterior cruciate ligament. J Orthop Res 1996;14:526-533.

120 Wojtys EM, Huston LJ, Lindenfeld TN, Hewett TE, Greenfield ML. Association between the menstrual cycle and anterior cruciate ligament injuries in female athletes. Am J Sports Med 1998;26:614-619.

121 Slauterbeck J, Clevenger C, Lundberg W, Burchfield DM. Estrogen level alters the failure load of the rabbit anterior cruciate ligament. J Orthop Res 1999;17:405-408.

122 Wojtys EM, Huston LJ, Boynton MD, Spindler KP, Lindenfeld TN. The effect of the menstrual cycle on anterior cruciate ligament injuries in women as determined by hormone levels. Am J Sports Med 2002;30:182-188.

123 Beynnon BD, Johnson RJ, Braun S, Sargent M, Bernstein IM, Skelly JM, Vacek PM. The relationship between menstrual cycle phase and anterior cruciate ligament injury: a case-control study of recreational alpine skiers. Am J Sports Med 2006;34:757-764.

124 Hewett TE, Zazulak BT, Myer GD. Effects of the menstrual cycle on anterior cruciate ligament injury risk: a systematic review. Am J Sports Med 2007;35:659-668.

125 More RC, Karras BT, Neiman R, Fritschy D, Woo SL, Daniel DM. Hamstrings--an anterior cruciate ligament protagonist. An in vitro study. Am J Sports Med 1993;21:231-237.

126 McLean SG, Huang X, Su A, Van Den Bogert AJ. Sagittal plane biomechanics cannot injure the ACL during sidestep cutting. Clin Biomech 2004;19:828-838.

127 Malinzak RA, Colby SM, Kirkendall DT, Yu B, Garrett WE. A comparison of knee joint motion patterns between men and women in selected athletic tasks. Clin Biomech 2001;16:438-445.

128 Ficek K, Cieszczyk P, Kaczmarczyk M, Maciejewska-Karlowska A, Sawczuk M, Cholewinski J, Leonska-Duniec A, Stepien-Slodkowska M, Zarebska A, Stepto NK, Bishop DJ, Eynon N. Gene variants within the COL1A1 gene are associated with reduced anterior cruciate ligament injury in professional soccer players. J Sci Med Sport 2013;16:396-400.

129 Posthumus M, September AV, Schwellnus MP, Collins M. Investigation of the Sp1-binding site polymorphism within the COL1A1 gene in participants with Achilles tendon injuries and controls. J Sci Med Sport 2009;12:184-189.

130 Collins M, Mokone GG, September AV, van der Merwe L, Schwellnus MP. The COL5A1 genotype is associated with range of motion measurements. Scand J Med Sci Sports 2009;19:803-810.

131 Gajko-Galicka A. Mutations in type I collagen genes resulting in osteogenesis imperfecta in humans. Acta Biochim Pol 2002;49:433-441.

132 Lian K, Zmuda JM, Nevitt MC, Lui L, Hochberg MC, Greene D, Li J, Wang J, Lane NE. Type I collagen alpha1 Sp1 transcription factor binding site polymorphism is associated with reduced risk of hip osteoarthritis defined by severe joint space narrowing in elderly women. Arthritis Rheum 2005;52:1431-1436.

133 Mann V, Hobson EE, Li B, Stewart TL, Grant SF, Robins SP, Aspden RM, Ralston SH. A COL1A1 Sp1 binding site polymorphism predisposes to osteoporotic fracture by affecting bone density and quality. J Clin Invest 2001;107:899-907.

134 Tilkeridis C, Bei T, Garantziotis S, Stratakis CA. Association of a COL1A1 polymorphism with lumbar disc disease in young military recruits. J Med Genet 2005;42.

135 Speer G, Szenthe P, Kosa JP, Tabak AG, Folhoffer A, Fuszek P, Cseh K, Lakatos P. Myocardial infarction is associated with Spl binding site polymorphism of collagen type 1A1 gene. Acta Cardiol 2006;61:321-325.

136 Skorupski P, Krol J, Starega J, Adamiak A, Jankiewicz K, Rechberger T. An alpha-1 chain of type I collagen Sp1-binding site polymorphism in women suffering from stress urinary incontinence. Am J Obstet Gynecol 2006;194:346-350.

137 Deak SB, van der Rest M, Prockop DJ. Altered helical structure of a homotrimer of alpha 1(I)chains synthesized by fibroblasts from a variant of osteogenesis imperfecta. Coll Relat Res 1985;5:305-313.

138 Niyibizi C, Kavalkovich K, Yamaji T, Woo SL. Type V collagen is increased during rabbit medial collateral ligament healing. Knee Surg Sports Traumatol Arthrosc 2000;8:281-285.

139 Laguette MJ, Abrahams Y, Prince S, Collins M. Sequence variants within the 3'-UTR of the COL5A1 gene alters mRNA stability: implications for musculoskeletal soft tissue injuries. Matrix Biol 2011;30:338-345.

140 Collins M, Posthumus M. Type V collagen genotype and exercise-related phenotype relationships: a novel hypothesis. Exerc Sport Sci Rev 2011;39:191-198.

141 Young BB, Zhang G, Koch M, Birk DE. The roles of types XII and XIV collagen in fibrillogenesis and matrix assembly in the developing cornea. J Cell Biochem 2002;87:208-220.

142 Raleigh SM, van der Merwe L, Ribbans WJ, Smith RK, Schwellnus MP, Collins M. Variants within the MMP3 gene are associated with Achilles tendinopathy: possible interaction with the COL5A1 gene. Br J Sports Med 2009;43:514-520.

143 Barlas IO, Sezgin M, Erdal ME, Sahin G, Ankarali HC, Altintas ZM, Turkmen E. Association of (-1,607) 1G/2G polymorphism of matrix metalloproteinase-1 gene with knee osteoarthritis in the Turkish population (knee osteoarthritis and MMPs gene polymorphisms). Rheumatol Int 2009;29:383-388.

144 Ye S, Patodi N, Walker-Bone K, Reading I, Cooper C, Dennison E. Variation in the matrix metalloproteinase-3, -7, -12 and -13 genes is associated with functional status in rheumatoid arthritis. Int J Immunogenet 2007;34:81-85.

145 Kalichman L, Hunter DJ. The genetics of intervertebral disc degeneration. Associated genes. Joint Bone Spine 2008;75:388-396.

146 Aulisa L, Papaleo P, Pola E, Angelini F, Aulisa AG, Tamburrelli FC, Pola P, Logroscino CA. Association between IL-6 and MMP-3 gene polymorphisms and adolescent idiopathic scoliosis: a case-control study. Spine 2007;32:2700-2702.

147 Malik MH, Jury F, Bayat A, Ollier WE, Kay PR. Genetic susceptibility to total hip arthroplasty failure: a preliminary study on the influence of matrix metalloproteinase 1, interleukin 6 polymorphisms and vitamin D receptor. Ann Rheum Dis 2007;66:1116-1120.

148 NCBI gene. MMP1 matrix metallopeptidase 1 (interstitial collagenase) [Homo sapians]. NCBI 2011;2011.

149 Rutter JL, Mitchell TI, Buttice G, Meyers J, Gusella JF, Ozelius LJ, Brinckerhoff CE. A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter creates an Ets binding site and augments transcription. Cancer Res 1998;58:5321-5325.

150 Ye S, Eriksson P, Hamsten A, Kurkinen M, Humphries SE, Henney AM. Progression of coronary atherosclerosis is associated with a common genetic variant of the human stromelysin-1 promoter which results in reduced gene expression. J Biol Chem 1996;271:13055-13060.

151 NCBI gene. MMP10 matrix metallopeptidase 10 (stromelysin 2) [*Homo sapiens*]. NCBI 2011;2011.

152 Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009;4:1073-1081.

153 NCBI gene. MMP12 matrix metallopeptidase 12 (macrophage elastase) [Homo sapians]. NCBI 2011;2011.

154 Jormsjo S, Ye S, Moritz J, Walter DH, Dimmeler S, Zeiher AM, Henney A, Hamsten A, Eriksson P. Allele-specific regulation of matrix metalloproteinase-12 gene activity is associated with coronary artery luminal dimensions in diabetic patients with manifest coronary artery disease. Circ Res 2000;86:998-1003.

155 Lee B, D'Alessio M, Ramirez F. Modifications in the organization and expression of collagen genes associated with skeletal disorders. Crit Rev Eukaryot Gene Expr 1991;1:173-187.

156 O'Connell K, Saunders C, Collins M. Collagen gene sequence variants in exercise-related traits. CEJSSM 2013;1:3-17.

157 Symoens S, Malfait F, Renard M, Andre J, Hausser I, Loeys B, Coucke P, De Paepe A. COL5A1 signal peptide mutations interfere with protein secretion and cause classic Ehlers-Danlos syndrome. Hum Mutat 2009;30:395-403.

158 Salmivirta M, Jalkanen M. Syndecan family of cell surface proteoglycans: developmentally regulated receptors for extracellular effector molecules. Experientia 1995;51:863-872.

159 Stattin E, Wiklund F, Lindblom K, Önnerfjord P, Jonsson B, Tegner Y, Sasaki T, Struglics A, Lohmander S, Dahl N, Heinegård D, Aspberg A. A Missense Mutation in the Aggrecan C-type Lectin Domain Disrupts Extracellular Matrix Interactions and Causes Dominant Familial Osteochondritis Dissecans. The American Journal of Human Genetics 2010;86:126-137.

160 Tompson SW, Merriman B, Funari VA, Fresquet M, Lachman RS, Rimoin DL, Nelson SF, Briggs MD, Cohn DH, Krakow D. A Recessive Skeletal Dysplasia, SEMD Aggrecan Type, Results from a Missense Mutation Affecting the C-Type Lectin Domain of Aggrecan. The American Journal of Human Genetics 2009;84:72-79.

161 Jackson RL, Busch SJ, Cardin AD. Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes. Physiol Rev 1991;71:481-539.

162 Kannus P. Structure of the tendon connective tissue. Scand J Med Sci Sports 2000;10:312-320.

163 Schaefer ,Liliana, Schaefer ,Roland. Proteoglycans: from structural compounds to signaling molecules. Cell Tissue Res 2010:237-246.

164 Haddock G, Cross AK, Allan S, Sharrack B, Callaghan J, Bunning RA, Buttle DJ, Woodroofe MN. Brevican and phosphacan expression and localization following transient middle cerebral artery occlusion in the rat. Biochem Soc Trans 2007;35:692-694.

165 Vogel KG, Heinegard D. Characterization of proteoglycans from adult bovine tendon. J Biol Chem 1985;260:9298-9306.

166 Somerman MJ, Young MF, Foster RA, Moehring JM, Imm G, Sauk JJ. Characteristics of human periodontal ligament cells in vitro. Arch Oral Biol 1990;35:241-247.

167 Yukawa M, Itabashi T, Takeuchi K, Narita H, Takeda Y, Okada A, Ueyama K, Harata S. The role of proteoglycans in ossification of spinal ligaments. Int Congr Ser 2001;1223:193-198.

168 Ilic MZ, Carter P, Tyndall A, Dudhia J, Handley CJ. Proteoglycans and catabolic products of proteoglycans present in ligament. Biochem J 2005;385:381-388.

169 Young K, Samiric T, Feller J, Cook J. Extracellular matrix content of ruptured anterior cruciate ligament tissue. Knee 2011;18:242-246.

170 Vogel KG, Ordog A, Pogany G, Olah J. Proteoglycans in the compressed region of human tibialis posterior tendon and in ligaments. J Orthop Res 1993;11:68-77.

171 Berenson MC, Blevins FT, Plaas AH, Vogel KG. Proteoglycans of human rotator cuff tendons. J Orthop Res 1996;14:518-525.

172 Waggett AD, Ralphs JR, Kwan APL, Woodnutt D, Benjamin M. Characterization of collagens and proteoglycans at the insertion of the human achilles tendon. Matrix Biology 1998;16:457-470.

173 Samiric T, Ilic MZ, Handley CJ. Characterisation of proteoglycans and their catabolic products in tendon and explant cultures of tendon. Matrix Biology 2004;23:127-140.

174 Larjava H, Hakkinen L, Rahemtulla F. A biochemical analysis of human periodontal tissue proteoglycans. Biochem J 1992;284:267-274.

175 Henninger HB, Maas SA, Underwood CJ, Whitaker RT, Weiss JA. Spatial distribution and orientation of dermatan sulfate in human medial collateral ligament. J Struct Biol 2007;158:33-45.

176 Yamada S, Murakami S, Matoba R, Ozawa Y, Yokokoji T, Nakahira Y, Ikezawa K, Takayama S, Matsubara K, Okada H. Expression profile of active genes in human periodontal ligament and isolation of PLAP-1, a novel SLRP family gene. Gene 2001;275:279-286.

177 Liu W, Burton-Wurster N, Glant TT, Tashman S, Sumner DR, Kamath RV, Lust G, Kimura JH, Cs-Szabo G. Spontaneous and experimental osteoarthritis in dog: similarities and differences in proteoglycan levels. Journal of Orthopaedic Research 2003;21:730-737.

178 Chatuparisute P, Shinohara Y, Kirchhoff C, Fischer F, Milz S. Immunohistochemical composition of the human lunotriquetral interosseous ligament. Appl Immunohistochem Mol Morphol 2012;20:318-324.

179 Rees SG, Waggett AD, Kerr BC, Probert J, Gealy EC, Dent CM, Caterson B, Hughes CE. Immunolocalisation and expression of keratocan in tendon. Osteoarthritis and Cartilage 2009;17:276-279.

180 Deere M, Johnson J, Garza S, Harrison WR, Yoon S, Elder FFB, Kucherlapati R, Hook M, Hecht JT. Characterization of Human DSPG3, a Small Dermatan Sulfate Proteoglycan. Genomics 1996;38:399-404.

181 Corps AN, Robinson AH, Movin T, Costa ML, Ireland DC, Hazleman BL, Riley GP. Versican splice variant messenger RNA expression in normal human Achilles tendon and tendinopathies. Rheumatology 2004;43:969-972.

182 Nelson F, Billinghurst RC, Pidoux I, Reiner A, Langworthy M, McDermott M, Malogne T, Sitler DF, Kilambi NR, Lenczner E, Poole AR. Early post-traumatic osteoarthritis-like changes in human articular cartilage following rupture of the anterior cruciate ligament. Osteoarthritis Cartilage 2006;14:114-119.

183 Vogel KG, Sandy JD, Pogány G, Robbins JR. Aggrecan in bovine tendon. Matrix Biology 1994;14:171-179.

184 Worapamorn W, Li H, Haas HR, Pujic Z, Girjes AA, Bartold PM. Cell surface proteoglycan expression by human periodontal cells. Connect Tissue Res 2000;41:57-68.

185 Sawaguchi N, Majima T, Iwasaki N, Funakoshi T, Shimode K, Onodera T, Minami A. Extracellular matrix modulates expression of cell-surface proteoglycan genes in fibroblasts. Connect Tissue Res 2006;47:141-148.

186 Elsaid KA, Fleming BC, Oksendahl HL, Machan JT, Fadale PD, Hulstyn MJ, Shalvoy R, Jay GD. Decreased lubricin concentrations and markers of joint inflammation in the synovial fluid of patients with anterior cruciate ligament injury. Arthritis Rheum 2008;58:1707-1715.

187 Marcelino J, Carpten JD, Suwairi WM, Gutierrez OM, Schwartz S, Robbins C, Sood R, Makalowska I, Baxevanis A, Johnstone B, Laxer RM, Zemel L, Kim CA, Herd JK, Ihle J, Williams C, Johnson M, Raman V, Alonso LG, Brunoni D, Gerstein A, Papadopoulos N, Bahabri SA, Trent JM, Warman ML. CACP, encoding a secreted proteoglycan, is mutated in camptodactyly-arthropathy-coxa vara-pericarditis syndrome. Nat Genet 1999;23:319-322.

188 Rees SG, Davies JR, Tudor D, Flannery CR, Hughes CE, Dent CM, Caterson B. Immunolocalisation and expression of proteoglycan 4 (cartilage superficial zone proteoglycan) in tendon. Matrix Biology 2002;21:593-602.

189 Iozzo RV. The family of the small leucine-rich proteoglycans: key regulators of matrix assembly and cellular growth. Crit Rev Biochem Mol Biol 1997;32:141-174.

190 McEwan PA, Scott PG, Bishop PN, Bella J. Structural correlations in the family of small leucine-rich repeat proteins and proteoglycans. J Struct Biol 2006;155:294-305.

191 Chen S, Birk DE. Focus on Molecules: Decorin. Exp Eye Res 2011;92:444-445.

192 Ansorge HL, Hsu JE, Edelstein L, Adams S, Birk DE, Soslowsky LJ. Recapitulation of the Achilles tendon mechanical properties during neonatal development: a study of differential healing during two stages of development in a mouse model. J Orthop Res 2012;30:448-456.

193 Schaefer L, Iozzo RV. Biological functions of the small leucine-rich proteoglycans: from genetics to signal transduction. J Biol Chem 2008;283:21305-21309.

194 Waddington RJ, Hall RC, Embery G, Lloyd DM. Changing profiles of proteoglycans in the transition of predentine to dentine. Matrix Biology 2003;22:153-161.

195 Chiu R, Li W, Herber RP, Marshall SJ, Young M, Ho SP. Effects of biglycan on physico-chemical properties of ligament-mineralized tissue attachment sites. Arch Oral Biol 2012;57:177-187.

196 Hey NJ, Handley CJ, Chee Keng Ng, Oakes BW. Characterization and synthesis of macromolecules by adult collateral ligament. BBA 1990;1034:73-80.

197 Reed CC, Iozzo RV. The role of decorin in collagen fibrillogenesis and skin homeostasis. Glycoconj J 2002;19:249-255.

198 Mochizuki Y, Shinomura T, Kimata K. Leucine-rich Proteoglycan. http://www.glycoforum gr jp/science/word/proteoglycan/PGA04E html March 1998;October 2013.

199 Jepsen KJ, Wu F, Peragallo JH, Paul J, Roberts L, Ezura Y, Oldberg A, Birk DE, Chakravarti S. A syndrome of joint laxity and impaired tendon integrity in lumican- and fibromodulin-deficient mice. J Biol Chem 2002;277:35532-35540.

200 Ezura Y, Chakravarti S, Oldberg A, Chervoneva I, Birk DE. Differential expression of lumican and fibromodulin regulate collagen fibrillogenesis in developing mouse tendons. J Cell Biol 2000;151:779-788.

201 Johnson J, Shinomura T, Eberspaecher H, Pinero G, Decrombrugghe B, Hook M. Expression and localization of PG-Lb/epiphycan during mouse development. Dev Dyn 1999;216:499-510.

202 Iozzo RV, Murdoch AD. Proteoglycans of the extracellular environment: clues from the gene and protein side offer novel perspectives in molecular diversity and function. FASEB J 1996;10:598-614.

203 Iozzo RV. Matrix proteoglycans: from molecular design to cellular function. Annu Rev Biochem 1998;67:609-652.

204 Campbell MA, Tester AM, Handley CJ, Checkley GJ, Chow GL, Cant AE, Winter AD, Cain WE. Characterization of a Large Chondroitin Sulfate Proteoglycan Present in Bovine Collateral Ligament. Arch Biochem Biophys 1996;329:181-190.

205 Zimmermann DR, Ruoslahti E. Multiple domains of the large fibroblast proteoglycan, versican. EMBO J 1989;8:2975-2981.

206 Schonherr E, Jarvelainen HT, Sandell LJ, Wight TN. Effects of platelet-derived growth factor and transforming growth factor-beta 1 on the synthesis of a large versican-like chondroitin sulfate proteoglycan by arterial smooth muscle cells. J Biol Chem 1991;266:17640-17647.

207 Wu YJ, La Pierre DP, Wu J, Yee AJ, Yang BB. The interaction of versican with its binding partners. Cell Res 2005;15:483-494.

208 Yada T. Aggrecan and Versican. http://www.glycoforum.gr.jp/science/word/proteoglycan/PGA03Ehtml March 1998;October 2013.

209 Roughley P, Martens D, Rantakokko J, Alini M, Mwale F, Antoniou J. The involvement of aggrecan polymorphism in degeneration of human intervertebral disc and articular cartilage. Eur Cell Mater 2006;11:1-7.

210 Rees SG, Waggett AD, Dent CM, Caterson B. Inhibition of aggrecan turnover in short-term explant cultures of bovine tendon. Matrix Biology 2007;26:280-290.

211 Li H, Schwartz NB, Vertel BM. cDNA cloning of chick cartilage chondroitin sulfate (aggrecan) core protein and identification of a stop codon in the aggrecan gene associated with the chondrodystrophy, nanomelia. J Biol Chem 1993;268:23504-23511.

212 Watanabe H, Kimata K, Line S, Strong D, Gao LY, Kozak CA, Yamada Y. Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in the aggrecan gene. Nat Genet 1994;7:154-157.

213 Gleghorn L, Ramesar R, Beighton P, Wallis G. A Mutation in the Variable Repeat Region of the Aggrecan Gene (AGC1) Causes a Form of Spondyloepiphyseal Dysplasia Associated with Severe, Premature Osteoarthritis. The American Journal of Human Genetics 2005;77:484-490.

214 Bernfield M, Gotte M, Park PW, Reizes O, Fitzgerald ML, Lincecum J, Zako M. Functions of cell surface heparan sulfate proteoglycans. Annu Rev Biochem 1999;68:729-777.

215 Couchman JR. Syndecans: proteoglycan regulators of cell-surface microdomains? Nat Rev Mol Cell Biol 2003;4:926-937.

216 Filmus J, Capurro M, Rast J. Glypicans. Genome Biol 2008;9:224-2008-9-5-224. Epub 2008.

217 Schumacher BL, Hughes CE, Kuettner KE, Caterson B, Aydelotte MB. Immunodetection and partial cDNA sequence of the proteoglycan, superficial zone protein, synthesized by cells lining synovial joints. J Orthop Res 1999;17:110-120.

218 Flannery CR, Hughes CE, Schumacher BL, Tudor D, Aydelotte MB, Kuettner KE, Caterson B. Articular Cartilage Superficial Zone Protein (SZP) Is Homologous to Megakaryocyte Stimulating Factor Precursor and Is a Multifunctional Proteoglycan with Potential Growth-Promoting, Cytoprotective, and Lubricating Properties in Cartilage Metabolism. Biochem Biophys Res Commun 1999;254:535-541.

219 Cordell HJ, Clayton DG. Genetic association studies. Lancet 2005;366:1121-1131.

220 Li G, Diogo D, Wu D, Spoonamore J, Dancik V, Franke L, Kurreeman F, Rossin EJ, Duclos G, Hartland C, Zhou X, Li K, Liu J, De Jager PL, Siminovitch KA, Zhernakova A, Raychaudhuri S, Bowes J, Eyre S, Padyukov L, Gregersen PK, Worthington J, Rheumatoid Arthritis Consortium International (RACI), Gupta N, Clemons PA, Stahl E, Tolliday N, Plenge RM. Human genetics in rheumatoid arthritis guides a high-throughput drug screen of the CD40 signaling pathway. PLoS Genet 2013.

221 Riley JH, Allan CJ, Lai E, Roses A. The use of single nucleotide polymorphisms in the isolation of common disease genes. Pharmacogenomics 2000;1:39-47.

222 Nell E. A pathway-based approach investigating DNA sequence variants to implicate the inflammatory pathway in the genetic predisposition to Achilles tendinopathy. 2011:1-204.

223 Burton PR, Tobin MD, Hopper JL. Key concepts in genetic epidemiology. Lancet 2005;366:941-951.

224 Hattersley AT, McCarthy MI. What makes a good genetic association study? Lancet 2005;366:1315-1323.

225 Pettersson FH, Anderson CA, Clarke GM, Barrett JC, Cardon LR, Morris AP, Zondervan KT. Marker selection for genetic case-control association studies. Nat Protoc 2009;4:743-752.

226 Lowe CE, Cooper JD, Chapman JM, Barratt BJ, Twells RC, Green EA, Savage DA, Guja C, Ionescu-Tirgoviste C, Tuomilehto-Wolf E, Tuomilehto J, Todd JA, Clayton DG. Cost-effective analysis of candidate genes using htSNPs: a staged approach. Genes Immun 2004;5:301-305.

227 Doege KJ, Sasaki M, Kimura T, Yamada Y. 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 1991;266:894-902.

228 Watanabe H, Yamada Y, Kimata K. Roles of aggrecan, a large chondroitin sulfate proteoglycan, in cartilage structure and function. J Biochem 1998;124:687-693.

229 Ensembl. Gene ACAN: ENSG00000157766. Ensembl September 2013;October 2013.

230 Sebastiani P, Solovieff N, Hartley SW, Milton JN, Riva A, Dworkis DA, Melista E, Klings ES, Garrett ME, Telen MJ, Ashley-Koch A, Baldwin CT, Steinberg MH. Genetic modifiers of the severity of sickle cell anemia identified through a genome-wide association study. Am J Hematol 2010;85:29-35.

231 Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, Freathy RM, Perry JR, Stevens S, Hall AS, Samani NJ, Shields B, Prokopenko I, Farrall M, Dominiczak A, Diabetes Genetics Initiative, Wellcome Trust Case Control Consortium, Johnson T, Bergmann S, Beckmann JS, Vollenweider P, Waterworth DM, Mooser V, Palmer CN, Morris AD, Ouwehand WH, Cambridge GEM Consortium, Zhao JH, Li S, Loos RJ, Barroso I, Deloukas P, Sandhu MS, Wheeler E, Soranzo N, Inouye M, Wareham NJ, Caulfield M, Munroe PB, Hattersley AT, McCarthy MI, Frayling TM. Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet 2008;40:575-583.

232 N'Diaye A, Chen GK, Palmer CD, Ge B, Tayo B, Mathias RA, Ding J, Nalls MA, Adeyemo A, Adoue V, Ambrosone CB, Atwood L, Bandera EV, Becker LC, Berndt SI, Bernstein L, Blot WJ, Boerwinkle E, Britton A, Casey G, Chanock SJ, Demerath E, Deming SL, Diver WR, Fox C, Harris TB, Hernandez DG, Hu JJ, Ingles SA, John EM, Johnson C, Keating B, Kittles RA, Kolonel LN, Kritchevsky SB, Le Marchand L, Lohman K, Liu J, Millikan RC, Murphy A, Musani S, Neslund-Dudas C, North KE, Nyante S, Ogunniyi A, Ostrander EA, Papanicolaou G, Patel S, Pettaway CA, Press MF, Redline S, Rodriguez-Gil JL, Rotimi C, Rybicki BA, Salako B, Schreiner PJ, Signorello LB, Singleton AB, Stanford JL, Stram AH, Stram DO, Strom SS, Suktitipat B, Thun MJ, Witte JS, Yanek LR, Ziegler RG, Zheng W, Zhu X, Zmuda JM, Zonderman AB, Evans MK, Liu Y, Becker DM, Cooper RS, Pastinen T, Henderson BE, Hirschhorn JN, Lettre G, Haiman CA. Identification, replication, and fine-mapping of Loci associated with adult height in individuals of african ancestry. PLoS Genet 2011.

233 Videman T, Saarela J, Kaprio J, Nakki A, Levalahti E, Gill K, Peltonen L, Battie MC. Associations of 25 structural, degradative, and inflammatory candidate genes with lumbar disc desiccation, bulging, and height narrowing. Arthritis Rheum 2009;60:470-481.

234 Yip SP, Leung KH, Ng PW, Fung WY, Sham PC, Yap MK. Evaluation of proteoglycan gene polymorphisms as risk factors in the genetic susceptibility to high myopia. Invest Ophthalmol Vis Sci 2011;52:6396-6403.

235 Fisher LW, Termine JD, Young MF. Deduced protein sequence of bone small proteoglycan I (biglycan) shows homology with proteoglycan II (decorin) and several nonconnective tissue proteins in a variety of species. J Biol Chem 1989;264:4571-4576.

236 Waddington RJ, Roberts HC, Sugars RV, Schonherr E. Differential roles for small leucine-rich proteoglycans in bone formation. Eur Cell Mater 2003;6:12-21.

237 Tenorio DM, Santos MF, Zorn TM. Distribution of biglycan and decorin in rat dental tissue. Braz J Med Biol Res 2003;36:1061-1065.

238 Dourte LM, Pathmanathan L, Mienaltowski MJ, Jawad AF, Birk DE, Soslowsky LJ. Mechanical, compositional, and structural properties of the mouse patellar tendon with changes in biglycan gene expression. J Orthop Res 2013;31:1430-1437.

239 Young MF, Bi Y, Ameye L, Chen XD. Biglycan knockout mice: new models for musculoskeletal diseases. Glycoconj J 2002;19:257-262.

240 Chen WB, Lenschow W, Tiede K, Fischer JW, Kalthoff H, Ungefroren H. Smad4/DPC4-dependent regulation of biglycan gene expression by transforming growth factor-beta in pancreatic tumor cells. J Biol Chem 2002;277:36118-36128.

241 Kolb M, Margetts PJ, Sime PJ, Gauldie J. Proteoglycans decorin and biglycan differentially modulate TGF-beta-mediated fibrotic responses in the lung. Am J Physiol Lung Cell Mol Physiol 2001;280:1327-34.

242 Frey H, Schroeder N, Manon-Jensen T, Iozzo RV, Schaefer L. Biological interplay between proteoglycans and their innate immune receptors in inflammation. FEBS J 2013;280:2165-2179.

243 Ensembl. Gene BGN: ENSG00000182492. Ensembl September 2013;October 2013.

244 Ungefroren H, Krull NB. Transcriptional regulation of the human biglycan gene. J Biol Chem 1996;271:15787-15795.

245 Ungefroren H, Cikós T, Krull NB, Kalthoff H. Biglycan Gene Promoter Activity in Osteosarcoma Cells Is Regulated by Cyclic AMP. Biochem Biophys Res Commun 1997;235:413-417.

246 Westergren-Thorsson G, Antonsson P, Malmström A, Heinegård D, Oldberg Å. The Synthesis of a Family of Structurally Related Proteoglycans in Fibroblasts is Differently Regulated by TGF-β. Matrix 1991;11:177-183.

247 Roughley PJ, Melching LI, Recklies AD. Changes in the expression of decorin and biglycan in human articular cartilage with age and regulation by TGF-β. Matrix Biology 1994;14:51-59.

248 Geerkens C, Vetter U, Just W, Fedarko NS, Fisher LW, Young MF, Termine JD, Robey PG, Wohrle D, Vogel W. The X-chromosomal human biglycan gene BGN is subject to X inactivation but is transcribed like an X-Y homologous gene. Hum Genet 1995;96:44-52.

249 Bhide VM, Laschinger CA, Arora PD, Lee W, Hakkinen L, Larjava H, Sodek J, McCulloch CA. Collagen phagocytosis by fibroblasts is regulated by decorin. J Biol Chem 2005;280:23103-23113.

250 Moscatello DK, Santra M, Mann DM, McQuillan DJ, Wong AJ, Iozzo RV. Decorin suppresses tumor cell growth by activating the epidermal growth factor receptor. J Clin Invest 1998;101:406-412.

251 Schonherr E, Sunderkotter C, Iozzo RV, Schaefer L. Decorin, a novel player in the insulin-like growth factor system. J Biol Chem 2005;280:15767-15772.

252 Scott JE. Proteoglycan-fibrillar collagen interactions. Biochem J 1988;252:313-323.

253 Pogany G, Vogel KG. The interaction of decorin core protein fragments with type I collagen. Biochem Biophys Res Commun 1992;189:165-172.

254 Vogel KG, Paulsson M, Heinegard D. Specific inhibition of type I and type II collagen fibrillogenesis by the small proteoglycan of tendon. Biochem J 1984;223:587-597.

255 Scott JE, Orford CR. Dermatan sulphate-rich proteoglycan associates with rat tail-tendon collagen at the d band in the gap region. Biochem J 1981;197:213-216.

256 Whinna HC, Choi HU, Rosenberg LC, Church FC. Interaction of heparin cofactor II with biglycan and decorin. J Biol Chem 1993;268:3920-3924.

257 Dourte LM, Pathmanathan L, Jawad AF, Iozzo RV, Mienaltowski MJ, Birk DE, Soslowsky LJ. Influence of decorin on the mechanical, compositional, and structural properties of the mouse patellar tendon. J Biomech Eng 2012;134.

258 Kresse H, Schonherr E. Proteoglycans of the extracellular matrix and growth control. J Cell Physiol 2001;189:266-274.

259 Markmann A, Hausser H, Schonherr E, Kresse H. Influence of decorin expression on transforming growth factor-beta-mediated collagen gel retraction and biglycan induction. Matrix Biol 2000;19:631-636.

260 Hildebrand A, Romaris M, Rasmussen LM, Heinegard D, Twardzik DR, Border WA, Ruoslahti E. Interaction of the small interstitial proteoglycans biglycan, decorin and fibromodulin with transforming growth factor beta. Biochem J 1994;302:527-534.

261 Iozzo RV, Moscatello DK, McQuillan DJ, Eichstetter I. Decorin is a biological ligand for the epidermal growth factor receptor. J Biol Chem 1999;274:4489-4492.

262 Jarvelainen H, Puolakkainen P, Pakkanen S, Brown EL, Hook M, Iozzo RV, Sage EH, Wight TN. A role for decorin in cutaneous wound healing and angiogenesis. Wound Repair Regen 2006;14:443-452.

263 Schonherr E, Levkau B, Schaefer L, Kresse H, Walsh K. Decorin-mediated signal transduction in endothelial cells. Involvement of Akt/protein kinase B in up-regulation of p21(WAF1/CIP1) but not p27(KIP1). J Biol Chem 2001;276:40687-40692.

264 Merline R, Moreth K, Beckmann J, Nastase MV, Zeng-Brouwers J, Tralhao JG, Lemarchand P, Pfeilschifter J, Schaefer RM, Iozzo RV, Schaefer L. Signaling by the matrix proteoglycan decorin controls inflammation and cancer through PDCD4 and MicroRNA-21. Sci Signal 2011;4:75.

265 Ensembl. Gene DCN: ENSG00000011465. Ensembl September 2013;October 2013.

266 Mauviel A, Korang K, Santra M, Tewari D, Uitto J, Iozzo RV. Identification of a bimodal regulatory element encompassing a canonical AP-1 binding site in the proximal promoter region of the human decorin gene. J Biol Chem 1996;271:24824-24829.

267 Heino J, Kahari VM, Mauviel A, Krusius T. Human recombinant interleukin-1 regulates cellular mRNA levels of dermatan sulphate proteoglycan core protein. Biochem J 1988;252:309-312.

268 Wegrowski Y, Paltot V, Gillery P, Kalis B, Randoux A, Maquart FX. Stimulation of sulphated glycosaminoglycan and decorin production in adult dermal fibroblasts by recombinant human interleukin-4. Biochem J 1995;307:673-678.

269 Iozzo RV, Cohen I. Altered proteoglycan gene expression and the tumor stroma. Experientia 1993;49:447-455.

270 van Diemen CC, Postma DS, Vonk JM, Bruinenberg M, Nolte IM, Boezen HM. Decorin and TGF-beta1 polymorphisms and development of COPD in a general population. Respir Res 2006;7:89.

271 Kelemen LE, Couch FJ, Ahmed S, Dunning AM, Pharoah PD, Easton DF, Fredericksen ZS, Vierkant RA, Pankratz VS, Goode EL, Scott CG, Rider DN, Wang X, Cerhan JR, Vachon CM. Genetic variation in stromal proteins decorin and lumican with breast cancer: investigations in two case-control studies. Breast Cancer Res 2008;10:98.

272 Dimasi DP, Burdon KP, Hewitt AW, Savarirayan R, Healey PR, Mitchell P, Mackey DA, Craig JE. Candidate gene study to investigate the genetic determinants of normal variation in central corneal thickness. Mol Vis 2010;16:562-569.

273 Ensembl. Gene FMOD: ENSG00000122176. Ensembl September 2013;October 2013.

274 Blochberger TC, Vergnes JP, Hempel J, Hassell JR. cDNA to chick lumican (corneal keratan sulfate proteoglycan) reveals homology to the small interstitial proteoglycan gene family and expression in muscle and intestine. J Biol Chem 1992;267:347-352.

275 Chakravarti S, Magnuson T, Lass JH, Jepsen KJ, LaMantia C, Carroll H. Lumican regulates collagen fibril assembly: skin fragility and corneal opacity in the absence of lumican. J Cell Biol 1998;141:1277-1286.

276 Funderburgh JL, Funderburgh ML, Mann MM, Conrad GW. Arterial lumican. Properties of a corneal-type keratan sulfate proteoglycan from bovine aorta. J Biol Chem 1991;266:24773-24777.

277 Sztrolovics R, Alini M, Mort JS, Roughley PJ. Age-related changes in fibromodulin and lumican in human intervertebral discs. Spine 1999;24:1765-1771.

278 Raouf A, Ganss B, McMahon C, Vary C, Roughley PJ, Seth A. Lumican is a major proteoglycan component of the bone matrix. Matrix Biol 2002;21:361-367.

279 Qin H, Ishiwata T, Wang R, Kudo M, Yokoyama M, Naito Z, Asano G. Effects of extracellular matrix on phenotype modulation and MAPK transduction of rat aortic smooth muscle cells in vitro. Exp Mol Pathol 2000;69:79-90.

280 Grover J, Chen XN, Korenberg JR, Roughley PJ. The human lumican gene. Organization, chromosomal location, and expression in articular cartilage. J Biol Chem 1995;270:21942-21949.

281 Nikitovic D, Berdiaki A, Zafiropoulos A, Katonis P, Tsatsakis A, Karamanos NK, Tzanakakis GN. Lumican expression is positively correlated with the differentiation and negatively with the growth of human osteosarcoma cells. FEBS J 2008;275:350-361.

282 D'Onofrio MF, Brezillon S, Baranek T, Perreau C, Roughley PJ, Maquart FX, Wegrowski Y. Identification of beta1 integrin as mediator of melanoma cell adhesion to lumican. Biochem Biophys Res Commun 2008;365:266-272.

283 Chakravarti S, Stallings RL, SundarRaj N, Cornuet PK, Hassell JR. Primary Structure of Human Lumican (Keratan Sulfate Proteoglycan) and Localization of the Gene (LUM) to Chromosome 12q21.3-q22. Genomics 1995;27:481-488.

284 Ensembl. Gene LUM: ESNG00000139329. Ensembl September 2013;October 2013.

285 Chen ZT, Wang I-, Shih Y, Lin LL. The Association of Haplotype at the Lumican Gene with High Myopia Susceptibility in Taiwanese Patients. Ophthalmology 2009;116:1920-1927.

286 Zhang F, Zhu T, Zhou Z, Wu Y, Li Y. Association of lumican gene with susceptibility to pathological myopia in the northern han ethnic chinese. J Ophthalmol 2009;2009:514306.

287 Couch FJ, Wang X, McWilliams RR, Bamlet WR, de Andrade M, Petersen GM. Association of breast cancer susceptibility variants with risk of pancreatic cancer. Cancer Epidemiol Biomarkers Prev 2009;18:3044-3048.

288 Toole BP. Solubility of collagen fibrils formed in vitro in the presence of sulphated acid mucopolysaccharide-protein. Nature 1969;222:872-873.

289 Connizzo BK, Yannascoli SM, Soslowsky LJ. Structure-function relationships of postnatal tendon development: a parallel to healing. Matrix Biol 2013;32:106-116.

290 Svensson L, Aszodi A, Reinholt FP, Fassler R, Heinegard D, Oldberg A. Fibromodulin-null mice have abnormal collagen fibrils, tissue organization, and altered lumican deposition in tendon. J Biol Chem 1999;274:9636-9647.

291 Hocking AM, Shinomura T, McQuillan DJ. Leucine-rich repeat glycoproteins of the extracellular matrix. Matrix Biology 1998;17:1-19.

292 Vogel KG, Trotter JA. The effect of proteoglycans on the morphology of collagen fibrils formed in vitro. Coll Relat Res 1987;7:105-114.

293 Xu T, Bianco P, Fisher LW, Longenecker G, Smith E, Goldstein S, Bonadio J, Boskey A, Heegaard AM, Sommer B, Satomura K, Dominguez P, Zhao C, Kulkarni AB, Robey PG, Young MF. Targeted disruption of the biglycan gene leads to an osteoporosis-like phenotype in mice. Nat Genet 1998;20:78-82.

294 Ferdous Z, Wei VM, Iozzo R, Hook M, Grande-Allen KJ. Decorin-transforming growth factor-interaction regulates matrix organization and mechanical characteristics of three-dimensional collagen matrices. J Biol Chem 2007;282:35887-35898.

295 Hakkinen L, Strassburger S, Kahari VM, Scott PG, Eichstetter I, Lozzo RV, Larjava H. A role for decorin in the structural organization of periodontal ligament. Lab Invest 2000;80:1869-1880.

296 Droguett R, Cabello-Verrugio C, Riquelme C, Brandan E. Extracellular proteoglycans modify TGF-beta bio-availability attenuating its signaling during skeletal muscle differentiation. Matrix Biol 2006;25:332-341.

297 Dunkman AA, Buckley MR, Mienaltowski MJ, Adams SM, Thomas SJ, Satchell L, Kumar A, Pathmanathan L, Beason DP, Iozzo RV, Birk DE, Soslowsky LJ. The Tendon Injury Response is Influenced by Decorin and Biglycan. Ann Biomed Eng 2013.

298 Grahame R. Joint hypermobility and genetic collagen disorders: are they related? Arch Dis Child 1999;80:188-191.

299 Lee JH, Ki CS, Chung ES, Chung TY. A novel decorin gene mutation in congenital hereditary stromal dystrophy: a Korean family. Korean J Ophthalmol 2012;26:301-305.

300 Bredrup C, Knappskog PM, Majewski J, Rodahl E, Boman H. Congenital stromal dystrophy of the cornea caused by a mutation in the decorin gene. Invest Ophthalmol Vis Sci 2005;46:420-426.

301 Majava M, Bishop PN, Hagg P, Scott PG, Rice A, Inglehearn C, Hammond CJ, Spector TD, Ala-Kokko L, Mannikko M. Novel mutations in the small leucine-rich repeat protein/proteoglycan (SLRP) genes in high myopia. Hum Mutat 2007;28:336-344.

302 Liu CY, Birk DE, Hassell JR, Kane B, Kao WW. Keratocan-deficient mice display alterations in corneal structure. J Biol Chem 2003;278:21672-21677.

303 Meij JT, Carlson EC, Wang L, Liu CY, Jester JV, Birk DE, Kao WW. Targeted expression of a lumican transgene rescues corneal deficiencies in lumican-null mice. Mol Vis 2007;13:2012-2018.

304 Bost F, Diarra-Mehrpour M, Martin JP. Inter-alpha-trypsin inhibitor proteoglycan family--a group of proteins binding and stabilizing the extracellular matrix. Eur J Biochem 1998;252:339-346.

305 Parkinson J, Samiric T, Ilic MZ, Cook J, Handley CJ. Involvement of proteoglycans in tendinopathy. J Musculoskelet Neuronal Interact 2011;11:86-93.

306 Robbins JR, Evanko SP, Vogel KG. Mechanical Loading and TGF-β Regulate Proteoglycan Synthesis in Tendon. Arch Biochem Biophys 1997;342:203-211.

307 Attia M, Scott A, Duchesnay A, Carpentier G, Soslowsky LJ, Huynh MB, Van Kuppevelt TH, Gossard C, Courty J, Tassoni MC, Martelly I. Alterations of overused supraspinatus tendon: a possible

role of glycosaminoglycans and HARP/pleiotrophin in early tendon pathology. J Orthop Res 2012;30:61-71.

308 Attia M, Scott A, Carpentier G, Lian O, Van Kuppevelt T, Gossard C, Papy-Garcia D, Tassoni MC, Martelly I. Greater glycosaminoglycan content in human patellar tendon biopsies is associated with more pain and a lower VISA score. Br J Sports Med 2013.

309 Parkinson J, Samiric T, Ilic MZ, Cook J, Feller JA, Handley CJ. Change in proteoglycan metabolism is a characteristic of human patellar tendinopathy. Arthritis Rheum 2010;62:3028-3035.

310 Turunen SM, Han SK, Herzog W, Korhonen RK. Cell deformation behavior in mechanically loaded rabbit articular cartilage 4 weeks after anterior cruciate ligament transection. Osteoarthritis Cartilage 2013;21:505-513.

311 Matyas JR, Atley L, Ionescu M, Eyre DR, Poole AR. Analysis of cartilage biomarkers in the early phases of canine experimental osteoarthritis. Arthritis Rheum 2004;50:543-552.

312 Lohmander LS. Markers of cartilage metabolism in arthrosis. A review. Acta Orthop Scand 1991;62:623-632.

313 Lohmander LS, Ionescu M, Jugessur H, Poole AR. Changes in joint cartilage aggrecan after knee injury and in osteoarthritis. Arthritis Rheum 1999;42:534-544.

314 Westergren-Thorsson G, Hernnas J, Sarnstrand B, Oldberg A, Heinegard D, Malmstrom A. Altered expression of small proteoglycans, collagen, and transforming growth factor-beta 1 in developing bleomycin-induced pulmonary fibrosis in rats. J Clin Invest 1993;92:632-637.

315 Hogemann B, Edel G, Schwarz K, Krech R, Kresse H. Expression of biglycan, decorin and proteoglycan-100/CSF-1 in normal and fibrotic human liver. Pathol Res Pract 1997;193:747-751.

316 Davies M. Glomerular cell proteoglycans: their possible role in progressive glomerular disease. Exp Nephrol 1995;3:122-126.

317 Scott PG, Dodd CM, Tredget EE, Ghahary A, Rahemtulla F. Chemical characterization and quantification of proteoglycans in human post-burn hypertrophic and mature scars. Clin Sci 1996;90:417-425.

318 Kinsella MG, Tsoi CK, Jarvelainen HT, Wight TN. Selective expression and processing of biglycan during migration of bovine aortic endothelial cells. The role of endogenous basic fibroblast growth factor. J Biol Chem 1997;272:318-325.

319 Lo IK, Marchuk LL, Hart DA, Frank CB. Comparison of mRNA levels for matrix molecules in normal and disrupted human anterior cruciate ligaments using reverse transcription-polymerase chain reaction. J Orthop Res 1998;16:421-428.

320 Lo IK, Marchuk L, Hart DA, Frank CB. Messenger ribonucleic acid levels in disrupted human anterior cruciate ligaments. Clin Orthop Relat Res 2003;407:249-258.

321 Provenzano PP, Alejandro-Osorio AL, Valhmu WB, Jensen KT, Vanderby R,Jr. Intrinsic fibroblast-mediated remodeling of damaged collagenous matrices in vivo. Matrix Biol 2005;23:543-555.

322 Buraschi S, Neill T, Owens RT, Iniguez LA, Purkins G, Vadigepalli R, Evans B, Schaefer L, Peiper SC, Wang ZX, Iozzo RV. Decorin protein core affects the global gene expression profile of the tumor microenvironment in a triple-negative orthotopic breast carcinoma xenograft model. PLoS One 2012;7.

323 Little J, Higgins JPT, Ioannidis JPA, Moher D, Gagnon F, von Elm E, Khoury MJ, Cohen B, Davey-Smith G, Grimshaw J, Scheet P, Gwinn M, Williamson RE, Zou GY, Hutchings K, Johnson CY, Tait V, Wiens M, Golding J, van Duijn C, McLaughlin J, Paterson A, Wells G, Fortier I, Freedman M, Zecevic M, King R, Infante-Rivard C, Stewart AF, Birkett N. Strengthening the reporting of genetic association studies (STREGA)—an extension of the strengthening the reporting of observational studies in epidemiology (STROBE) statement. J Clin Epidemiol 2009;62:597-608.

324 Marshall S, Padua D, McGrath M. Incidence of ACL injuries. In: Hewett T, Schultz S, Griffin L, eds. Understanding and preventing noncontact ACL injuries. Champaign, IL: Human Kinetics 2007:5-30.

325 Lahiri DK, Nurnberger JI,Jr. A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. Nucleic Acids Res 1991;19:5444.

326 Mokone GG, Gajjar M, September AV, Schwellnus MP, Greenberg J, Noakes TD, Collins M. The guanine-thymine dinucleotide repeat polymorphism within the tenascin-C gene is associated with achilles tendon injuries. Am J Sports Med 2005;33:1016-1021.

327 Peat RA, Gécz J, Fallon JR, Tarpey PS, Smith R, Futreal A, Stratton MR, Lamandé SR, Yang N, North KN. Exclusion of biglycan mutations in a cohort of patients with neuromuscular disorders. Neuromuscular Disorders 2008;18:606-609.

328 Gaunt T, Rodríguez S, Day I. Cubic exact solutions for the estimation of pairwise haplotype frequencies: implications for linkage disequilibrium analyses and a web tool 'CubeX'. BMC Bioinformatics 2007;8:428.

329 Heinegard D, Hascall VC. Aggregation of cartilage proteoglycans. 3. Characteristics of the proteins isolated from trypsin digests of aggregates. J Biol Chem 1974;249:4250-4256.

330 Kiani C, Chen L, Lee V, Zheng PS, Wu Y, Wen J, Cao L, Adams ME, Sheng W, Yang BB. Identification of the motifs and amino acids in aggrecan G1 and G2 domains involved in product secretion. Biochemistry 2003;42:7226-7237.

331 Aspberg A, Adam S, Kostka G, Timpl R, Heinegard D. Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and versican. J Biol Chem 1999;274:20444-20449.

332 Rodahl E, Van Ginderdeuren R, Knappskog PM, Bredrup C, Boman H. A second decorin frame shift mutation in a family with congenital stromal corneal dystrophy. Am J Ophthalmol 2006;142:520-521.

333 Salgado RM, Favaro RR, Martin SS, Zorn TM. The estrous cycle modulates small leucine-rich proteoglycans expression in mouse uterine tissues. Anat Rec 2009;292:138-153.

334 Hong SH, Nah HY, Lee JY, Gye MC, Kim CH, Kim MK. Analysis of estrogen-regulated genes in mouse uterus using cDNA microarray and laser capture microdissection. J Endocrinol 2004;181:157-167.

335 Rodrigo MC, Martin DS, Eyster KM. Estrogen decreases biglycan mRNA expression in resistance blood vessels. Am J Physiol Regul Integr Comp Physiol 2003;285:754-61.

336 Salgado RM, Favaro RR, Zorn TM. Modulation of small leucine-rich proteoglycans (SLRPs) expression in the mouse uterus by estradiol and progesterone. Reprod Biol Endocrinol 2011;9:22-7827-9-22.

337 Markiewicz M, Znoyko S, Stawski L, Ghatnekar A, Gilkeson G, Trojanowska M. A role for estrogen receptor-alpha and estrogen receptor-beta in collagen biosynthesis in mouse skin. J Invest Dermatol 2013;133:120-127.

338 Hay M, Patricios J, Collins R, Branfield A, Cook J, Handley CJ, September AV, Posthumus M, Collins M. Association of type XI collagen genes with chronic Achilles tendinopathy in independent populations from South Africa and Australia. Br J Sports Med 2013;47:569-574.

339 September AV, Nell EM, O'Connell K, Cook J, Handley CJ, van der Merwe L, Schwellnus M, Collins M. A pathway-based approach investigating the genes encoding interleukin-1{beta}, interleukin-6 and the interleukin-1 receptor antagonist provides new insight into the genetic susceptibility of Achilles tendinopathy. Br J Sports Med 2011.

340 Janssen KW, Orchard JW, Driscoll TR, van Mechelen W. High incidence and costs for anterior cruciate ligament reconstructions performed in Australia from 2003-2004 to 2007-2008: time for an anterior cruciate ligament register by Scandinavian model? Scand J Med Sci Sports 2012;22:495-501.

341 Dallinga JM, Benjaminse A, Lemmink KA. Which screening tools can predict injury to the lower extremities in team sports?: a systematic review. Sports Med 2012;42:791-815.

342 September AV, Posthumus M, Collins M. Application of genomics in the prevention, treatment and management of achilles tendinopathy and anterior cruciate ligament ruptures. Recent Pat DNA Gene Seq 2012;6:216-223.

343 Drenos F, Whittaker JC, Humphries SE. The use of meta-analysis risk estimates for candidate genes in combination to predict coronary heart disease risk. Ann Hum Genet 2007;71:611-619.

344 Islam M, Gor J, Perkins SJ, Ishikawa Y, Bachinger HP, Hohenester E. The concave face of decorin mediates reversible dimerisation and collagen binding. J Biol Chem 2013.

345 Chanut-Delalande H, Fichard A, Bernocco S, Garrone R, Hulmes DJ, Ruggiero F. Control of heterotypic fibril formation by collagen V is determined by chain stoichiometry. J Biol Chem 2001;276:24352-24359.

346 NCI-NHGRI Working Group on Replication in Association Studies, Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, Hirschhorn JN, Abecasis G, Altshuler D, Bailey-Wilson JE, Brooks LD, Cardon LR, Daly M, Donnelly P, Fraumeni JF,Jr, Freimer NB, Gerhard DS, Gunter C, Guttmacher AE, Guyer MS, Harris EL, Hoh J, Hoover R, Kong CA, Merikangas KR, Morton CC, Palmer LJ, Phimister EG, Rice JP, Roberts J, Rotimi C, Tucker MA, Vogan KJ, Wacholder S, Wijsman EM, Winn DM, Collins FS. Replicating genotype-phenotype associations. Nature 2007;447:655-660.

347 Tishkoff SA, Williams SM. Genetic analysis of African populations: human evolution and complex disease. Nat Rev Genet 2002;3:611-621.

348 Ackerman CM, Lowe LP, Lee H, Hayes MG, Dyer AR, Metzger BE, Lowe WL, Urbanek M, Hapo Study Cooperative Research Group. Ethnic variation in allele distribution of the androgen receptor (AR) (CAG)n repeat. J Androl 2012;33:210-215.

349 Cortes A, Field J, Glazov EA, Hadler J, ANZgene Consortium, Stankovich J, Brown MA. Resequencing and fine-mapping of the chromosome 12q13-14 locus associated with multiple sclerosis refines the number of implicated genes. Hum Mol Genet 2013;22:2283-2292.

350 Blobe GC, Schiemann WP, Lodish HF. Role of transforming growth factor beta in human disease. N Engl J Med 2000;342:1350-1358.

351 Klein MB, Yalamanchi N, Pham H, Longaker MT, Chang J. Flexor tendon healing in vitro: effects of TGF-beta on tendon cell collagen production. J Hand Surg Am 2002;27:615-620.

352 Wang X, Smith P, Pu LL, Kim YJ, Ko F, Robson MC. Exogenous transforming growth factor beta(2) modulates collagen I and collagen III synthesis in proliferative scar xenografts in nude rats. J Surg Res 1999;87:194-200.

353 Fu SC, Wong YP, Cheuk YC, Lee KM, Chan KM. TGF-beta1 reverses the effects of matrix anchorage on the gene expression of decorin and procollagen type I in tendon fibroblasts. Clin Orthop Relat Res 2005;431:226-232.

354 Marti HP, Lee L, Kashgarian M, Lovett DH. Transforming growth factor-beta 1 stimulates glomerular mesangial cell synthesis of the 72-kd type IV collagenase. Am J Pathol 1994;144:82-94.

355 Imai K, Hiramatsu A, Fukushima D, Pierschbacher MD, Okada Y. Degradation of decorin by matrix metalloproteinases: identification of the cleavage sites, kinetic analyses and transforming growth factor-beta1 release. Biochem J 1997;322:809-814.

356 Erridge C. Endogenous ligands of TLR2 and TLR4: agonists or assistants? J Leukoc Biol 2010;87:989-999.

357 Chen GY, Nunez G. Sterile inflammation: sensing and reacting to damage. Nat Rev Immunol 2010;10:826-837.

358 Schaefer L, Babelova A, Kiss E, Hausser HJ, Baliova M, Krzyzankova M, Marsche G, Young MF, Mihalik D, Gotte M, Malle E, Schaefer RM, Grone HJ. The matrix component biglycan is proinflammatory and signals through Toll-like receptors 4 and 2 in macrophages. J Clin Invest 2005;115:2223-2233.

359 Kim S, Takahashi H, Lin WW, Descargues P, Grivennikov S, Kim Y, Luo JL, Karin M. Carcinoma-produced factors activate myeloid cells through TLR2 to stimulate metastasis. Nature 2009;457:102-106.

360 Noble PW, McKee CM, Cowman M, Shin HS. Hyaluronan fragments activate an NF-kappa B/I-kappa B alpha autoregulatory loop in murine macrophages. J Exp Med 1996;183:2373-2378.

361 Moreth K, Iozzo RV, Schaefer L. Small leucine-rich proteoglycans orchestrate receptor crosstalk during inflammation. Cell Cycle 2012;11:2084-2091.

362 Carlson EC, Lin M, Liu CY, Kao WW, Perez VL, Pearlman E. Keratocan and lumican regulate neutrophil infiltration and corneal clarity in lipopolysaccharide-induced keratitis by direct interaction with CXCL1. J Biol Chem 2007;282:35502-35509.

363 Wu F, Vij N, Roberts L, Lopez-Briones S, Joyce S, Chakravarti S. A novel role of the lumican core protein in bacterial lipopolysaccharide-induced innate immune response. J Biol Chem 2007;282:26409-26417.

364 Wang VM, Bell RM, Thakore R, Eyre DR, Galante JO, Li J, Sandy JD, Plaas A. Murine tendon function is adversely affected by aggrecan accumulation due to the knockout of ADAMTS5. J Orthop Res 2012;30:620-626.

365 El Khoury L, Posthumus M, Collins M, Handley CJ, Cook J, Raleigh SM. Polymorphic variation within the ADAMTS2, ADAMTS14, ADAMTS5, ADAM12 and TIMP2 genes and the risk of Achilles tendon pathology: A genetic association study. J Sci Med Sport 2013;16:493-498.

366 Sarrazin S, Adam E, Lyon M, Depontieu F, Motte V, Landolfi C, Lortat-Jacob H, Bechard D, Lassalle P, Delehedde M. Endocan or endothelial cell specific molecule-1 (ESM-1): A potential novel endothelial cell marker and a new target for cancer therapy. BBA 2006;1765:25-37.

367 Rivera Z, Ferrone S, Wang X, Jube S, Yang H, Pass HI, Kanodia S, Gaudino G, Carbone M. CSPG4 as a target of antibody-based immunotherapy for malignant mesothelioma. Clin Cancer Res 2012;18:5352-5363.

368 Wang J, Svendsen A, Kmiecik J, Immervoll H, Skaftnesmo KO, Planaguma J, Reed RK, Bjerkvig R, Miletic H, Enger PO, Rygh CB, Chekenya M. Targeting the NG2/CSPG4 proteoglycan retards tumour growth and angiogenesis in preclinical models of GBM and melanoma. PLoS One 2011;6.

369 Samiric T, Ilic MZ, Handley CJ. Sulfated polysaccharides inhibit the catabolism and loss of both large and small proteoglycans in explant cultures of tendon. FEBS J 2006;273:3479-3488.

370 Yoshida H, Kojima T, Kurokouchi K, Takahashi S, Hanamura H, Kojima M, Poole AR, Ishiguro N. Relationship between pre-radiographic cartilage damage following anterior cruciate ligament injury and biomarkers of cartilage turnover in clinical practice: a cross-sectional observational study. Osteoarthritis Cartilage 2013;21:831-838.

371 Ruan MZ, Erez A, Guse K, Dawson B, Bertin T, Chen Y, Jiang MM, Yustein J, Gannon F, Lee BH. Proteoglycan 4 expression protects against the development of osteoarthritis. Sci Transl Med 2013;5.

372 Xie X, Wu H, Zhao S, Xie G, Huangfu X, Zhao J. The effect of platelet-rich plasma on patterns of gene expression in a dog model of anterior cruciate ligament reconstruction. J Surg Res 2013;180:80-88.

373 Nixon AJ, Goodrich LR, Scimeca MS, Witte TH, Schnabel LV, Watts AE, Robbins PD. Gene therapy in musculoskeletal repair. Ann N Y Acad Sci 2007;1117:310-327.

374 Kim SH, Lechman ER, Kim S, Nash J, Oligino TJ, Robbins PD. Ex vivo gene delivery of IL-1Ra and soluble TNF receptor confers a distal synergistic therapeutic effect in antigen-induced arthritis. Mol Ther 2002;6:591-600.

375 Bresnihan B, Alvaro-Gracia JM, Cobby M, Doherty M, Domljan Z, Emery P, Nuki G, Pavelka K, Rau R, Rozman B, Watt I, Williams B, Aitchison R, McCabe D, Musikic P. Treatment of rheumatoid arthritis with recombinant human interleukin-1 receptor antagonist. Arthritis Rheum 1998;41:2196-2204.

376 Oligino TJ, Yao Q, Ghivizzani SC, Robbins P. Vector systems for gene transfer to joints. Clin Orthop Relat Res 2000;379:17-30.

Appendices

Appendix A

Study information

THE GENETIC BASIS OF EXERCISE-INDUCED LIGAMENT INJURY

Although there is a high incidence of ligament injuries as a result of participation in exercise and sporting activities, the cause(s) of these injuries are poorly understood. Some researchers have suggested that there is a genetic component to exercise-induced ligament injuries. In an attempt to determine whether there is a genetic basis for ligament pathology, we are interested in studying whether certain genes are associated with ligament injuries. This project is being done in collaboration with the UCT/MRC Research Unit for Exercise Science and Sports Medicine within the Department of Human Biology of the University of Cape Town.

You will be required to visit the Sports Science Institute of South Africa (SSISA) in Boundary Road, Newlands. During the visit, which should take 15 minutes, you will be asked to donate 5ml (1 teaspoon) of a blood sample for DNA analysis. You will also be required to complete personal particulars, sporting details, medical history and stretching and warm up questionnaires.

All the information retrieved from this study will be treated with the strictest confidentiality and will be used only for scientific research purposes. Your name and personal particulars will not be released under any circumstances and all data will be analysed anonymously. Your DNA sample will be destroyed on completion of the study on the genetic basis of ligament injury. You are also free to request that your DNA sample be destroyed before the completion of the study.

If you are part of the ligament pathology group, we would appreciate it if you could help us by recruiting two other people of same (or similar) age whom you know and who has trained without suffering any ligament or tendon injuries for the control group.

We will keep you informed about the outcomes of this study and look forward to working together with you. If you have any questions about this study, please feel free to contact us at:-

Dr Mike Posthumus, PhD
(021) 650 4572
michael.posthumus@uct.ac.za

Dr Alison September, PhD
(021) 650 4559
alison.september@uct.ac.za

Prof Malcolm Collins, PhD
(021) 650 4574
malcolm.collins@uct.ac.za

Informed Consent for adults >18 years

GENETIC BASIS OF EXERCISE-INDUCED LIGAMENT INUJURY

INFORMED CONSENT

I, the undersigned, have been fully informed about the UCT/MRC Research Unit for Exercise Science and Sports Medicine within the Department of Human Biology of the University of Cape Town's study on the genetic basis of exercise induced chronic ligament pathology. I have agreed to donate five millilitres of venous blood or a buccal mouthwash sample, which will be used for the extraction and analysis of genetic material (DNA). I have also agreed to complete personal particulars, sporting participation, medical history, stretching and warm up questionnaires and understand that all the information that is collected during the study will be treated with the strictest confidentiality and will only be used for scientific research purposes. I also understand that my name and personal particulars will be not released under any circumstances and that all data will be analysed anonymously.

I agree to participate in the study and I have been informed that I will be free to withdraw from the study at any time if I so wish. I understand that my DNA sample will be destroyed on completion of the study on the genetic basis of ligament pathology. I also understand that I will be free to request that my DNA sample be destroyed before the completion of the study.

I understand that the DNA will be genotyped (analysed) for variations (polymorphisms) within genes relating to the genetic basis of ligament injuries. I understand that whilst there is no direct benefit to myself, if a genetic predisposition for ligament injuries can be established, then future generations will be able to establish their risk for this condition. This may allow better prevention and treatment options in the future. I understand that I will receive the overall results of the study. I have read (or where appropriate, have had read to me) and understand the information about this study, and any questions I have asked have been answered to my satisfaction. I agree to participate in the study, realising that I have the right to request that my DNA sample be destroyed at any time. I agree that research data provided by me or with my permission during the project may be included in a thesis, presented at conferences and published in journals on the condition that either my name not any other identifying information is used.

FULL NAME OF SUBJECT: _____

SUBJECT'S SIGNATURE: _____

DATE: _____

INVESTIGATOR : _____

INVESTIGATOR'S SIGNATURE: _____

Informed assent and consent for minors <18 years

GENETIC BASIS OF EXERCISE-INDUCED LIGAMENT INJURY

INFORMED ASSENT (for minors <18 years)

I, the undersigned, have been fully informed about the UCT/MRC Research Unit for Exercise Science and Sports Medicine within the Department of Human Biology of the University of Cape Town's study on the genetic basis of exercise-induced chronic ligament pathology. I have agreed to donate five millilitres of venous blood or a buccal mouthwash sample, which will be used for the extraction and analysis of genetic material (DNA). I have also agreed to complete personal particulars, sporting participation, medical history, stretching and warm up questionnaires and understand that all the information that is collected during the study will be treated with the strictest confidentiality and will only be used for scientific research purposes. I also understand that my name and personal particulars will be not released under any circumstances and that all data will be analysed anonymously.

I agree to participate in the study and I have been informed that I will be free to withdraw from the study at any time if I so wish. I understand that my DNA sample will be destroyed on completion of the study on the genetic basis for ligament pathology. I also understand that I will be free to request that my DNA sample be destroyed before the completion of the study.

I understand that the DNA will be genotyped (analysed) for variations (polymorphisms) within genes relating to the genetic basis of ligament injuries. I understand that whilst there is no direct benefit to myself, if a genetic predisposition for ligament injuries can be established, future generations will be able to establish their risk for this condition. This may allow better prevention and treatment options in the future. I understand that I will receive the overall results of the study. I have read (or where appropriate, have had read to me) and understand the information about this study, and any questions I have asked have been answered to my satisfaction. I agree to participate in the study, realising that I have the right to request that my DNA sample be destroyed at any time. I agree that research data provided by me or with my permission during the project may be included in a thesis, presented at conferences and published in journals on the condition that either my name not any other identifying information is used.

FULL NAME OF PARTICIPANT: _____

PARTICIPANT'S SIGNATURE: _____

DATE: _____

INVESTIGATOR: _____

INVESTIGATOR'S SIGNATURE: _____

PARENTAL/GUARDIAN CONSENT (for minors <18 years)

I, the undersigned and parent/ legal guardian of _____, have been fully informed about the UCT/MRC Research Unit for Exercise Science and Sports Medicine within the Department of Human Biology of the University of Cape Town's study on the genetic basis of exercise-induced chronic ligament pathology. I, the parent/ legal guardian, hereby give written permission that my child may participate in this study. I give permission for my child to donate five millilitres of venous blood or a buccal mouthwash sample, which will be used for the extraction and analysis of genetic material (DNA). In addition, I give permission that my child may complete personal particulars, sporting participation, medical history, stretching and warm up questionnaires and understand that all the information that is collected during the study will be treated with the strictest confidentiality and will only be used for scientific research purposes. I also understand that all our names and personal particulars will be not released under any circumstances and that all data will be analysed anonymously.

I agree that I have been informed that my child will be free to withdraw from the study at any time if I so wish. I understand that the DNA sample will be destroyed on completion of the study on the genetic basis for ligament pathology. I also understand that I will be free to request that the DNA sample be destroyed before the completion of the study.

I understand that the DNA will be genotyped (analysed) for variations (polymorphisms) within genes relating to the genetic basis of ligament injuries. I understand that whilst there is no direct benefit to myself, if the susceptibility factors for ligament injuries can be established, future generations will be able to establish their risk for this condition. This may allow better prevention and treatment options in the future. I understand that I will receive the overall results of the study. I have read (or where appropriate, have had read to me) and understand the information about this study, and any questions I have asked have been answered to my satisfaction. I agree for my child to participate in the study, realising that I have the right to request that the DNA sample be destroyed at any time. I agree that research data provided by my child or with my permission during the project may be included in a thesis, presented at conferences and published in journals on the condition that either of our names or any other identifying information is used.

The University of Cape Town (UCT) has an appropriate insurance policy to cover payment for any trial-related injury. This study has obtained ethical approval from the UCT Faculty of Health Sciences Research Ethic Committee (FHS REC). If you have any complaints or queries that the investigator has not been able to answer to your satisfaction, you may contact Prof. Marc Blockman from the FHS REC on telephone number 021 406 6452.

FULL NAME OF PARENT /LEGAL GUARDIAN: _____
PARENT /LEGAL GUARDIAN'S SIGNATURE: _____
DATE: _____
INVESTIGATOR: _____
INVESTIGATOR'S SIGNATURE: _____

Participant Questionnaire

GENETIC BASIS OF LIGAMENT INJURY QUESTIONNAIRES

A. PERSONAL PARTICULARS			
Surname			
First Name			
Postal Address		Code	
E-mail address		Phone (day time)	
Date of birth	Y Y Y Y / M M / D D	Cell	
Height (cm)		Gender	Male ☐ Female ☐
Weight (kg)	Pre-Injury:	Current:	
Ethnic group (Only Required and Used for Research Purposes)	Black/African ☐ White ☐ Indian ☐ Mixed Ancestry (Coloured) ☐ Asian ☐ Other ☐		
Ancestry: Tribal or national background (eg Xhosa, Dutch, Zulu, German, Italian)	Father	Unknown ☐	
	Mother	Unknown ☐	
Country of Birth			
Dominant Hand	Left ☐ Right ☐ Both ☐	Dominant Leg	Left ☐ Right ☐ Both ☐
Smoker	Yes (Current) ☐	Yes (Ex-smoker) ☐	No, never ☐
	If yes, Number of years _____	If stopped, when _____	
	If yes, number per day _____		

(If you participate or have participated in more than 6 sports, please complete additional Sporting Details Questionnaires, Part B)

B. SPORTING DETAILS **Please record your sporting activities in order of importance**			
Type of sport(s) you have participated in (please name)	**Main sport 1**	**Other sport 2**	**Other sport 3**
Current or past participation	Current ☐ Past ☐	Current ☐ Past ☐	Current ☐ Past ☐
Year started participation			
Number of years involved in the sport			
Position played prior to injury (if appropriate)			
Playing level prior to injury (if appropriate)			
Number of years played prior to the injury.			

Type of sport(s) you have participated in (please name)	**Other sport 4**	**Other sport 5**	**Other sport 6**
Current or past participation	Current ☐ Past ☐	Current ☐ Past ☐	Current ☐ Past ☐
Year started participation			
Number of years involved in the sport			
Position played prior to injury (if appropriate)			
Playing level prior to injury (if appropriate)			
Number of years played prior to the injury.			

C. ANTERIOR CRUTIATE LIGAMENT INJURY DETAILS

Date of ACL injury?	
Which side was injured?	☐ Left ☐ Right ☐ Both
To what extent was your ligament ruptured?	☐ Complete ☐ Partial ☐ None ☐ Unknown
Investigation done to confirm the diagnosis	☐ MRI ☐ Surgery
How bad is your pain today? (mark line: e.g.I----------**I**----------I)	I----------I No pain Pain as bad as it can be
How was the ACL ruptured?	☐ Direct impact (directly to the injured knee) ☐ Twisting and bending with indirect contact (i.e. contact elsewhere on the body) ☐ Twisting and bending without contact (no external contact) ☐ Skiing ☐ Other (please specify)……
Please describe the exactly how the injury occurred (If uncertain please state that you do not know)	…… …… …… ……
What was the initial treatment? (You may tick more than one block.)	☐ Ice application ☐ Compression ☐ Immobilisation ☐ Medication ☐ Other……..
What was the final treatment?	☐ Surgery ☐ Rehabilitation ☐ Other……..
What are your current symptoms? (You may tick more than one block.)	☐ Pain ☐ Swelling ☐ Instability ☐ Weakness ☐ Other……

What is your current sports participation?	☐ None ☐ Limited to non-weight bearing exercise ☐ Limited, not to same level as pre-injury ☐ Full participation
If you are able to recall, what were the weather and pitch conditions like at the time of injury?	☐ Wet and soft ground ☐ Dry, but soft ground ☐ Dry and firm ground ☐ Wet, but firm ground ☐ Other……..
Associated injuries?	☐ Meniscal tear ☐ MCL tear ☐ Other ligament tear ☐ Bone bruising ☐ Other……..

D. HISTORY OF OTHER LIGAMENT AND TENDON INJURIES IN THE PAST				
Have you ever injured a ligament in the past?	Yes ☐ No ☐			
If yes, please specify which ligaments? (You may tick more than one block, please select either L (left) or R (right))		L R		L R
	Knee (ACL)	☐ ☐	Wrist ligaments	☐ ☐
	Knee (MCL)	☐ ☐	Finger ligaments	☐ ☐
	Ankle lateral ligaments	☐ ☐	Knee (PCL)	☐ ☐
	Spinal ligaments	☐ ☐	Knee (LCL)	☐ ☐
	Shoulder ligaments	☐ ☐	Ankle medial ligaments	☐ ☐
	Elbow ligaments	☐ ☐	Other ligaments	☐ ☐
To your knowledge, have any other members of your family suffered from any ligament injury?	Yes ☐ No ☐	If Yes, please specify the family member ☐ Mother ☐ Father ☐ Sibling ☐ Son / daughter ☐ Other family member…… and condition: Please choose ligament injury from the list above ……		
Have you ever injured a tendon in the past?	Yes ☐ No ☐			
If yes, please specify which tendon? (You may tick more than one block, please select either L (left) or R (right))	Foot and ankle:		L R	
		Achilles tendon	☐ ☐	
		Tibialis posterior	☐ ☐	
		Plantar fascia	☐ ☐	
	Knee:	Patellar tendon	☐ ☐	
	Elbow and wrist:	Wrist extensor tendons	☐ ☐	
	Shoulder:	Subscapularis	☐ ☐	
		Supraspinatus	☐ ☐	
		Infraspinatus	☐ ☐	
		Teres minor	☐ ☐	
	Other:……..			

To your knowledge, have any other members of your family suffered from any tendon pathology?	Yes ☐ No ☐	If Yes, please specify the family member ☐ Mother ☐ Father ☐ Sibling ☐ Son / daughter ☐ Other family member:……..…….. Condition: Please choose tendon injury from the list above ……
Have you ever suffered from any of the following joint capsule injuries?	☐ Acute shoulder dislocation ☐ Chronic shoulder instability ☐ Chronic ankle instability ☐ Other: _____ _____	

E. MEDICAL HISTORY

Do you currently suffer from any of these medical conditions:

☐ High Blood Pressure	☐ Angina/Heart Attack	☐ Asthma
☐ Emphysema	☐ Rheumatoid arthritis	☐ Osteoarthritis (wear & tear)
☐ Malignant disease (cancer)	☐ Elevated Blood Cholesterol	☐ Adrenal disorders
If Yes, what type? _____	☐ Diabetes mellitus	☐ Thyroid disorders
	☐ Renal disease	☐ Amyloidosis

Do you currently suffer from any other Connective Tissue & Rheumatological Diseases & Disorders?	Yes ☐ No ☐	If Yes, please select from the list below

List of some Connective Tissue and/or Rheumatic Diseases and Disorders

☐ Ankylosing Spondylitis	☐ Lipid Storage Diseases	☐ Pseudogout
☐ Aspartylglycosaminuria (AGU)	☐ Marfan Syndrome	☐ Reactive Arthritis
☐ Behcet's Syndrome	☐ Menkes Kinky Hair Syndrome	☐ Reiter's Syndrome
☐ Crohn's Disease	☐ Mucopolysaccharidoses	☐ Relapsing Polychrondritis
☐ Discoid Lupus Erythematosus	☐ Myopathies and Dystrophies	☐ Scleroderma
☐ Ehlers-Danlos syndrome (EDS)	☐ Ochronosis (Homocystinuria)	☐ Sjogren's Syndrome
☐ Eosinophilic Fascitis	☐ Osteogenesis imperfecta (OI)	☐ Systemic Lupus Erythematosus (SLE)
☐ Giant Cell (Temporal) Arthritis	☐ Polyarteritis Nodosa	☐ Systemic Sclerosis
☐ Gout	☐ Polymyalgia Rheumatica	☐ Wegener's Granulomatosis
☐ Hypersentive Vasulatis	☐ Polymyositis & Dermatomyositis	☐ Other_____

What surgical operations have you had? (please list and give dates)	Operation	Date

If female:

At what age did you start menstruating? (years)	
Are you currently using any type of contraception?	☐ Yes ☐ No
If Yes, what type of contraception are you using?	☐ Pill ☐ Injection ☐ IUD

Are you currently?	☐ Pre-menopausal (±12 cycles per year at intervals of 23–33 days & bleeding lasts 3-7 days) ☐ Menopausal (cycles are irregular and less frequent) ☐ Post-menopausal (no longer menstruating)	
Family History		
Do any other members of your family suffer from elevated blood cholesterol?	Yes ☐ No ☐	If Yes, which relative? ☐ Mother ☐ Father ☐ Sibling ☐ Son / daughter ☐ Other relative:……..…….
Is there any history of arthritis in your family?	Yes ☐ No ☐	If Yes, which relative? ☐ Mother ☐ Father ☐ Sibling ☐ Son / daughter ☐ Other relative:……..……. & What type of aethritis? Rheumatoid ☐ Osteoarthritis ☐ Other ☐

Drug and Allergy History	If yes, how long ago (or how many times, where applicable) did you use the medication?	
Have you ever used oral corticosteroids (cortisone tablets)?	Yes ☐ No ☐	☐ 3 months ☐ 6 months ☐ 12 months ☐ 24 or more months
Have you ever been given an injection with corticosteroids?	Yes ☐ No ☐	☐ 3 months ☐ 6 months ☐ 12 months ☐ 24 or more months
Have you ever been given an injection of corticosteroids in or around a tendon?	Yes ☐ No ☐	☐ Once ☐ Twice ☐ 3 times ☐ >3 times
Have you ever used anabolic steroids?	Yes ☐ No ☐	☐ 3 months ☐ 6 months ☐ 12 months ☐ 24 or more months

Have you ever used fluoroquinoline antibiotics?	Yes ☐ No ☐	☐ 3 months ☐ 6 months ☐ 12 months ☐ 24 or more months
If yes, please select from the list below:		
☐ ADCO-CIPRIN ☐ AVELON ☐ BACTIDRON ☐ CIFLOC ☐ CIFRAN ☐ CIPLA-CIPROFLOXACIN ☐ CIPLOXX ☐ CIPRO-HEXAL ☐ Other_____	☐ CIPROBAY ☐ CIPROGEN ☐ CPL ALLIANCE CIPROFLOXACIN ☐ DYNAFLOC ☐ FLOXIN ☐ MAXAQUIN ☐ NOROXIN ☐ ORPIC	☐ SANDOZ CIPROFLOXACIN ☐ TAFLOC ☐ TARIVID ☐ TAVANIC ☐ TEQUIN ☐ UNIQUIN ☐ UTN-400 ☐ ZANOCIN
What medication, if any, are you currently using? (please list)		
What allergies do you have? (please list)		

F. OCCUPATIONAL DETAILS	
What is your current occupation?	
What was your occupation prior to injuring your ligament?	
Prior to injury, did your occupation involve lower limb activity?	☐ Yes ☐ No
If yes please indicate which legs.	Right leg ☐ Both legs ☐ Left leg ☐ None ☐

Appendix B

Ethics Approval

UNIVERSITY OF CAPE TOWN

Health Sciences Faculty
Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory 7925
Telephone [021] 406 6338 • Facsimile [021] 406 6411
e-mail: preaward@curie.uct.ac.za

26 April 2006

REC REF: 164/2006

Dr M Collins
Human Biology

Dear Dr Collins

PROJECT TITLE: THE COL5A1 AND TNC GENES AND THEIR ASSOCIATION WITH ANTERIOR CRUCIATE LIGAMENT INJURIES

Thank you for submitting your study to the Research Ethics Committee for review.

It is a pleasure to inform you that the Ethics Committee has **formally approved** the above-mentioned study on the 21 April 2006.

This serves to confirm that the University of Cape Town Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP) and Declaration of Helsinki guidelines.

The Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.

Please quote the REC. REF in all your correspondence.

Yours sincerely

DR. M. BLOCKMAN
CHAIRPERSON, HSF HUMAN ETHICS

lemjedi

UNIVERSITY OF CAPE TOWN

FACULTY OF HEALTH SCIENCES

Human Research Ethics Committee

Amendment Form

Date	24 May 2012
HREC Ref Number	164/2006
Protocol number (if applicable) & Protocol title	The identification of the genetic risk factors underlying anterior cruciate ligament injuries in all the populations of South Africa.
Principal Investigator	Dr Alison September
Department / Office Internal Mail Address	ESSM, Human Biology, UCT, 3rd Floor SISSA, Newlands

List of Proposed Amendments with Revised Version Numbers and Dates

Partial amendment 164/2006 Renewal of new candidate genes and subject cohorts are provided

RESEARCH ETHICS COMMITTEE
2012 -05- 25
HEALTH SCIENCES FACULTY
UNIVERSITY OF CAPE TOWN

HREC office use only (FWA00001637; IRB00001938)

☑ Approved ☐ Type of review Expedited ☐ Re-submission

This serves as notification that all changes and documentation described above are approved

Signature Chairperson of the HREC	Lesley Berry	Date	28/5/12

PP

1 October 2010 Page 4 of 4 FM-0006

Appendix C

DNA Extraction

As described by Lahiri and Nurnberger (1991) and modified by Mokone *et al.* (2005).

The 5ml of collected blood was transferred from the ethylenediaminetetraacetic acid (EDTA) vacutainer tube to a sterile 15ml polypropolene tube. 2 volumes (10ml) of TKM1 buffer (see solutions below) containing 2.5% NP40 was added to the polypropolene tube. The tube was mixed by inversion several times, and incubated at room temperature for 10 minutes in order to enhance the haemolysis of red blood cells. Centrifugation at 3000rpm at room temperature for 10 minutes then allows the formation of a pellet of white blood cells at the bottom of the tube, and the supernatant of red blood cells may be discarded. This occurs because the heavy white blood cells gravitate to the bottom of the tube during centrifugation, whereas the haemolysed red blood cells are lighter and so remain in the supernatant. 1 volume (5ml) of TKM1 buffer was added to the pellet of white blood cells and the tube was shaken vigorously to re-suspend the pellet and wash it. The tube was then centrifuged at 3000rpm at room temperature for 10 minutes and the supernatant discarded once again. The wash with TKM1 buffer, centrifugation and discarding of supernatant was repeated until the pellet at the bottom of the tube was clean and white. 800µl of TKM2 buffer and 50µl of 10% SDS (see solutions) was added to the pellet. The solution was then shaken again to assist the lyses of the white blood to release the DNA. The solution was then incubated at 55˚C for 1 hour to dissolve the pellet and free the contents of the white blood cells. 150µl of 5M $NaClO_4$ (see solutions) was added, followed by 500µl of chloroform. The solution was vortexed and transferred to 1.5ml microfuge tubes. The tubes were centrifuged at room temperature for 5 minutes at 1300rpm, and 500µl of the top aqueous phase containing the DNA was transferred to a new sterile 1.5ml microfuge tube. 1ml of absolute ethanol was added, and the tube inverted until the DNA precipitated. The tube was centrifuged once again at 1300rpm at room temperature for 5-10 minutes, and the supernatant carefully poured off leaving behind the pellet of DNA at the bottom of the tube. The pellet was allowed to air dry until being re-suspended in 100µl of 1XTE buffer (see solutions). The tube was incubated at 65˚C for 15 minutes in a heating block, after which the DNA was stored at -20˚C for later use.

Solutions:

TKM1 buffer (pH 7.6)

	Molecular Weight	Final concentration	Amount for 500ml
Tris-HCl	121.00	10mM	0.6056g
KCl	74.56	10mM	0.3728g
$MgCl_2.6H_2O$	203.20	10mM	1.016g
EDTA	372.24	2mM	0.372g
dH_2O			To 500ml

TKM2 buffer (pH 7.6)

	Molecular Weight	Final concentration	Amount for 200ml
Tris-HCl	121.00	10mM	0.242g
KCl	74.56	10mM	0.149g
$MgCl_2.6H_2O$	203.20	10mM	0.406g
EDTA	372.24	2mM	0.1488g
NaCl	58.44	0.4M	4.675g
dH_2O			To 200ml

10% SDS

	Final concentration	Amount for 200ml
SDS (10x)	10%	20g
dH_2O		To 200ml

5M $NaClO_4$

	Molecular Weight	Final concentration	Amount for 100ml
$NaClO_4$	122.4	5M	61.2ml
dH_2O			To 100ml

1X TE buffer (pH 8)

	Molecular Weight	Final concentration	Amount for 100ml
Tris-HCl	121.00	10mM	0.121g
EDTA	372.24	1mM	0.037g
dH_2O			To 100ml

Appendix D

Supplementary Tables

Supplementary Table 1: P values of the genotype effect, of the variants within the aggrecan (*ACAN*), biglycan (*BGN*), decorin (*DCN*), fibromodulin (*FMOD*) and lumican (*LUM*) genes, on descriptive measures of participants.

	Age	Sex	Weight	Height	BMI	COB
ACAN						
rs2351491	0.245	0.220	0.521	0.910	0.263	0.347
rs1042631	0.839	**0.033**	0.996	0.329	0.748	0.528
rs1516797	0.605	0.812	0.528	0.905	0.670	0.518
***BGN* - MALE PARTICIPANTS**						
rs1126499	0.777		0.966	0.098	0.086	0.153
rs1042103	0.274		0.318	0.247	0.206	0.455
***BGN* - FEMALE PARTICIPANTS**						
rs1126499	0.092		0.949	0.672	0.984	0.621
rs1042103	0.923		0.170	0.607	0.703	0.571
DCN						
rs13312816	0.125	0.970	0.574	0.748	0.572	0.933
rs516115	0.780	0.549	0.395	0.669	0.276	0.606
FMOD						
rs7543148	0.308	0.967	0.338	0.466	0.144	0.758
rs10800912	0.693	0.866	0.785	0.259	0.489	0.832
LUM						
rs2268578	0.875	0.753	0.924	0.204	0.235	0.274

Significant p values are noted in bold.
Abbreviations: BMI, body mass index; COB, country of birth.

Supplementary Table 2: Genotype and minor allele frequency distributions, as well as the Hardy-Weinberg equilibrium (HWE) values, of the sequence variants investigated within the aggrecan (*ACAN*), biglycan (*BGN*), decorin (*DCN*), fibromodulin (*FMOD*) and lumican (*LUM*) genes in all participants, male and female control (CON) and anterior cruciate ligament rupture (ACL) groups, as well as the ACL subgroup with a non-contact (NON) mechanism of injury.

	All participants			Male			Female		
	CON	**ACL**	**NON**	**CON**	**ACL**	**NON**	**CON**	**ACL**	**NON**
ACAN									
rs2351491									
N	233	225	125	144	165	93	89	60	32
CC	12.0 (28)	15.6 (35)	16.0 (20)	12.5 (18)	17.6 (29)	17.2 (16)	11.2 (10)	10.0 (6)	12.5 (4)
CT	47.2 (110)	45.3 (102)	48.8 (61)	50.7 (73)	44.2 (73)	47.3 (44)	41.6 (37)	48.3 (29)	53.1 (17)
TT	40.8 (95)	39.1 (88)	35.2 (44)	36.8 (53)	38.2 (63)	35.5 (33)	47.2 (42)	41.7 (25)	34.4 (11)
p Value		0.547	0.437		0.365	0.598		0.717	0.445
HWE	0.773	0.573	1.000	0.478	0.335	0.833	0.638	0.777	0.713
C allele	35.6	38.2	40.4	37.9	39.7	40.7	32.0	34.2	39.1
p Value		0.415	0.208		0.638	0.511		0.699	0.308
rs1042631									
N	233	227	126	144	166	94	89	61	32
TT	3.0 (7)	4.4 (10)	4.0 (5)	2.8 (4)	4.2 (7)	3.2 (3)	3.4 (3)	4.9 (3)	6.3 (2)
TC	28.8 (67)	35.7 (81)	36.5 (46)	34.0 (49)	38.0 (63)	40.4 (38)	20.2 (18)	29.5 (18)	25.0 (8)
CC	68.2 (159)	59.9 (136)	59.5 (75)	63.1 (91)	57.8 (96)	56.4 (53)	76.4 (68)	65.6 (40)	68.8 (22)
p Value		0.168	0.254		0.564	0.575		0.349	0.635
HWE	1.000	0.709	0.796	0.598	0.519	0.388	0.184	0.684	0.276
T allele	17.4	22.3	22.2	19.8	23.2	23.4	13.5	19.7	18.8
p Value		0.064	0.115		0.305	0.346		0.151	0.310
rs1516797									
N	233	227	126	144	166	94	89	61	32
TT	51.1 (119)	42.3 (96)	47.6 (60)	50.0 (72)	44.0 (73)	51.1 (48)	52.8 (47)	37.7 (23)	37.5 (12)
TG	42.9 (100)	46.7 (106)	42.9 (54)	44.4 (64)	44.0 (73)	40.4 (38)	40.5 (36)	54.1 (33)	50.0 (16)
GG	6.0 (14)	11.0 (25)	9.5 (12)	5.6 (8)	12.1 (20)	8.5 (8)	6.7 (6)	8.2 (5)	12.5 (4)
p Value		0.059	0.451		0.122	0.617		0.188	0.275
HWE	0.321	0.662	1.000	0.299	0.863	1.000	1.000	0.261	1.000
G allele	27.5	34.4	31.0	27.8	34.0	28.7	27.0	35.3	37.5
p Value		**0.024**	0.325		0.093	0.823		0.125	0.114

	All participants			Male			Female		
	CON	**ACL**	**NON**	**CON**	**ACL**	**NON**	**CON**	**ACL**	**NON**
BGN									
rs1126499									
N				143	166	94	88	61	32
CC				51.6 (74)	48.2 (80)	48.9 (46)	23.9 (21)	18.0 (11)	21.9 (7)
CT							55.7 (49)	45.9 (28)	43.8 (14)
TT				48.3 (69)	51.8 (86)	51.1 (48)	20.5 (18)	36.1 (22)	34.4 (11)
p Value					0.533	0.672		0.105	0.278
HWE							0.394	0.790	0.496
T allele				48.3	51.8	51.1	48.3	59.0	56.3
P Value					0.378	0.549		0.068	0.276
rs1042103									
N				143	166	94	89	61	32
GG				55.9 (80)	60.8 (101)	59.6 (56)	38.2 (34)	26.2 (16)	28.1 (9)
GA							49.4 (44)	54.1 (33)	56.3 (18)
AA				44.1 (63)	39.2 (65)	40.4 (38)	12.4 (11)	19.7 (12)	15.6 (5)
p Value					0.383	0.580		0.226	0.584
HWE							0.652	0.612	0.718
A allele				44.1	39.2	40.4	37.1	46.7	43.8
p Value					0.218	0.434		0.095	0.348
DCN									
rs13312816									
N	234	227	126	144	166	94	90	61	32
AA	1.3 (3)	0.0 (0)	0.0 (0)	1.4 (2)	0.0 (0)	0.0 (0)	1.1 (1)	0.0 (0)	0.0 (0)
AT	16.2 (38)	15.4 (35)	15.9 (20)	14.6 (2)	17.5 (29)	17.0 (16)	18.9 (17)	9.8 (6)	12.5 (4)
TT	82.5 (193)	84.6 (192)	84.1 (106)	84.0 (121)	82.5 (137)	83.0 (78)	80.0 (73)	90.2 (55)	87.5 (28)
p Value		0.221	0.438		0.256	0.465		0.214	0.584
HWE	0.437	0.372	1.000	0.281	0.612	1.000	1.000	1.000	1.000
A allele	9.4	7.7	7.9	8.7	8.7	8.5	10.6	4.9	6.3
p Value		0.359	0.510		0.981	0.949		0.081	0.311
rs516115									
N	234	227	126	144	166	94	90	61	32
AA	45.7 (107)	45.4 (103)	46.8 (59)	50.0 (72)	42.2 (70)	40.4 (38)	38.9 (35)	54.1 (33)	65.6 (21)
AG	43.2 (101)	44.5 (101)	42.9 (54)	40.3 (58)	44.6 (74)	45.7 (43)	47.8 (43)	44.3 (27)	34.4 (11)
GG	11.1 (26)	10.2 (23)	10.3 (13)	19.7 (14)	13.3 (22)	13.8 (13)	13.3 (12)	1.6 (1)	0.0 (0)
p Value		0.926	0.965		0.334	0.305		**0.021**	**0.012**
HWE	0.767	0.881	1.000	0.692	0.736	1.000	1.000	0.156	0.554
G allele	32.7	32.4	31.8	29.9	35.5	36.7	37.2	23.8	17.2
p Value		0.919	0.796		0.133	0.120		**0.014**	**0.003**

	All participants			Male			Female		
	CON	**ACL**	**NON**	**CON**	**ACL**	**NON**	**CON**	**ACL**	**NON**
FMOD									
rs7543148									
N	232	226	125	143	165	93	89	61	32
GG	4.7 (11)	3.1 (7)	3.2 (4)	5.6 (8)	2.4 (4)	3.2 (3)	3.4 (3)	4.9 (3)	3.1 (1)
GA	31.5 (73)	28.3 (64)	26.4 (33)	29.4 (42)	29.7 (49)	28.0 (26)	34.8 (31)	24.6 (15)	21.9 (7)
AA	63.8 (148)	68.6 (155)	70.4 (88)	65.0 (93)	67.9 (112)	68.8 (64)	61.8 (55)	70.5 (43)	75.0 (24)
p Value		0.458	0.428		0.355	0.657		0.392	0.389
HWE	0.554	0.820	0.744	0.303	0.788	0.728	0.755	0.355	0.481
G allele	20.5	17.3	16.3	20.3	17.3	17.2	20.8	17.2	14.1
p Value		0.214	0.186		0.340	0.406		0.441	0.270
rs10800912									
N	233	225	124	144	164	92	89	61	32
CC	3.0 (7)	3.1 (7)	3.2 (4)	3.5 (5)	3.1 (5)	4.4 (4)	2.3 (2)	3.3 (2)	0.0 (0)
CT	23.2 (54)	27.1 (61)	24.2 (30)	22.9 (33)	28.1 (46)	25.0 (23)	23.6 (21)	24.6 (15)	21.9 (7)
TT	73.8 (172)	69.8 (157)	72.6 (90)	73.6 (106)	68.9 (113)	70.7 (65)	74.2 (66)	72.1 (44)	78.1 (25)
p Value		0.616	0.968		0.586	0.868		0.914	0.671
HWE	0.295	0.637	0.487	0.316	0.787	0.275	0.676	0.618	1.000
C allele	14.6	16.7	15.3	14.9	17.1	16.9	14.0	15.6	10.9
p Value		0.387	0.794		0.470	0.576		0.713	0.668
LUM									
rs2268578									
N	234	227	126	144	166	94	90	61	32
TT	2.6 (6)	1.3 (3)	0.8 (1)	3.5 (5)	1.2 (2)	1.1 (1)	1.1 (1)	1.6 (1)	0.0 (0)
TC	26.5 (62)	25.6 (58)	23.8 (30)	23.6 (34)	27.1 (45)	25.5 (24)	31.1 (28)	21.3 (13)	18.8 (6)
CC	70.9 (166)	73.1 (166)	75.4 (95)	72.9 (105)	71.7 (119)	73.4 (69)	67.8 (61)	77.1 (47)	81.3 (26)
p Value		0.598	0.412		0.343	0.497		0.406	0.325
HWE	1.000	0.585	0.689	0.329	0.525	1.000	0.445	1.000	1.000
T allele	15.8	14.1	12.7	15.3	14.8	13.8	16.7	12.3	9.4
p Value		0.466	0.261		0.857	0.663		0.295	0.218

Genotype and minor allele frequencies are reported as a percentage with the number of observations (n) in parenthesis. The total number of samples genotyped (N) is also indicated. ACL and NON p values are the allelic differences between the ACL and CON, and NON and CON groups respectively.