The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

Extracellular matrix gene sequence variant analyses and Achilles tendinopathy

By

Colleen Jayne Saunders

Submitted to the University of Cape Town

in fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in the Department of Human Biology, Faculty of Health Sciences

UNIVERSITY OF CAPE TOWN

April 2013

University of Cape Town

Supervisors:

Prof. Malcolm Collins[2,1]

Dr. Alison September[1]

Prof. Martin Schwellnus[1]

[1]MRC/UCT Research Unit for Exercise Science and Sports Medicine, Department of Human Biology, Faculty of Health Sciences, University of Cape Town, Newlands, South Africa

[2]South African Medical Research Council, Cape Town, South Africa

DECLARATION

Doctoral Degrees Board
University of Cape Town
Private Bag X3, Rondebosch
7701 South Africa
Tel: (021) 650-2202
Fax: (021) 650-4913
janine.isaacs@uct.ac.za

PhD THESIS TITLE: EXTRACELLULAR MATRIX GENE SEQUENCE VARIANT ANALYSES AND ACHILLES TENDINOPATHY

1. DECLARATIONS:

I, Colleen Jayne Saunders, hereby:

(a) grant the University of Cape Town free licence to reproduce the above thesis in whole or in part, for the purpose of research;

(b) declare that:

(i) the above thesis is my own unaided work, both in concept and execution, and that apart from the normal guidance from my supervisor, I have received no assistance except as stated below:

I gratefully received guidance and assistance with the more complex statistical analyses included in this thesis from Dr. Lize van der Merwe, as included in the acknowledgments._____

(ii) neither the substance nor any part of the above thesis has been submitted in the past, or is being, or is to be submitted for a degree at this University or at any other university, except as stated below:

_____N/A_____

_____

I am now presenting the thesis for examination for the degree of PhD.

Candidate's Signature		**Date**	

University of Cape Town

DEDICATION

This thesis is dedicated to my father, Butch, who will never see the completion of this dream, and my mother, Teena, for their love and uncompromising support throughout my student career. Words do not adequately convey my appreciation for them allowing me the freedom to follow my dreams.

TEACH YOUR CHILDREN
Crosby, Stills, Nash and Young

You who are on the road, must have a code that you can live by.
And so, become yourself, because the past is just a good bye.
Teach your children well, their father's hell did slowly go by,
And feed them on your dreams, the one they picked is the one you're known by.
Don't you ever ask them why, if they told you you would cry,
So just look at them and sigh, and know they love you.

And you of the tender years, can't know the fears that your elders grew by.
And so please help them with your years, they seek the truth before they can die.
Teach your parents well, their children's hell will slowly go by.
And feed them on your dreams, the one they picked is the one you're known by.
Don't you ever ask them why, if they told you you would cry,
So just look at them and sigh, and know they love you.

University of Cape Town

ACKNOWLEDGEMENTS

Thank-you to my extremely patient supervisors, Dr. Alison September and Prof. Malcolm Collins, for sharing this experience with me. Your knowledge, guidance and support have been invaluable to me and I will be forever grateful for the solid foundation you have given my academic career. Alison – your technical expertise in the laboratory and pep talks helped me push through many obstacles. Malcolm – your open-door policy and engaging teaching practice has been invaluable and helped me build confidence in my own abilities.

Thank-you to my "stats guru", Dr. Lize van der Merwe, for your guidance and assistance with the more complex statistical analyses undertaken in this thesis.

The love and mountains of support I received from my partner, (soon-to-be Dr.) Benoit Capostagno, has been invaluable. I hope I can return the favour!

Thank-you to my sister, Janice, who has been continuously supportive and kept me supplied with home-cooked meals and cookies whenever the going got tough.

Thank-you to all my colleagues and friends at ESSM who help foster a supportive and productive environment. In particular Dr. Kristina Plattner, Dr. Caroline D'Alton, Nicholas Tam, Dr. Liesl de Milander, Dr. Yumna Albertus-Kajee, Dr. Michael Posthumus, Dr. Robert Lamberts and Mark Kirkman have travelled this path with me and made it an enjoyable journey. I could not have wished for more supportive and inspiring office companions than the hard-working Dr. Sharief Hendricks, James Brown and Benoit Capostagno – keep the coffee and laughs rolling! My close friends, particularly Sarah Sandmann and Kim Germishuys, have loved and supported me for many years – I hope to spend more time on the beach with you now!

Thank-you to the wonderful administrative staff at ESSM who have assisted me with a smile so often, as well as Adri Winkler and her team at the Faculty of Heath Sciences postgraduate office who are always helpful, efficient and supportive. Thank you also to Neezaam Kariem and Trevino Larry for their willing assistance in the laboratory.

University of Cape Town

Thank you to the South African and Australian participants who donated their time and DNA samples to these studies. Thank you to Dr. George Mokone, Prof. Martin Schwellnus, Prof. Jill Cook and Prof. Christopher Handley who were involved in the initial recruitment and diagnosis of the South African and Australian participants. I would also like to acknowledge the contribution of Abrahams et al.,[6] El Khoury et al.,[76] Nell et al.,[205] Posthumus et al.[224] and September et al.[255,257] who completed the previously published genotyping for the polymorphisms included in chapter 5 of this thesis.

Thank you to the examiners who kindly agreed to review and examine this thesis. Your time and valuable input is greatly appreciated, and your constructive comments greatly helped to improve the quality of this thesis.

Finally, my studies would not have been possible without financial assistance gratefully received from the University of Cape Town postgraduate funding office (Benfarra scholarship, Harry Crossley Foundation scholarship, Marion Beatrice Waddell scholarship, University Research Committee scholarship and UCT conference travel grant) and the National Research Foundation (NRF Grantholder bursary).

University of Cape Town

TABLE OF CONTENTS

University of Cape Town

University of Cape Town

University of Cape Town

University of Cape Town

University of Cape Town

SCIENTIFIC OUTPUTS ASSOCIATED WITH THIS THESIS

PUBLICATIONS IN PEER-REVIEWED JOURNALS

Saunders, C.J., Van Der Merwe, L., Posthumus, M., Cook, J., Handley, C.J., Collins, M., & September, A.V. Investigation of variants within the *COL27A1* and *TNC* genes and Achilles tendinopathy in two populations. Journal of Orthopaedic Research. 2013. 31(4), 632–637

Saunders, C.J., Van Der Merwe, L., Cook, J., Handley, C.J., Collins, M., & September, A.V. Variants within the *COMP* and *THBS2* genes are not associated with Achilles tendinopathy in a case-control study of South African and Australian populations. Journal of Sports Sciences. *In press*

PRESENTATIONS AT INTERNATIONAL CONFERENCES

Posthumus, M., Saunders, C., September, A.V., Collins, M. The polygenic profiles in participants with Achilles tendinopathy and controls. IOC World Conference on Prevention of Injury and Illness in Sport, Monte-Carlo, Principality of Monaco, 2011 (Poster Presentation).

Saunders, C.J., September, A.V., Posthumus, M., Schwellnus, M., Cook, J., Handley, C., van der Merwe, L., Collins, M. Haplotype analysis of the *COL27A1* and *TNC* genes as potential risk factors for Achilles tendon pathology. 2011 Joint AfSHG and SASHG Conference, Cape Town, March 2011 (Poster Presentation).

Posthumus M., Saunders C., September A.V. and Collins M. The polygenic profiles in participants with Achilles tendinopathy and controls. The International Scientific Tendinopathy Symposium, Umeå, Sweden, 2010 (Poster Presentation).

Saunders,C.J.; September,A.V.; Schwellnus,M.; Collins,M. Haplotype analysis of a candidate gene as a risk factor for Achilles tendon pathology. European College of Sport Science 13th Annual congress, Lisbon, Portugal, 2008 (Poster Presentation).

University of Cape Town

PRESENTATIONS AT LOCAL CONFERENCES

Saunders, C.J, September, A.V., Schwellnus. M.P., Collins, M. Genetic analysis of an additional candidate gene as a risk factor for Achilles tendon pathology. MRC research day. 18 October 2007 (Poster Presentation).

ABBREVIATIONS

A:	Adenine
AAS:	Androgenic anabolic steroids
ACL:	Anterior cruciate ligament
ADAM:	A disintegrin and metalloproteinase
ADAMTS:	A disintegrin-like and metalloproteinase with thrombospondin motifs
AGE:	Advanced glycation end product
ANOVA:	Analysis of variance
AT:	Achilles tendinopathy
ATI:	Achilles tendon injury
ATR:	Achilles tendon rupture
AUC:	Area under the curve
AUS:	Australian participants
bp:	Base pairs
BMI:	Body mass index
C:	Cytosine
CASP8:	Gene encoding caspase-8
Chr:	Chromosome
CI:	Confidence interval
COL3A1:	Gene encoding α1(III) collagen
COL5A1:	Gene encoding α1(V) collagen
COL5A2:	Gene encoding α2(V) collagen
COL5A3:	Gene encoding α3(V) collagen
COL27A1:	Gene encoding α1(XXVII) collagen
COMP:	Cartilage oligomeric matrix protein
COMP:	Gene encoding cartilage oligomeric matrix protein
CON:	Control
COOH:	Carboxy terminal
DAMP:	Damage associated molecular pattern
dATP:	Deoxyadenosine triphosphate
dCTP:	Deoxycytidine triphosphate
dGTP:	Deoxyguanosine triphosphate
DNA:	Deoxyribonucleic acid
dTTP:	Deoxythymidine triphosphate
ECM:	Extracellular matrix

University of Cape Town

EDS: Ehlers-Danlos syndrome
EDTA: Ethylenediaminetetraacetic acid
EGF: Epidermal growth factor
ERT: Estradiol replacement therapy
ESE: Exonic splicing enhancer
ESS: Exonic splicing silencer

FACIT: Fibril associated collagen with interrupted triple helices
FNIII: Fibronectin type III domain

G: Guanine
GAG: Glycosaminoglycan
GDF5: Gene encoding growth/differentiation factor-5
Gln: Glutamine
Glu: Glutamic acid
Gly: Glycine
GRA: Genetic risk assessment
GVS: Genome variation server
GWAS: Genome-wide association studies

HRT: Hormone replacement therapy
HWE: Hardy-Weinberg Equilibrium

IGF-1: Insulin-like growth factor-1
IL-1β: Interleukin-1β
IL-1β: Gene encoding interleukin-1β
IL-1ra: Interleukin-1 receptor antagonist
IL-1RN: Gene encoding interleukin-1 receptor antagonist
IL-6: Interleukin-6
IL-6: Gene encoding interleukin-6
Ile: Isoleucine

Kbp: Kilo base pairs (1000 base pairs)

lamG: Laminin-G domain
LD: Linkage disequilibrium
Leu: Leucine

miRNA: Micro ribonucleic acid
MIR608: Gene encoding micro RNA Has-miR-608
MMP: Matrixmetalloproteinase

University of Cape Town

MMP3:	Gene encoding matrix metalloproteinase-3
mRNA:	Messenger ribonucleic acid
MTJ:	Myotendinous junction
NC:	Non-collagenous
NCBI:	National Centre for Biotechnology Information
NFL:	National Football League
NH_2:	Amino terminal
OC:	Oral contraceptives
OR:	Odds Ratio
OTJ:	Osteotendinous junction
p:	Short arm of chromosome
P:	Probability
PAGE:	Polyacrylamide gel electrophoresis
PARP:	Proline and arginine rich protein domain
PCR:	Polymerase chain reaction
Phe:	Phenylalanine
q:	Long arm of chromosome
RCT:	Randomised controlled trial
RFLP:	Restriction fragment length polymorphism
ROC:	Receiver operating characteristic
ROM:	Range of motion
RR:	Relative risk
SA:	South African participants
SDFT:	Superficial digital flexor tendon
SLRP:	Small leucine rich proteoglycan
SNP:	Single nucleotide polymorphism
T:	Thymine
T-ARMS:	Tetra-primer amplification refractory mutation system
TA:	Tenascin Assembly
TEN:	Tendinopathy participants
THBS2:	Gene encoding thrombospondin-2
TIMP:	Tissue inhibitor of metalloproteinases
TIMP2:	Gene encoding tissue inhibitor of metalloproteinases-2
TLR:	Toll-like receptor

University of Cape Town

TN-C:	Tenascin-C
TNC:	Gene encoding tenascin-C
TRIPP:	Translating research into injury prevention practice
TSP:	Thrombospondin
TSPN:	Thrombospondin NH_2-terminal like
TSR:	Thrombospondin type-1 repeat
UTR:	Untranslated region
vWFC:	von Willebrand Factor type-C

University of Cape Town

LIST OF TABLES

University of Cape Town

University of Cape Town

LIST OF FIGURES

University of Cape Town

University of Cape Town

University of Cape Town

University of Cape Town

University of Cape Town

ETHICAL APPROVAL & FUNDING

The work included in this thesis was approved by the Research Ethics Committee of the Faculty of Health Sciences within the University of Cape Town, South Africa (reference numbers 289/2004 and 086/2005) and the Human Ethics Committee of La Trobe University, Melbourne, Australia [Appendix A].

This research was supported by funds from the National Research Foundation (NRF grant numbers SUR2008060500012 and FA2005021700015), the South African Medical Research Council, the University of Cape Town and Discovery Health.

University of Cape Town

THESIS ABSTRACT

BACKGROUND

The incidence rates of Achilles tendon rupture and Achilles tendinopathy (AT) are increasing. There are clear clinical, histopathological and imaging criteria for the diagnosis of AT, however its exact aetiology has not been elucidated. The number of intrinsic and extrinsic risk factors associated with AT emphasise the multifactorial nature of this condition. The risk factors with the strongest supportive evidence and highest degree of certainty are previous Achilles tendon injury, neovascularisation of the tendon, treatment with fluoroquinolone antibiotics, middle- and long-distance running and genetic variation. The certainty for the existence of a genetic predisposition towards AT is increasing as more studies are published reporting the association of gene sequence variants with risk of AT. In addition, the increasing number of genes that are reported to be associated with AT emphasises the polygenic nature of genetic susceptibility to AT. It is therefore vital that the full spectrum of polymorphisms which influence risk of AT is eventually uncovered.

AIM OF THE THESIS

The primary aim of this thesis was to identify additional genetic elements predisposing individuals to risk of AT using a candidate gene, case-control genetic association approach, and to propose the biological mechanisms underlying this genetic risk. Candidate genes (*COMP*, *THBS2*, *COL27A1*, *TNC*, *COL3A1, COL5A2* and *COL5A3*) were selected based on their chromosomal location and/or the biological function of their encoded proteins within the extracellular matrix (ECM) of the tendon. The objectives of the specific studies which addressed this primary aim of the thesis were:

- To test the association of sequence variants within the *COMP* and *THBS2* candidate genes with risk of AT in participants from South Africa and Australia (Chapter 2)
- To test the association of several sequence variants spanning a 780.9kb region which includes the *COL27A1* and *TNC* genes, and a potential haplotype consisting of two or more of these SNPs, with AT in participants from South Africa and Australia (Chapter 3)

University of Cape Town

- To investigate the association of polymorphisms within the *COL5A2*, *COL5A3* and *COL3A1* genes with AT in participants from South Africa and Australia (Chapter 4)
- To investigate interactions between polymorphisms within the *COL5A2*, *COL5A3* and *COL3A1* genes and the previously associated *BstUI* RFLP within the *COL5A1* gene in modulating risk of AT (Chapter 4)

The secondary aim of this thesis was to develop and test preliminary polygenic models assessing risk of developing AT. The objectives of the specific studies which addressed this secondary aim of the thesis were:

- To investigate the relative contribution of polymorphisms within genes encoding components of the collagen fibril and components of cell-signalling pathways within the ECM to overall genetic risk in a polygenic risk model (Chapter 5)
- To develop and evaluate a preliminary, clinically relevant polygenic risk assessment model for AT using (i) polymorphisms that have previously been independently associated with AT in two populations, and (ii) polymorphisms with a definitive AT risk genotype that were implicated in this thesis in a haplotype or gene-gene interaction (Chapter 5)

METHODS

One hundred and thirty one (131) physically active South African Caucasian participants without any history of tendon or ligament pathology (SA-CON), as well as 94 physically active South African Caucasian participants clinically diagnosed with chronic AT (SA-TEN) were recruited for the studies presented in chapter's two to five. In addition, 209 physically active Australian Caucasian participants without any history of tendon or ligament pathology (AUS-CON), as well as 85 physically active Australian Caucasian participants clinically diagnosed with chronic AT (AUS-TEN) were also recruited for the studies presented in this thesis. To address the primary aim of this thesis, participants were genotyped for variants within *COMP*, *THBS2* (Chapter 2), *COL27A1*, *TNC* (Chapter 3), *COL3A1*, *COL5A2*, and *COL5A3* (Chapter 4) using standard PCR based methods. Where appropriate, haplotypes or pseudohaplotypes were inferred using genotype data from the respective variants. Stepwise logistic regression and genotype risk scores were used to address the secondary

University of Cape Town

aim of this thesis by identifying polymorphisms which contribute to overall genetic risk of AT and to develop a series of AT genetic risk assessment models.

RESULTS AND DISCUSSION

Polymorphisms within the *COMP* (rs284945050, P=0.994; rs730079, P=0.727) and *THBS2* genes (rs9505888, P=0.244; rs6422747, P=0.645; rs9283850, P=0.604) were found not to independently associate with AT in the South African and Australian groups (Chapter 2). In addition, polymorphisms within the *COL5A3* (rs2303099, P=0.890; rs1559186, P=0.050; rs2161468, P=0.473), *COL3A1* (rs2056156, P=0.928; rs3106796, P=0.900; rs1800255, P=0.726) and *COL5A2* (rs13031549, P=0.969; rs4667264, P=0.915) genes were also found not to be significantly independently associated with AT in these two groups (Chapter 4). Furthermore, four polymorphisms within the *COL27A1* gene (rs4143245, P=0.610; rs1249744, P=0.106; rs753085, P=0.199; rs946053, P=0.248) and the rs13321 SNP within the *TNC* gene (P=0.291) were also found not to be independently associated with AT (Chapter 3). However, allele distributions of two SNPs within the *TNC* gene (rs2104772, P=0.017; rs1330363, P=0.020) were found to differ significantly between the TEN and CON participants. In addition, the GCA haplotype constructed from the *COL27A1* rs946053 (T>G), *TNC* rs13321 (G>C) and *TNC* rs2104772 (T>A) SNPs was significantly associated with AT in the South African and Australian participants (TEN: 27% vs CON: 18%, P=0.019). Bioinformatic analysis of this region suggests that this haplotype may have functional effects on mRNA transcription by influencing the position and accessibility of transcription factor binding sites and splicing regulatory elements. The potential changes in structure and properties of the TN-C and α1(XXVII) proteins may modify protein-protein and protein environment interactions.

In addressing the secondary aim of this thesis, stepwise logistic regression with receiver operating characteristic (ROC) curve analysis was used to develop two risk models for AT which identified the polymorphisms contributing, and their relative contribution, to overall genetic risk of AT. The best fit risk model with the largest area under the curve included the variables age (years), sex (male), *COL27A1* rs946053 (GT,TT), *COL5A1* rs12722 (TC,CC), *COL5A3* rs1559186 (CG,GG), *IL-6* rs1800795 (GC,CC), *CASP8* rs1045485 (GC,CC) and *CASP8* rs3834129 (del/CTTACT,del/del), and had sensitivity and specificity indicators of 58.1% and

University of Cape Town

86.5% respectively. These risk models are valuable at a population level, but are not easy to interpret in a clinical setting. Therefore, a series of potentially clinically relevant genetic risk assessment models were developed and evaluated using genotype risk scores. A genetic risk assessment model which included risk genotype scores for polymorphisms rs16399 (*COL5A1*), rs4919510 (*MIR608*), rs143383 (*GDF5*), rs3834129 (*CASP8*), rs2104772 (*TNC*) and rs946053 (*COL27A1*) was developed and accurately assessed 90% of CON participants as not at-risk of AT but only accurately assessed 37% of TEN participants as at-risk of AT (OR 5.89, 95% CI:2.39-14.52; P<0.001).

CONCLUSION

The novel findings of this thesis provide further evidence that the long arm of chromosome 9 harbours genetic elements which contribute to risk of AT, and narrows the implicated interval to a 759kbp region containing the 3'-end of the *COL27A1* gene and the 5'-end of the *TNC* gene. Polymorphisms investigated within the *COMP*, *THBS2*, *COL5A3*, *COL3A1* and *COL5A2* genes were, however, not associated with AT in these two participant groups.

In addition, this thesis provides proof of concept that genetic screening tests for AT can be of value as a clinical tool. Future studies are required in order to replicate these findings in independent populations and elucidate the biological mechanisms underpinning these associations. Large, prospective cohort studies are required to confirm the involvement of genetic and other risk factors in the development of AT and to test the ability of genetic screening tests to effectively discriminate risk of AT.

University of Cape Town

CHAPTER 1: ACHILLES TENDINOPATHY: A REVIEW

1.1. INTRODUCTION AND SCOPE OF THESIS

Regular physical activity is important in the maintenance of a healthy lifestyle and the reduction of mortality risk.[28] However, the risk of musculoskeletal injuries rises with increasing physical activity.[200] In particular, tendon injuries contribute an estimated 30-50% to all sporting injuries.[115,131] The Achilles tendon is particularly susceptible to tendinopathy which accounts for 55-65% of Achilles tendon disorders.[132,233,306] Achilles tendinopathy (AT) presents a large burden of morbidity and is the focus of this thesis. The current chapter (Chapter 1) introduces the scope and objectives of this thesis (Section 1.1), presents a brief review of tendon anatomy (Section 1.2) and molecular structure (Section 1.3) as well as a more detailed review of Achilles tendon pathology (Section 1.4) and the risk factors associated with AT (Section 1.5).

A framework for translating research into injury prevention practice (TRIPP) proposed by Finch (2006)[82] emphasises the importance of injury surveillance and elucidating the aetiology and mechanisms of injury before developing and evaluating preventative measures. The identification of both intrinsic and extrinsic risk factors for AT is an important element in understanding the aetiology and mechanism of injury. A large volume of research has been undertaken to establish the prevalence and incidence of AT, as well as the risk factors that are associated with the development of AT. A number of preventative measures and treatments have been developed and evaluated based on this research. However, research linking genetic variation to risk of AT has yet to be fully translated into injury prevention practice and remains at stage 2 of the TRIPP model in which the mechanisms for injury that are influenced by genetic variation are still being expounded.

The primary aim of this thesis was therefore to further identify specific genetic elements predisposing individuals to risk of AT and postulate on the biological mechanisms underlying this genetic risk. Candidate genes were selected based on their chromosomal location and/or their biological function within the extracellular matrix (ECM) of tendon tissue. A case-control genetic association study approach was used to investigate the association of

University of Cape Town

polymorphisms within these genes with risk of AT (Chapters 2, 3 and 4). The ultimate purpose of injury research is the development and implementation of relevant and practical injury prevention strategies which reduce risk of injury. With this in mind, the second objective of this thesis was to develop and test preliminary models assessing genetic risk of developing AT. In chapter 5 of this thesis, polymorphisms previously independently associated with risk of AT, as well as new polymorphisms identified in this thesis, were incorporated into several genetic risk assessment models which were tested in an applied case-control analysis. The final chapter of this thesis, chapter 6, summarises the main findings of the thesis and discusses the limitations, implications and context of this research as well as suggesting avenues of future research.

1.2. TENDON ANATOMY

1.2.1. MACRO-STRUCTURE OF TENDONS

The general structure and metabolism of tendons is extensively reviewed elsewhere,[60,208] but will be briefly elaborated here. Tendons are connective tissue structures that link muscle to bone, and vary greatly in both shape and size. The myotendinous junction (MTJ) is the intersection between the muscle and tendon and is the point from which the tendon elongates. The MTJ is subjected to immense mechanical stress during the transmission of contractile force from the muscle to the tendon. The osteotendinous junction (OTJ), also known as the enthesis, is the gradual insertion of the tendon into lamellar bone. The Achilles tendon, in particular, is the merging of the myotendinous areas of the gastrocnemius and soleus muscles of the lower leg (Figure 1.1).[132] It is the largest and strongest tendon in the human body and has a high capacity to withstand tensional forces.[132] Healthy tendons can, in fact, withstand tensile loads much greater than those required for normal function.[239]

University of Cape Town

Figure 1.1: Gross anatomy of the Achilles tendon
[Source: www.photobucket.com]

University of Cape Town

Tendons are relatively avascular with lower oxygen consumption than other tissues. However, they are metabolically active and respond to external stimuli. Blood supply to the tendons extends predominantly from the muscle and divides into three regions - the MTJ, tendon length and OTJ, although very few blood vessels cross the OTJ.[9,208] The lower metabolic rate of tendons allows them to sustain tension for extended periods of time with a minimal risk of ischaemia and tissue necrosis.[115] However, blood supply to tendons is often restricted in areas of friction, torsion or compression, and this is particularly the case in the Achilles tendon where the tendon fibres can spiral by up to 90°.[109,208] This twisting increases the strength of the tendon under tension and endows greater elastic recoil properties but may cause constriction of the associated blood vessels.[109,280]

Nerve supply to the tendon is predominantly afferent with mechanoreceptors being found near the MTJ.[138] These mechanoreceptors are responsible for proprioception and provide feedback control of muscle activity to protect the joints.[115] An interrupted nerve supply to the tendon results in the atrophy of both muscles and tendons.

1.2.2. HIERARCHICAL MICRO-STRUCTURE OF TENDONS

Tendon tissue consists of a cellular component and the surrounding ECM. The spindle shaped fibroblasts, known as tenocytes, are the cellular elements within tendons and are responsible for secreting collagen and other molecules, and maintenance of the ECM.[115,306] The ECM consists of tightly packed parallel collagen fibres surrounded by ground substance. Ground substance is produced by the tenocytes and is a highly viscous matrix which consists of 60-80% water, glycosaminoglycan's (GAG) and structural proteoglycans.[208] The base unit of the hierarchical collagen network is a type I procollagen molecule which conglomerates in groups of five to form fibrils that are 20-150nm in diameter (Figure 1.2). Several fibrils embedded in parallel in the ECM will form fibres, which in turn group together to form fascicles individually encapsulated by the endotenon. The endotenon is a connective tissue sheath that accommodates the nerves and blood vessels. The fascicles are further encapsulated by the epitenon which is surrounded by a fluid layer and the outer paratenon.[115,208] Crimp is a significant feature of tendon which represents a sinusoidal pattern in the ECM.[208] It functions as a shock-absorbing buffer so that slight elongation of the tendon can occur without the risk of damage.

Healthy tendon tissue is white in colour with a firm fibroelastic texture. The parallel collagen bundles have a regular, uniform structure and the individual fibres are uniform in diameter and orientation. The tenocytes have uniform, spindle shaped nuclei and are arranged in parallel with the collagen fibres.[306]

University of Cape Town

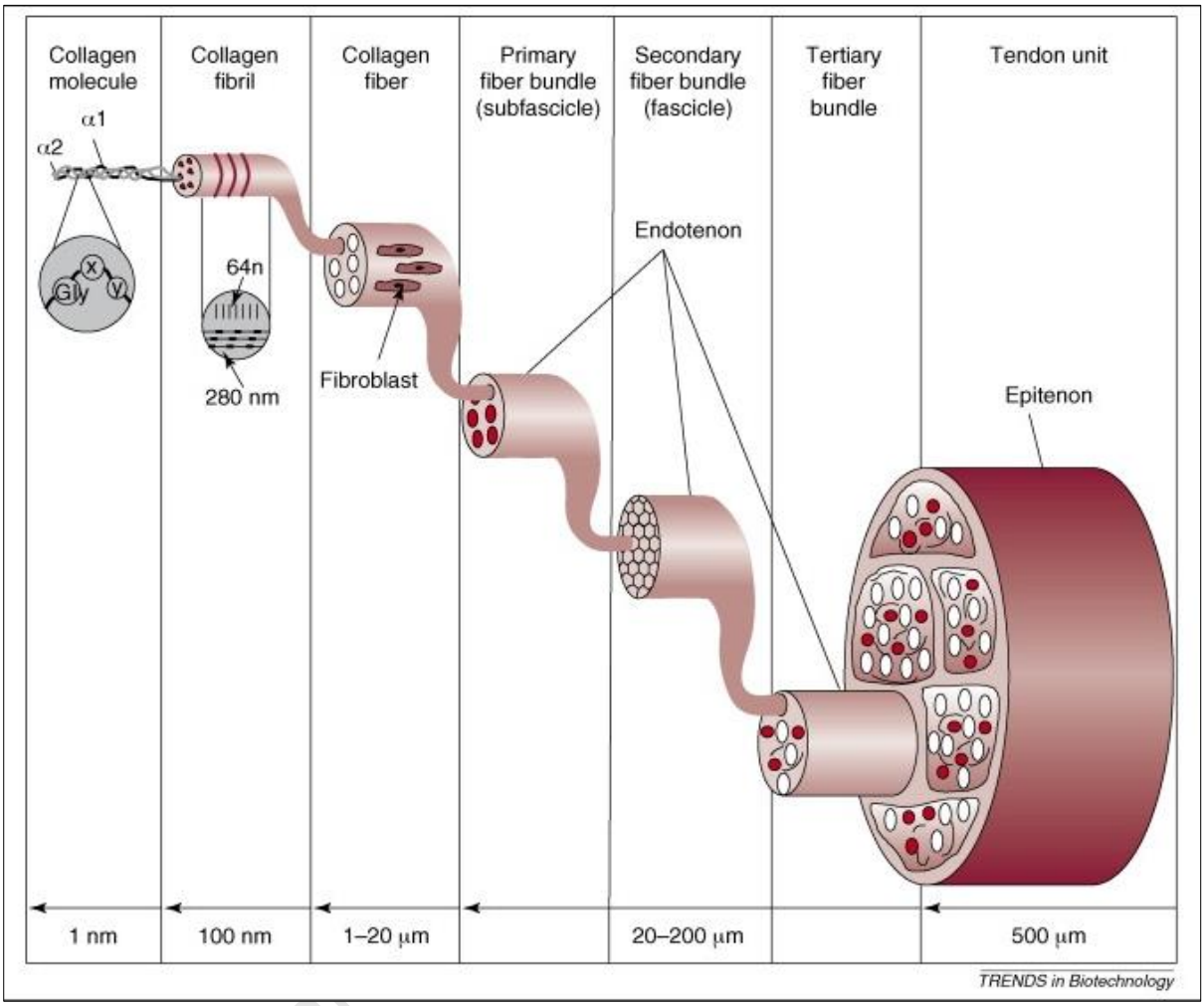

Figure 1.2: The hierarchical structure of tendon.
[Reprinted from Liu et al. (2008)[174] with permission from Elsevier]

1.3. MOLECULAR STRUCTURE OF TENDONS

The major molecular components of the tendon ECM include the collagens, elastin and other glycoproteins, and proteoglycans (Table 1.1). As the focus of this thesis is on genes that code for selected collagens and glycoproteins, the following review will predominantly focus on these molecules.

Table 1.1: Molecular composition of tendon extracellular matrix

MOLECULE		STRUCTURE
COLLAGEN	Type I	Fibrillar
	Type II*	Fibrillar
	Type III	Fibrillar
	Type V	Fibrillar
	Type XI	Fibrillar
	Type IX*	FACIT
	Type XII	FACIT
	Type XIV	FACIT
	Type IV	Forms meshwork
	Type X*	Forms meshwork
	Type VI	Beaded filaments
GLYCOPROTEINS	Elastin	Branched network
	Cartilage oligomeric matrix protein (COMP)	Branched molecule
	Fibrillin	Linear
	Fibronectin	Modular
	Laminin	Modular
	Link protein	Globular
	Tenascin-C	Branched molecule
	Thrombospondins	Modular
PROTEOGLYCANS	Biglycan	Small leucine-rich repeat proteoglycan
	Decorin	Small leucine-rich repeat proteoglycan
	Fibromodulin	Small leucine-rich repeat proteoglycan
	Lumican	Small leucine-rich repeat proteoglycan
	Aggrecan*	Hyalectan
	Versican	Hyalectan

** Predominant in fibrocartilaginous regions of tendon*
[Table modified from Riley et al. (2005)[240]]

University of Cape Town

1.3.1. COLLAGEN

Collagen is the most abundant protein, accounting for approximately 30% of all protein in the human body.[173] In tendons, 60-85% of the dry tissue mass is made up of collagen proteins.[208] A large number of distinct collagen types have been identified in the ECM and the variety in these collagen types allows connective tissue to have a structural diversity and adapt to many stimuli. The focus of this review will be on collagen types that have been identified in tendon (Table 1.1). Collagens are predominantly separated into two groups based on their structure and function, (i) fibrillar collagens which form the ECM scaffolding network and contain rigid, uninterrupted helices, and (ii) non-fibrillar collagens which include, amongst others, the fibril-associated collagens with interrupted triple helices (FACIT).[216,240] Fibrillar collagens are the most abundant, and are further separated into major (types I, II and III) and minor collagens (types V and XI).[32,285] More recently type XXVII collagen was identified as a unique fibrillar collagen that is primarily expressed in cartilage, but also in tendon.[35,114,133,216]

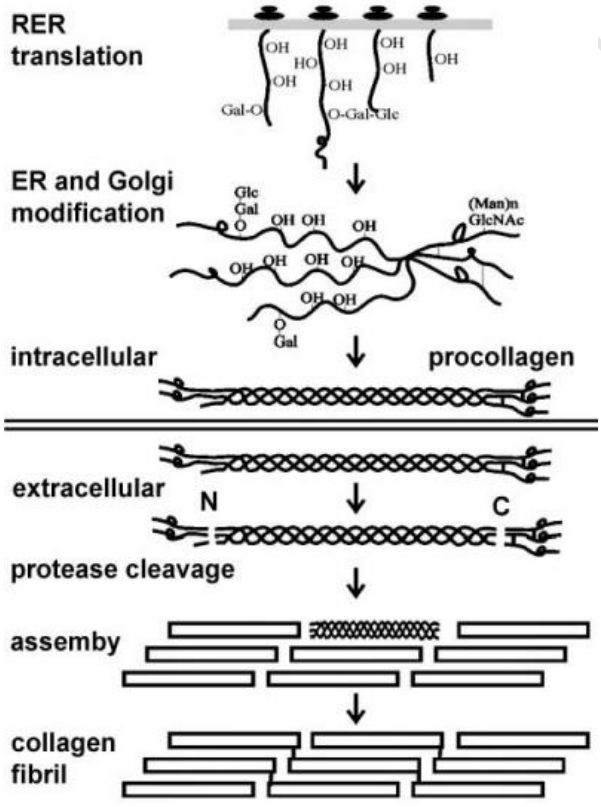

Figure 1.3: Collagen synthesis
[Source: http://php.med.unsw.edu.au]

The synthesis of collagen is extensively reviewed by Banos et al. (2008)[24] and summarised in Figure 1.3. Briefly, fibrillar collagens are formed within the tenocytes where collagen messenger RNA (mRNA) is translated and translocated to the rough endoplasmic reticulum. Typically, the polypeptide α-chains are composed of approximately 1000 amino acids in an uninterrupted repeating triplet which contains a glycyl residue in every third position while the other positions are predominantly occupied by proline and 4-hydroxyproline (Gly-X-Y).[142] These α-chains also contain uncommon hydroxylysine residues.[208] In the rough endoplasmic

University of Cape Town

reticulum, these α-chains undergo post-translational modification which includes glycosylation of some hydroxylysine residues, and hydroxylation of proline and lysine residues. Three polypeptide α-chains are then wound together in a triple helix to form procollagen units. The hydroxyproline residues form hydrogen bonds between the α-chains and contribute to the stability of the procollagen molecule.[208] These procollagen molecules are transported to the Golgi apparatus where they are aggregated into organised, lateral bundles of procollagen molecules. The procollagen triple helices are synthesised containing propeptides at both the amino (NH_2) and carboxyl (COOH) terminals of the triple helical domain. Although the exact process after release from the Golgi apparatus is debated, it is accepted that the COOH- and NH_2-propeptides are post-translationally cleaved and the resulting soluble procollagen bundles are transported into the ECM.[24,142,208] This cleavage exposes the non-collagenous (NC) telopeptide extensions of the triple helix, which contain binding sites for fibrillogenesis and are numbered from the COOH-terminus starting with NC_1.[142] In the ECM, the procollagen bundles aggregate to from collagen fibrils as described in section 1.2.2.[115,208] The hydroxylysine residues are important in establishing electrostatic cross-links between the procollagen molecules to reinforce the fibrillar structure.[208] These cross-links are vital in determining the tensile strength of collagen in that they allow for increased energy absorption and increase resistance to proteolytic activity. A reduction in the number of cross-links results in a weaker collagen fibre.[208]

The major structural domains of collagen proteins are typically highly conserved across species. Mutations within the majority of collagen genes result in serious connective tissue disorders including osteogenesis imperfecta (type I),[216,222] chondrodysplasias (types II and XI)[285] and Ehlers-Danlos syndrome (EDS)(types III and V).[183,184,193,268] This emphasises the importance of the collagen genes and proteins in the determination of a functional connective tissue.

The predominant collagen in tendon is the fibrillar, heterotrimeric type I collagen (two α1(I) chains and one α2(I) chain). Type I collagen is responsible for the hierarchical structure of tendons and determines the mechanical and tensile strength of tendon tissue.[78,115,234] Type II collagen, a homotrimer consisting of three α1(II) chains, is the main structural collagen in

University of Cape Town

cartilage and has been linked to the development of osteoarthritis.[144] Its expression in tendon is limited to the fibrocartilaginous zones.[240]

Type III collagen is also a major fibrillar collagen of tendon and is important in the healing process and during fibrillogenesis.[24,173] It co-localises with, and in some cases may form heterotypic fibrils with type I collagen during tendon development.[24] It is thought that type III collagen regulates the diameter of type I collagen fibrils during development and healing by limiting lateral growth.[24] Type III collagen is laid down early during the proliferative phase of the healing process and rapidly forms disulphide cross-links which stabilise the newly synthesized matrix of the repair site.[5,52,261] These cross-links are easily degraded which allows for rapid growth and remodelling of tissues, however these fibres are thinner and more extensible than type I and its accumulation within tendon tissues during a prolonged healing response therefore weakens the mechanical and tensile strength of the tissue.[52,78] Type III collagen occurs as a homotrimer of three α1(III) chains which are encoded by the *COL3A1* gene on chromosome 2q31-32.[59,268] Mutations within the *COL3A1* gene result in a form of EDS characterised by tissue friability.[268] In particular, mutations near the COOH-terminal result in dilatation of the rough endoplasmic reticulum and the secretion of smaller collagen fibrils, while mutations at the NH_2-terminal result in a more variable fibril diameter.[268]

Type V collagen is a minor fibrillar collagen which is co-expressed with type I collagen.[32] The predominant isoform of type V collagen is a heterotrimer of two α1(V) and one α2(V) chains, however homotrimers of three α3(V) chains also exist.[32,59] These chains are encoded by the *COL5A1* (chr9q34.2-34.3), *COL5A2* (chr2q14-32) and *COL5A3* (chr19p13.2) genes respectively.[59,193,285] The triple helical domain of type V collagen is embedded within the type I collagen molecule with the NH_2-terminal protruding from the fibril (Figure 1.4). Through this interaction it plays a vital role in the regulation of type I collagen fibril diameter.[32,75,193] Recent investigations of the function of type V collagen in tendon fibrillogenesis in a murine model suggest a synergistic role with type XI collagen in fibril nucleation and the regulation of fibrillogenesis during tendon development and growth of mature fibrils.[297]

University of Cape Town

Collagen XXVII is a fibrillar collagen expressed primarily in cartilage and sites of transition from cartilage to bone, but also in foetal organs and skeletal muscle.[35,114,216] The pro-α1 chain of type XXVII collagen has unusual molecular characteristics and a unique expression pattern compared with classical vertebrate fibrillar collagens.[35] As described above, typical fibrillar collagens share a highly conserved COOH-terminal NC domain and a long collagenous domain. They are traditionally classified into two distinct clades based on phylogenetic analyses and these clades differ in the exon arrangement of the relevant collagen genes, the sequence of the COOH-propeptide and the size of the NH_2-terminal NC domain.[35,216,265] Collagen XXVII protein is similar to other minor fibrillar collagens in that the pro-α1(XXVII) chain consists of a large globular NH_2-terminal propeptide and contains overlapping laminin-G (lamG) and thrombospondin NH_2-terminal like (TSPN) domains, as well as a proline/arginine rich (PARP) domain and the characteristic COOH-propeptide.[216] However, the protein differs from previously characterized fibrillar collagens in that the variable domain fuses directly with the major triple helical domain and therefore lacks the classical telopeptide and minor triple helical domain.[35,216] In addition, pro-α1(XXVII) has a shorter triple helical domain and two conserved interruptions in the Gly-X-Y repeat in exons 6 and 42.[35,216] The pro-α1(XXVII) chain is highly conserved in man, mouse and *Fugu* puffer fish, particularly the TSPN domain, collagenous domain and the COOH-propeptide, suggesting little tolerance for variation in these domains.[35] Collagen types III, V and XXVII are discussed further and in more detail in chapters three, four and five of this thesis.

In addition to these fibrillar collagens, other non-fibrillar collagens are also expressed in tendon. These include the FACITs (types IX, XII and XIV), and other types (type IV, VI and X).[240] The FACITs are important in mediating cell-matrix interactions between the type I collagen fibres and cell surfaces.[240] In addition, it has recently been confirmed that the NC_2 domain of FACITs, types IX and XIX collagen in particular, are responsible for α-chain selection, trimerisation and the stabilisation of attached collagen triple helices.[38,39] Further description of the non-fibrillar collagens is beyond the scope of this thesis, but is reviewed by Ricard-Blum et al. (2005).[237]

University of Cape Town

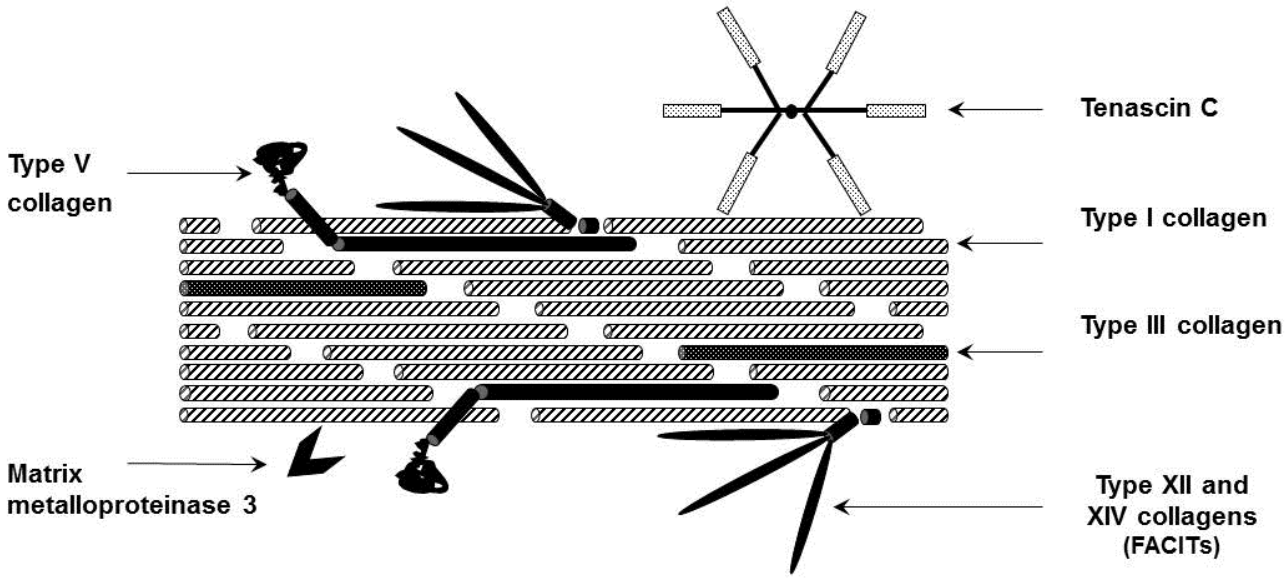

Figure 1.4: Components of a collagen fibril in tendon
[Modified from Collins & Raleigh (2009)[59]]

University of Cape Town

1.3.2. NON-COLLAGENOUS PROTEINS OF THE EXTRACELLULAR MATRIX

Although collagen is the major protein in the ECM, other non-collagenous macromolecules are important components of the ECM. These include glycoproteins, particularly elastin, proteoglycans and matrix enzymes.

1.3.2 (I) GLYCOPROTEINS

Glycoproteins are conjugated proteins with oligosaccharide chains covalently linked to polypeptide side chains. The major glycoprotein in tendon is elastin which makes up approximately 2% of the dry weight of tendon.[208,240] It is a hydrophobic, highly extensible molecule that contributes to the flexibility of the tendon.[208] Other glycoproteins include the thrombospondins (TSP), tenascin-C (TN-C), fibrillin, laminin and fibronectin. They fulfil a variety of functions in the ECM but are predominantly involved in mediating cell-matrix interactions.[240] The thrombospondins are a family of calcium binding glycoproteins with two subfamilies that have distinct functions.[7,36] Group A are the matricellular proteins, TSP-1 and TSP-2, which modulate cell functions and cell-matrix interactions.[36] Group B is made up of the structural proteins TSP-3, TSP-4 and the cartilage oligomeric matrix protein (COMP; TSP-5).[7,36,37] TSP-2 and COMP are described in more detail and investigated in

chapter two of this thesis. TSP-2 is involved in cell-matrix interactions and has an anti-angiogenic role during the healing response in connective tissues.[7,36,37] COMP is another major component of tendon and interacts with type I and type III collagens during matrix assembly and healing.[108,223,240,243,272] TN-C is also discussed in more detail and investigated in chapter three of this thesis. This mechanosensitive glycoprotein is expressed during wound healing and tissue remodelling, particularly in response to high tensile and compressive forces.[129,136,137] The review of other glycoproteins is beyond the scope of this thesis but is reviewed in more detail by Riley et al. (2005).[240]

1.3.2 (II) PROTEOGLYCANS

Proteoglycans are heavily glycosylated proteins which consist of a protein core with GAG side chains. The function and metabolism of proteoglycans in tendon is beyond the scope of this thesis and is expertly reviewed elsewhere.[217,234,309] Briefly, they interact with the fibrillar collagen network and are important in the normal functioning of the tendon ECM. They include the large hyalectan type proteins, aggrecan and versican, and the small leucine-rich repeat proteoglycans (SLRP) decorin, biglycan, fibromodulin, lumican, proteoglycan-4 and keratocan.[234,240] Aggrecan resists compressive and shear forces while versican is involved in cell migration, adhesion and differentiation.[234]. The SLRP's in particular bind to fibrillar collagen and growth-factors and regulate fibrillogenesis.[234,240] Proteoglycans in tendon have a high turnover rate and perturbations in their metabolism have been implicated in tendinopathy.[86,142,234]

1.3.2 (III) MATRIX ENZYMES AND THEIR INHIBITORS

Proteolytic activity in the tendon is essential to maintain homeostasis and to remodel tissue during healing.[125,240] Collagen remodelling in the ECM is mediated by a family of proteases, known as the matrix metalloproteinases (MMP), and the tissue inhibitors of metalloproteinases (TIMP). A large number of MMPs are expressed in tendon tissue and are tightly regulated in both the physiological and pathological remodelling of tendon.[240,306] In addition, a disintegrin and metalloproteinase (ADAM) as well as a disintegrin-like and metalloproteinase with thrombospondin motifs (ADAMTS) play important roles in cell-adhesion, and collagen and proteoglycan turnover in tendon.[306] Changes in expression of both MMPs and TIMPs have been implicated in tendinopathy and it has been suggested that

University of Cape Town

the homeostatic balance between MMPs and TIMPs is important in maintaining the tendon ECM integrity.[240,306] Further review of the role of MMPs and their inhibitors is beyond the scope of this thesis but is reviewed by Birkedal-Hansen et al. (1993).[33]

1.4. ACHILLES TENDON PATHOLOGY

1.4.1. PREVALENCE OF ACHILLES TENDON RUPTURE AND TENDINOPATHY

An increasing emphasis on the benefits of physical activity has led to an increase in the active adult population.[75] This, together with improvements in health care, has further led to an increase in activity in the elderly.[74,75] This increase in participation in physical activity has resulted in an increased exposure to, and incidence of, sporting injuries.[16,140,278] Tendon and ligament injuries are a source of considerable morbidity with an estimated 30-50% of all sporting injuries being tendon injuries.[115,131] The Achilles tendon is particularly susceptible to tendinopathy, which is the most common tendon disorder accounting for 55-65% of Achilles tendon disorders.[132,233,306] In particular, a prospective study of 427 runners reported that 7% of runners developed mid-portion Achilles tendinopathy within one year.[113]

Incidence rates for Achilles tendon ruptures (ATR) and Achilles tendinopathy (AT) within different populations vary according to the definitions and clinical diagnostic criteria used, and may also reflect international differences in participation rates for particular sports.[121] In a 2009 cross-sectional study of 57 725 Dutch patients registered for primary health care, the incidence of adult mid-portion AT in the general population was 2.35 per 1000 registered patients.[68] However, a prospective cohort study of 785 elite Finnish athletes showed that middle- and long-distance runners had a higher risk of developing AT before the age of 45 years when compared to 416 active controls (OR 31.2, 95% CI:13.5-71.8; P<0.001).[159]

There is a particularly large difference in the reported incidence values for ATR (6.0 to 37.3 per 100 000 people).[275] A prospective study of ATR in a Scottish population showed an increase in incidence from 4.7 per 100 000 people in 1982 to 6.0 per 100 000 people in 1994,[180] while a similar study in a Danish county showed an increase from 18.2 per 100 000

University of Cape Town

in 1984 to 37.3 per 100 000 in 1996.[121] A more recent study in Edmonton, Canada, showed an incidence of 8.3 per 100 000 inhabitants over a 5-year period between 1998 and 2002.[275] Notwithstanding the large variations in the reported incidence of ATR, it is clear that the incidence of ATR is increasing and that it presents a large burden of morbidity.

1.4.2. AETIOLOGY

1.4.2 (I) DEFINITIONS

There are a number of conditions that affect the Achilles tendon and surrounding tissues and as knowledge in this field has grown, there has been a lack of uniformity in the terminology used to describe these conditions.[178,288] For the purposes of this thesis, and in-line with current opinion, Achilles tendinopathy is defined as a clinical syndrome characterised by swelling in and around the tendon which is tender to palpation and results in a pain-induced impairment in physical performance.[5,89,128,132,178,234,288,306] This broader term assumes no knowledge of the underlying pathology and encompasses, for example, paratenonitis, which is characterised by inflammation of the paratenon, and tendinosis which is characterised by degeneration of the tendon without histological signs of inflammation.[233,250,288] The participants included in this thesis were diagnosed with chronic overuse AT that, in most cases, was confirmed to be a degenerative tendinosis by ultrasound. These participants will be described in more detail in chapter two.

1.4.2 (II) HISTOPATHOLOGY AND CLINICAL CHARACTERISTICS

Although the exact aetiology of tendinopathy is still unclear, there are several histopathological changes in the tendon and clinical criteria that are well recognised in the diagnosis of AT (Table 1.2).[5,17,69,89,128,306] Several types of degeneration have been proposed to take place in tendon, adjacent to healthy tissue.[128] These include: (i) hypoxic degeneration, (ii) mucoid degeneration, (iii) fibrinoid degeneration, (iv) hyaline degeneration, (v) fatty degeneration, (vi) calcification and (vii) fibrocartilaginous metaplasia. On a macroscopic level, AT is characterised by loose and poorly defined intratendinous regions with a focal loss of fibrillar tendon structure. The tendon loses its glistening white appearance and becomes grey and amorphous with diffuse, nodular or fusiform thickening. There are signs of neovascularisation with the proliferation of capillaries and arterioles.

University of Cape Town

Under light microscopy, there is an increase in ground substance (mucoid degeneration) and this increase in proteoglycans and GAGs, with the associated increase in bound water, results in the hypoechoic regions seen during ultrasound examination. There is disruption and separation of the collagen fibres which become disordered and lose their hierarchical structure. Collagen fibres become angulated and vary in their diameter and orientation. The characteristic crimping of the collagen fibres becomes irregular with increased waviness. Calcification and intratendinous lipid accumulation may accompany the degradation of the collagen fibres.[128]

There is large variation in the cellular density across the degenerated tendon with areas of low cell density and areas with large numbers of metabolically active tenocytes. Tenocytes lose their spindle shape and parallel alignment and have changes characteristic of hypoxia with round nuclei, lipid vacuoles, enlarged lysosomes and degranulated endoplasmic reticulum. There is also evidence for increased cellular apoptosis in AT.[306]

Although chronic AT was traditionally, and in some clinical practices still is, diagnosed as Achilles tendinitis, the role of inflammation in the development of AT has remained elusive. Many studies have found no signs of inflammation in degenerate tendons and, particularly, in ruptured tendons.[5,89,128,146,261,306] However, both ruptures and tendinopathy in human tendons are generally only investigated after the degeneration becomes symptomatic.[146] In addition, animal models of tendinopathy may not fully replicate the early stages of chronic overuse injuries.[5] The possibility of an acute inflammatory response can therefore not be excluded. There is, in fact, some evidence to suggest this acute inflammatory response in rabbit Achilles tendon and equine superficial digital flexor tendon (SDFT).[86,189,300] In addition, peritendinous injections of prostaglandin E_1 and inflammatory cytokines resulted in a histopathological profile similar to that of mild tendinopathy.[274,276] As originally hypothesized by Jarvinen et al. in 1997,[128] the current evidence suggests that there is an early, acute inflammatory response to overuse which may either be superseded by degenerative changes, or proceed in parallel with these degenerative changes in AT.[5,233,261] Both inflammatory and degenerative changes can co-exist in contiguous areas of tendons, and a 2009 review of the pathogenesis of tendinopathies by Abate et al. (2009)[5] concluded

University of Cape Town

by supporting the "entangled roles of inflammation and subsequent degeneration within tendons".

Table 1.2: Histopathological changes of tendon tissue in Achilles tendinopathy

	HEALTHY TENDON	DEGENERATIVE TENDON
Macroscopic:	• Glistening white in colour • Firm, fibroelastic texture	• Grey-brown in colour • Fragile and loose texture • Fusiform, nodular or diffuse thickening • Fibrocartilaginous and bony metaplasia • Increased nerve outgrowths • Neovascularisation
Microscopic: Collagen Fibres	• Densely packed, organised, parallel collagen fibres which are uniform in diameter & orientation	• Disordered collagen bundles with angulated collagen fibres which vary in diameter and orientation • Large vacuoles • Increased ground substance with proteoglycans and GAGs (mucoid degeneration) • Calcification • Intratendinous lipid accumulation
Microscopic: Cellular	• Tenocytes with spindle shaped nuclei, in parallel alignment with collagen fibres	• Tenocytes with round nuclei • Loss of parallel alignment • Hypoxic changes in tenocytes (lipid vacuoles, enlarged lysosomes, degranulated endoplasmic reticulum) • Increased apoptosis
Ultrasound:	• parallel hyperechoic/bright white lines	• localised widening of tendon • localised hypoechoic areas • irregular fibre structure • neovascularisation

[Compiled from [5,17,69,78,89,125,128,239,306]]

University of Cape Town

1.4.2 (III) MODELS FOR THE PATHOGENESIS OF ACHILLES TENDINOPATHY

The exact aetiology of AT has not been elucidated and research in this area is hampered in that human tendons are generally only studied after they become symptomatic and not during the early stages of the condition.[5] Previously, pain in the tendon was diagnosed as tendinitis and treated primarily as an inflammatory condition.[232] The 1990's however saw a new perspective on tendon disorders as investigators reported evidence which suggested tendinopathy was not primarily inflammatory but degenerative in nature.[232] Degenerative changes in tendons are predominantly considered a result of aging, dysregulated cell-signalling processes and chronic mechanical stress.[128] Several models and hypotheses have emerged to explain the development of chronic tendinopathy.

(a) Mechanical Overuse Theory

It is generally accepted that mechanical overuse of tendon is the primary inciting event in AT.[5,89,306] Although tendons can withstand loads much larger than those required in day to day functioning, repetitive exposure to high tensile loads can result in excessive strain on the tendon tissues.[115,132,233,239] Stress is defined as the force or load applied to a tendon, and strain is defined as the displacement of tissue resulting from that force. The relationship between these can be plotted as a stress-strain curve (Figure 1.5).[132,233,239] When stress is applied to a tendon, the crimp stretches out resulting in the lag or "toe" region of the stress-strain curve. As the crimp straightens out, collagen fibres experience strain in a relatively linear relationship to the load. Under loads resulting in less than 4% strain tendons behave elastically and will return to their original length once the load is removed. Loads resulting in more than 4% strain may exceed the tendons' physiological extensibility and the load is then directly absorbed by the collagen fibres resulting in microtears and damage. In particular, the collagen fibres start to slide past one another which disrupts the collagen cross-links.[5,89] It should be noted that recently strain loads of up to 6% have been suggested to be physiological.[233] The chronic accumulation of this damage over time has led to the suggestion that tendinopathy may be more degenerative than inflammatory in nature.[233] Further pathological strain will overwhelm the reparative processes, and/or the maximum tensile load of the tendon, and will lead to complete failure of the tendon.[132,145,239,261]

University of Cape Town

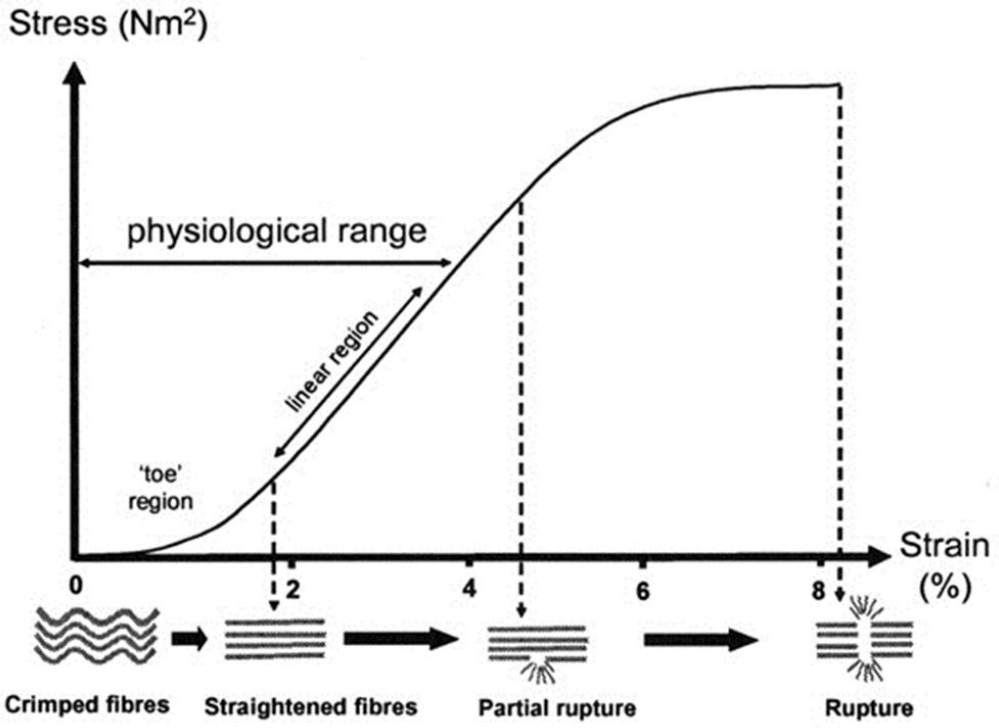

Figure 1.5: The relationship between stress and strain in tendons
[Reprinted from Riley et al. (2004)[239] with permission from Oxford University Press]

This mechanical overload model supports the observation that most ruptures are preceded by tendon degeneration[78] and the higher incidence of tendinopathy with age and exercise, however it does not explain why certain areas of tendons are more susceptible to tendinopathy than others, nor spontaneous ruptures in patients with no history of exercise.[233] The mechanical overuse model assumes that the structural micro-damage of the collagen fibrils is the primary cause of tendinopathy and that the cells of the tendon respond to this damage in an attempt at repair only after the micro-damage has occurred. This simple catastrophic model therefore implies that the cellular component of the ECM does not contribute to the mechanoresponse in the tendon during the early stages of tendinopathy to prevent micro-damage to the collagen fibril. The "iceberg" theory, an integrated model described in 1.4.2.(iii)(h), is based on this assumption of micro-damage during the intitial stages of tendinopathy with a cellular response that occurs as a consequence of this micro-damage. However, it is more likely that the tendon tissue, particularly the tenoctyes, responds to mechanical loading earlier than this through the mechanotransduciton of force. Mechanotransduction is the process of converting

University of Cape

physiological mechanical stimuli to biochemical responses and is comprised of three steps: (i) mechanocoupling, (ii) cell to cell communication and (iii) an effector response.[151] Briefly, mechanical loading causes the deformation of cells of the tendon as they experience tensile force, and these mechanical signals are transformed into chemical signals both within and amongst cells of the tendon. This chemical signal is transferred to other cells through cell-signalling pathways and results in the distribution of the effects of the mechanical loading stimulus to cells which did not directly receive the stimulus. Stimulated cells then interact with the ECM to effect repair and remodelling in adaptation and/or healing.[151]

(b) *Vascular insufficiency theory*

The vascular insufficiency theory acknowledges that tendons have a cellular component, are metabolically active and therefore require an adequate blood supply. As the Achilles tendon has been shown to have a poor blood supply, it has been suggested that this is a major factor which hampers healing and results in a weaker tendon.[9,233] In addition, blood flow in the tendon is further decreased during exercise and certain areas are therefore exposed to hypoxia. The subsequent hyperemia at cessation of exercise may result in reperfusion injury.[19,145] It has further been suggested that this reperfusion injury may enhance the production of reactive oxygen species and lead to oxidative damage in the tendon.[145,176] Contrary to this, there is increasing evidence that neovascularisation is a contributing factor to AT. In particular, a prospective trial of asymptomatic runners using power Doppler ultrasound found that the presence of intratendinous microvessels, indicative of neovascularisation, was a significant predictor of developing mid-portion AT within a year (OR 6.9, 95% CI:2.6-18.8; P=0.0001).[113] This neovascularisation may however occur in response to repeated hypoxic episodes. The vascular insufficiency theory may explain why the mid-portion of the Achilles tendon, which is proposed to be hypovascular,[233] may be more susceptible to injury. However, it should be noted that there is also evidence which suggests that blood flow is uniform along the length of the Achilles tendon.[19] In addition, the limited vascular supply results in limited heat dissipation during exercise,[5,31] and localised hyperthermia may therefore be the factor contributing to the development of AT.[233] This model is not sufficient to explain all the histological and clinical features of tendinopathy, however it does suggest that secondary factors such as hypoxia, hyperthermia and oxidative stress might contribute to the pathogenesis of tendinopathy.

University of Cape Town

(c) Failed healing theory

The failed healing theory is an extension of the mechanical overuse theory and proposes that overuse induced micro-damage may not stimulate the normal three-phase healing process – inflammation, proliferation, maturation - in tendons.[61] Compromised healing then results in a pathological tendon with decreased tendon strength which is less able to respond to load.[61] Fu et al. (2010)[91] recently proposed a three-stage process for the development of tendinopathy which centres around failed healing. Briefly, tendinopathy is initiated with an insult of various origin which results in damage to the collagen network. A healing response is activated but fails to repair the tissue adequately, possibly due to unfavourable environments, genetic predisposition and/or hormonal and pharmaceutical factors. The healing response is not only inadequate, but may also be incorrectly diverted to an abnormal pathway. Progressive degeneration of the collagen network will lead to the final stage which is the clinical presentation of either rupture or pain. As in the mechanical overuse theory, this catastrophic model of tendinopathy assumes that the metabolically active cellular component of tendons only responds after damage has occurred.

A number of theory's (described below) take into account the contribution of the cellular component, and the secretion or release of chemoactive substances from these cells, in the development of tendinopathy.

(d) Inflammatory theory

Although the term "tendinitis" has now been widely abandoned, there is emerging evidence that chronic tendinopathy is not completely non-inflammatory.[5,232,252] It is likely that much of the confusion around the role of inflammation in tendinopathy arises as a result of differences in the understanding of the term "inflammation". As inflammatory markers and mediators are not only involved in normal healing responses but also in several cell-signalling pathways, this thesis refers to the up- and down-regulation of these inflammatory mediators in tendinopathy as dysregulated cell-signalling. Although a traditional cellular immune response is not likely to be the primary underlying mechanism of tendinopathy, several inflammatory mediators in cell-signalling pathways are clearly dysregulated in tendinopathy.[232,252] For example, cytokines are produced not only by inflammatory cells but also by tenocytes, and peritendinous injections of prostaglandin E_1 and cytokines have been

University of Cape Town

shown to result in a histopathological profile that is similar to mild tendinopathy.[252,274,276] Interestingly, these inflammatory markers are increased in response to exercise.[89] A recent study has shown an association of genetic variation within several components of the inflammatory pathway and AT.[257] Allele combinations of five polymorphisms within the genes encoding interleukin-1β (IL-1β), interleukin-6 (IL-6) and the receptor antagonist of IL-1β, interleukin-1ra (IL-1ra), were strongly associated with AT and therefore implicated this pathway in the pathogenesis of AT.[257]

(e) Apoptosis theory

Cytokine mediated apoptosis of tenocytes is normal in the healing process of tendon following overload, however excessive tenocyte apoptosis may disrupt the homeostasis of the ECM.[205] High doses of cyclical strain have been shown to induce genes in both the oxidative stress and cartilage production pathways which ultimately leads to apoptosis, a more cartilaginous matrix and the loss of matrix integrity.[253,306] In addition, an independent study found high levels of apoptosis in both non-insertional AT and rotator cuff tendinopathy.[219,311] Recently, polymorphisms within the gene coding for caspase-8, which is involved in tenocyte apoptosis, were found to be associated with AT.[205]

(f) Neuropeptide theory

This theory proposes that chronic overuse results in excessive neural stimulation with a subsequent degranulation of mast cells and release of neurotransmitters such as substance-P and glutamate.[233,239,240] Increased levels of both substance-P and glutamate have been found in AT.[11,240] Furthermore, recent evidence has shown that loading results in increased levels of endogenous substance-P preceding any tissue changes characteristic of tendinosis, and this increase in substance-P may be responsible for tenocyte proliferation.[21,22] In addition, gene expression analysis has shown that loading induced increases in substance-P in tendons resulted in increased expression of *COL3A1* and *MMP3* mRNA as well as increased collagen remodelling.[84,232] Interestingly, substance-P has also been implicated as a pro-inflammatory mediator.[232] In summary, there is increasing evidence to support both a central and local neurogenic pathway in the aetiology of AT.

University of Cape Town

Although the inflammatory, apoptosis and neuropeptide theories take into account an initial mechanical stimulus as well as the effects of this stimulus on the metabolic and cellular components, it is not likely that any particular one of these cell-signalling pathways is the solo primary cause of tendinopathy. It is more likely to be a combination of dysregulation in a number of pathways involved in cell-signalling and the healing response.

(g) Pain in tendinopathy

The traditional model of pain in tendinopathy is that (i) inflammation and (ii) separation of the collagen fibres are the causes of pain.[89,150] However, inflammation is unlikely to be the cause of pain as inflammatory cells and markers are not always found in tendon tissue samples from patients experiencing chronic tendon pain.[11,149] In addition, there is incongruity between collagen integrity and chronic pain.[150] For example, there is no pain in donor sites during anterior cruciate ligament (ACL) reconstruction despite the excision of collagen.[150] Khan et al. (2000)[150] proposed a biochemical hypothesis for pain in patellar tendinopathy that attributed pain to biochemical irritants, such as glutamate and chondroitin sulphate, acting on nociceptors. This is supported by the observation of higher glutamate levels in Achilles tendons with painful nodes,[11] as well as recent findings that glutamate receptors are up-regulated and activated in tendinopathy.[251]

(h) Integrative models

While many models for the pathogenesis of tendinopathy have been proposed, they are not necessarily mutually exclusive. It is highly likely that several mechanisms act synergistically in the progression of tendinopathy (Figure 1.6). Mechanical loading results in several stimuli which are detected and mechanotransduced by tenocytes. The thresholds at which each of these stimuli activate either an appropriate or pathological adaptation and healing responses is individual and probably influenced by intrinsic risk factors such as biomechanics and genetic variation. An appropriate healing response will lead to repair, healing and adaptation. An inappropriate or failed healing response is characterised by dysregulated cell-signalling, failed adaptation, degradation of the matrix and eventually injury. The effectiveness of the healing response may also be influenced by both intrinsic and extrinsic risk factors.

University of Cape Town

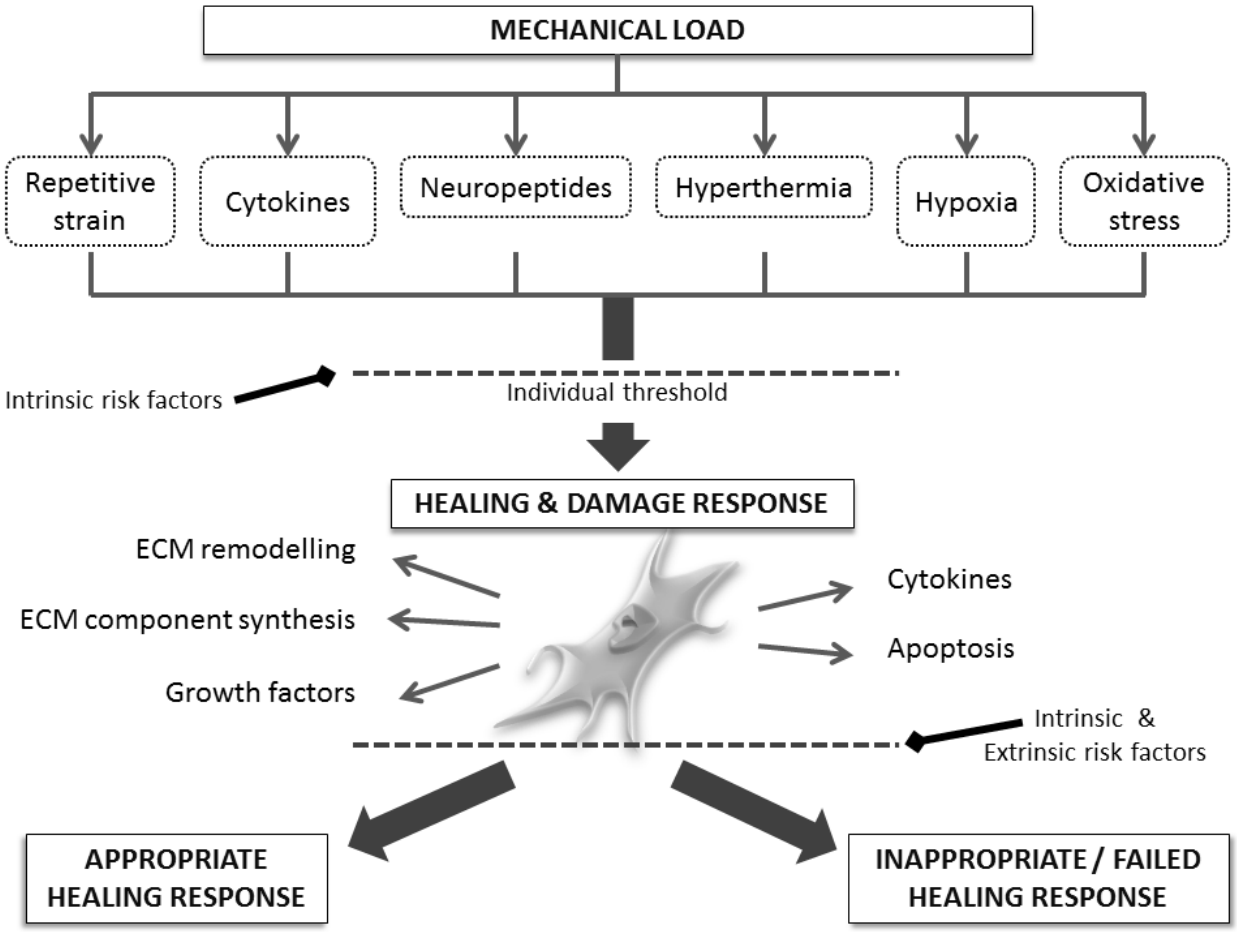

Figure 1.6: Response to mechanical loading in tendon

The "iceberg theory" is an integrative model first proposed by Fredberg et al. (2008)[89] and supported by Abate et al. (2009)[5] (Figure 1.7). This model suggests that AT occurs as a continuum from physiological tendon to an overt clinical presentation of AT.[5] Again, mechanical overuse is the primary inciting factor and leads to disruption of the collagen fibres and the simultaneous activation of healing and degenerative pathways. If the conditions for optimal healing – adequate blood supply, adequate recovery time and absence of further overloading – are not met, the healing mechanisms are overwhelmed and fail. Micro-damage accumulates in the tendon and the "pathogenetic cascade" of inflammation, degradation, neovascularisation and nerve proliferation is initiated.[5] The tendon is initially asymptomatic but reaches a threshold of neovascularisation and nerve

proliferation at which glutamate levels are increased and the patient experiences symptomatic pain. This pain is, therefore the "tip of the iceberg".[5,89] A short period of rehabilitation may allow the extent of damage to recede below the symptomatic threshold, but further loading before adequate healing will lead to relapse of symptoms.[5,89] The thresholds for damage, repair and pain vary between individuals and are influenced by various intrinsic and extrinsic risk factors which results in a very individual susceptibility to the progression of AT.

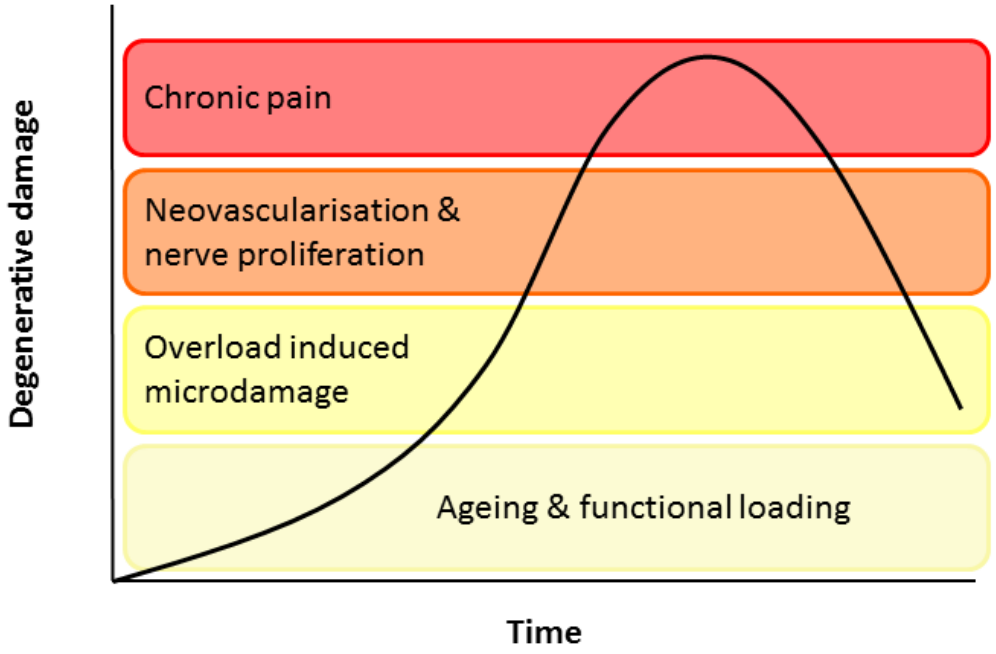

Figure 1.7: The "Iceberg" model of the pathogenesis of tendinopathy
[Compiled with modifications from Fredberg et al. (2008)[89] and Abate et al. (2009)[5]]

Another integrative model proposed by Cook et al. (2009)[62] suggests that tendinopathy is a continuum with different tendons, even different areas within one tendon, progressing back and forth between stages in response to treatment and interactions with various risk factors (Figure 1.8). Briefly, when healthy tendon is subjected to acute tensile or compressive overload it initiates short-term adaptations that result in thickening of the tendon and increased stiffness. Reactive tendinopathy is a non-inflammatory, proliferative stage characterised by an influx of large proteoglycans and an increase in bound water. However,

the integrity of the collagen network is maintained and there are no neurovascular changes. Tendon properties may revert to normal if load is reduced and with sufficient recovery time. Further pathological loading leads to tendon disrepair, an attempt at healing characterised by increased matrix breakdown, an increased number of cells, increased protein production which results in separation of the collagen fibres and disorganisation of the matrix, as well as increased vascularity and neuronal ingrowth. This stage has limited reversibility with load management and other treatment modalities. At the furthest end of the continuum is degenerative tendinopathy in which the tendon has very little capacity to recover. It is characterised by increased apoptosis, areas of acellularity and a vascularised, disordered ECM filled with matrix breakdown products. It is becoming increasingly evident that the balance between matrix degradation and synthesis is disrupted in tendinopathy,[64] and this model best captures the interplay between physiological adaptation of tendon to unaccustomed load and the progressive degeneration of tendon tissue as a result of the dysregulation of cell-signalling pathways and ECM homeostasis.

Cape Town

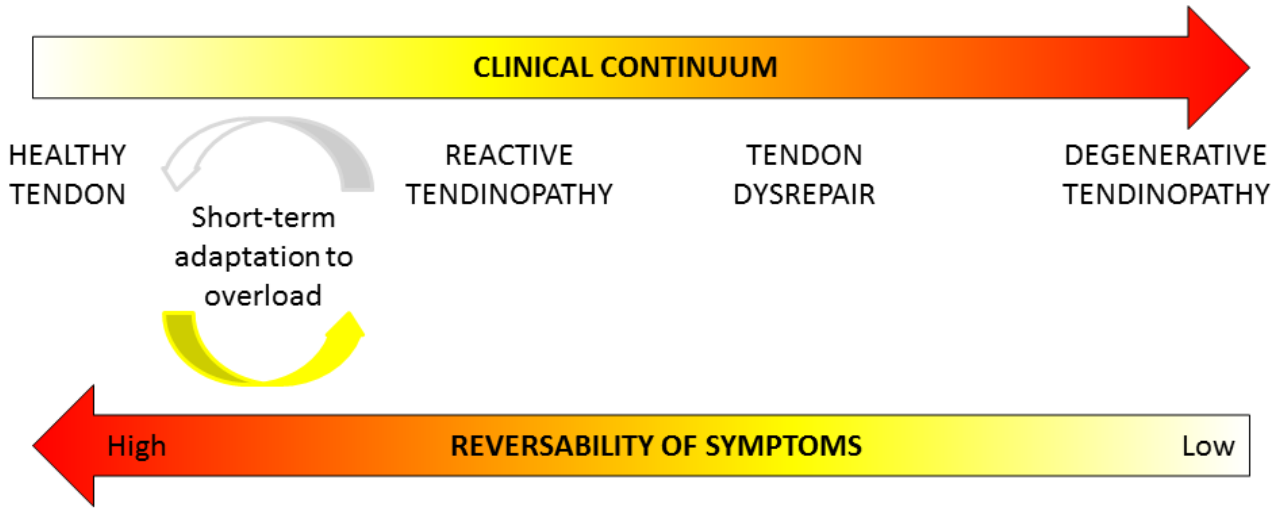

Figure 1.8: The tendinopathy continuum as described by Cook et al. (2009)[62]

1.4.3. GENE EXPRESSION IN TENDINOPATHY

Ireland et al. (2001)[125] reported a number of gene expression changes in chronic human AT. The expression of a large number of genes was up-regulated more than two-fold in degenerate tissue samples taken from patients undergoing surgery for Achilles tendon

disorders compared to normal cadaver specimens. These included, amongst others, collagen types I and III, tenascin-C, thrombospondin-2 and biglycan, as well as MMP-2, MMP-9 and MMP-14. MMP-3 mRNA was down-regulated in degenerate samples. As proteoglycans are one of MMP-3's many substrates, decreased MMP-3 activity may account for the increase in proteoglycans seen in AT. In addition, proteoglycan expression decreases with age in normal tendon, however in chronic tendinopathy there is increased expression of proteoglycans, particularly aggrecan and biglycan.[64] Aggrecan has been shown to be expressed at a six-fold higher level than versican, an expression ratio opposite to that seen in normal tendon.[64] These observations support the increasing evidence that suggests the balance between ECM synthesis and breakdown is disturbed during remodelling in tendinopathy.[64,125] Interestingly, there was no increase in expression of several markers for inflammatory cells (lymphocytes, monocytes and granulocytes), nor for other markers of cytokine mediated inflammation. However, the degenerate tissue samples examined in this study were taken from tendons in a very late stage of chronic tendinopathy.

Jelinsky et al. (2011)[133] recently investigated gene expression in human tendinopathy and found 983 mRNA transcripts, across many different biological pathways, to be differentially regulated in tendinopathy. Of particular interest for this thesis was the significant increase in expression of the *TNC* (5.3 fold, P<0.001), *COL3A1* (2.0 fold, P<0.05), *COL5A1* (2.0 fold, P<0.05), *COL5A2* (2.1 fold, P<0.005), *COL5A3* (2.0 fold, P<0.005) and *COL27A1* (2.5 fold, P<0.005) genes. Whilst most MMPs were differentially up-regulated in diseased tendons, MMP-3 was again shown to be down-regulated in tendinopathy (0.3 fold, P=0.055). In addition, there was some evidence for the differential regulation of inflammatory cytokines and the IL-6 pathway in tendinopathy, although the inflammatory pathway as a whole was not implicated. It should again be noted that the tendon samples investigated in this study were from tendons with late stage disease. Similar results were reported by Pingel et al. (2012),[221] who observed an increase in the expression of collagen types I and III, as well as tenascin-C, in tendinopathic areas compared to a healthy area of the same Achilles tendon. There were, again, no observed changes in the expression of inflammatory markers in the tendinopathic areas of these tendons with late stage AT.

University of Cape Town

In summary, gene expression analyses consistently show increased expression of type I, type III and type V collagen in tendinopathy, despite a decrease in total collagen content. In addition, the ratio of type III to type I collagen is also increased.[306] There is increased expression of several proteoglycans and most MMPs, except MMP-3 which is down-regulated in tendinopathy. There are no changes in expression of cytokine mediated inflammatory markers in late stage tendinopathy.

1.5. RISK FACTORS FOR ACHILLES TENDON PATHOLOGY

Although, the exact aetiology of tendon injuries is not well defined it is well accepted that tendinopathy is a complex condition influenced by several extrinsic and intrinsic factors.[255] Intrinsic risk factors, which influence risk from within, together with exposure to extrinsic risk factors, which influence risk by acting outside the body, determine an individual's susceptibility to developing disease.[59] There are varying levels of evidence that these factors influence an individual's risk of developing AT and the evidence supporting the inclusion of various traits and factors as recognised risk factors for AT is briefly reviewed below.

For the purposes of this thesis, level of evidence is evaluated using a modified version of the guidelines for prognostic studies outlined by Slobogean et al. (2012)[267] Level of evidence was described as strong, limited or weak. Strong evidence includes level I studies: well conducted prospective cohort studies which control for bias and confounding factors, and systematic reviews of level I studies. Limited evidence includes level II and III studies: retrospective studies, cross-sectional studies, prospective cohort studies with issues of bias and confounding, case-control studies, meta-analyses and systematic reviews of level II studies. Weak evidence includes level IV and V evidence: case reports, case series and published expert opinion.

Level of certainty is a more subjective evaluation based on the guidelines described by the U.S. preventative services task force (2007).[248] Level of certainty is described as high, moderate or low. High certainty is assumed when the evidence includes consistent results from well-conducted and well-designed studies in appropriate populations. Where level of certainty is high, future studies are not likely to strongly affect the conclusions of this

University of Cape Town

review. Moderate certainty is assumed where confidence in the available evidence is limited by the number, size or quality of studies, the repeatability of findings, and the use of inappropriate populations. Low certainty is assumed where the evidence is limited by the number and size of studies, flawed study design and inconsistent findings. Where the evidence is deemed to be of moderate or low certainty, future studies may alter the conclusions of this review.

The level of evidence and level of certainty for each risk factor is summarised schematically at the end of this chapter (Figure 1.10). The evidence reviewed below may relate to AT, ATR or other tendinopathies and tendon injuries, however the level of evidence and certainty is evaluated based only on the evidence for being a risk factor for AT.

1.5.1. EXTRINSIC RISK FACTORS

1.5.1 (I) CYCLIC LOADING AND SPORT

Physical training and exercise are important to tendon health as physiological loading induces short term adaptations which increase the stiffness, tensile strength and maximum static strength of tendons.[42,154,208] There is evidence to suggest that increased activity leads to increases in type I collagen turnover with net formation, and that prolonged activity will therefore result in increases in tendon tissue and strength.[154] Conversely, immobilisation leads to decreases in collagen synthesis and increased MMP activity which results in collagen degeneration and decreases in tensile strength.[154,208] These effects result from the activation of several signalling pathways by the mechanotransduction of force to a biochemical stimulus through integrins, growth factors, inflammatory and vasoactive substances, and subsequently changes in gene expression.[154,294]

Although tensile load has anabolic effects on tendon tissue, it also has catabolic effects.[62] In particular, load-induced strain is thought to be the chief mechanical factor in initiating the accumulation of tendon damage and, eventually, injury.[304] In a 2008 systematic review of the available knowledge of tendinopathy, Xu et al. (2008)[306] suggested that repetitive loading induced expression of genes in both the cartilage gene pathway and the oxidative stress pathway. This leads to apoptosis of tendon cells, a more cartilaginous tendon matrix and finally tendinopathy. It is also clear that both the volume and frequency of load are

University of Cape Town

important factors in initiating tendinopathy.[62,132,239] A review of the effects of exercise on tendon properties suggests that biochemical, biomechanical and structural tendon properties respond not only to the absolute load, and its resultant strain, but more so to the number of loading cycles.[42] In addition, cyclically loaded Achilles tendons were found to reach failure significantly faster than would be predicted by time dependant damage alone.[304] Logically then exposure to elements of training and exercise that influence either volume or frequency of loading, such as intensity, running speed, number of repetitions and hill running, will influence risk of developing AT.[132,159,239,306]

Certain sporting codes have also been shown to result in a high prevalence of Achilles tendon injuries. In a retrospective study of ATR in a Danish population, 46.3% and 23.3% of ruptures occurred while playing badminton and soccer (football) respectively, with handball, volleyball and athletics contributing to 24.6% of ruptures.[121] In a Canadian population, 75% of ruptures occurred during sporting activities with soccer (11.8%) and volleyball (9.6%) having the highest prevalence.[275] A 1987 cross-sectional study of 180 runners with injuries reported that 65% of these runners were long-distance runners, and that Achilles "tendinitis" was the third most prevalent diagnosis (11%).[127] A more recent retrospective cohort study of elite male Finnish athletes showed that 42% of middle- and long-distance runners developed AT before the age of 45 years compared to 3% in an active control group (OR 31.2, 95% CI:13.5-71.8; P<0.001), while sprinters sustained ATR more often than active controls (OR 14.9, 95% CI:4.35-50.7; P<0.001).[159] Differences in the sporting codes most associated with AT and ATR may reflect differences in national sporting cultures and preferences but, in general, sports which involve fast changes in direction and bursts of power are associated with tendon ruptures while sports such as running, which involve repetitive actions, are associated with tendinopathy.

In summary, there is a large volume of research at the tissue and molecular level describing the effects of exercise and physical activity on tendon tissue properties. It is clear that long-term cyclic loading affects tendon properties. However, many of these studies involve *in vitro* studies of tendon specimens from acutely overloaded animal models or human cadavers.[154] These studies provide valuable information of the basic physiology, or pathophysiology, of tendons but have limited applicability to what happens *in vivo* during

University of Cape Town

exercise.[154] To the authors knowledge, there are no prospective cohort studies specifically designed to investigate repetitive and/or excessive loading as a risk factor for AT. However, exercise, particularly running, is often used as an inciting factor in prospective studies of other potential risk factors for AT. Therefore, although there is only limited evidence directly associating exercise, specifically middle- and long-distance running, as risk factors for AT, it is of high certainty.

1.5.1 (II) TRAINING ERRORS

Training errors are repeatedly listed as risk factors for AT.[132,145,239] In a 1978 observational study of 180 runners with injuries by James et al.,[127] 60% of injuries were associated with training errors. These included excessive mileage, high intensity workouts, rapid changes in training routine, hill running and running on hard surfaces. As the methodology behind this study is not outlined, it is hard to draw conclusions from these findings. There is a scarcity of further evidence to support these claims and, therefore, the level of evidence for training errors as an independent risk factor for AT is weak and of a low certainty. It is more likely that these factors influence the volume and frequency of loading (reviewed in 1.5.1(i)), or are accompanied by biomechanical risk factors (reviewed in 1.5.2(vii)).

1.5.1 (III) TRAINING SURFACES

Training surface is also often cited as a risk factor for AT.[132,239] In particular, there is concern over the effect on injury risk of artificial playing surfaces compared to natural turf i.e. grass. A prospective study of two cohorts over two football seasons showed no difference in the incidence of muscle/tendon injuries sustained during either training or match play on artificial turf versus grass surfaces.[92,93] Furthermore, a recent systematic review of injury rates in all football codes (rugby union, soccer, American football) provided strong evidence that the injury rate on modern artificial turf was comparable to that on natural surfaces.[301] The only possible exception to this was in the incidence of ankle injury which may be higher when playing on artificial turf. It should however be noted that these studies do not specifically investigate AT, and there is some evidence to suggest that playing football on artificial turf influences plantar loading across the foot.[85]

University of Cape Town

There is, however, more evidence to support running surface as a risk factor for injury. Two independent cross-over studies on recreational runners have reported differences in maximum in-shoe plantar pressure on different running surfaces. Tessutti et al. (2012)[278] reported 16% less peak pressure in the rear foot and lateral forefoot when running on grass versus asphalt, concrete and rubber (P<0.005). Wang et al. (2012)[295] reported an increase in maximum plantar pressure when running on concrete compared to natural grass (95% CI:8.5-91.7 kPa; P=0.016). Although there were no differences in maximum plantar force, the contact time on grass was significantly longer (95% CI:4.4-32.6 ms; P=0.017). Similar results were found in another study by the same group comparing treadmill running to running on concrete and grass.[117] Total foot maximum plantar pressure was again found to be higher when running on concrete compared to grass (95% CI:22.97-93.08 kPa; P<0.0017), with no difference in maximum plantar forces between grass and concrete. In addition, the maximum plantar pressure across the total foot was found to be lower for treadmill running compared to running on concrete (95% CI:29.11-116.42 kPa; P<0.0017) and the total foot maximum plantar force was found to be decreased during treadmill running compared to both grass and concrete (P<0.0017).

In summary, there is strong evidence of high certainty that different running and playing surfaces affect pressure loads across the plantar foot, however there is also strong evidence of high certainty that there is no difference in injury rates when playing on artificial turf compared to grass. To the authors' knowledge, there is no research comparing incidence of AT when running on harder surfaces compared to grass or treadmills. Training surface can therefore not be included as a risk factor for AT at this stage.

University of Cape Town

1.5.1 (IV) ENVIRONMENTAL CONDITIONS

Exercise in extreme heat or cold, humidity, strong wind and at altitude have all been listed as risk factors for developing AT.[132,197,239] However, with the exception of cold weather training, little evidence exists to support these as independent risk factors, and they may actually exert their effect on risk of developing AT via the associated alterations in exercise intensity. A prospective study of 1500 military recruits reported a significant effect of training season on the incidence of AT, with more recruits developing AT in winter (P=0.001).[195] As the majority of cases in this AT group were more specifically diagnosed

with paratenonitis, the authors hypothesized that cold weather adversely affected the viscosity of the muccopolysaccharide lubrication of the paratenon during exercise. There is therefore strong evidence to support cold weather training as a risk factor for AT however, as there is only one study reporting this association, the level of certainty is low.

1.5.1 (V) FOOTWEAR

The issue of footwear as a risk factor for injury comes in two separate debates: (i) whether barefoot/minimalist running is more beneficial than shod running and, (ii) if different types of footwear in shod running will affect injury incidence.

The recent trend for runners to change to barefoot or minimalist shoe running and its effect on injury risk is a hotly contested debate. There is a lack of prospective and controlled research in this area to support either side of the argument. Briefly, the argument for barefoot, or minimalist shoe, running is that it encourages a forefoot strike pattern which results in decreased impact loading and a shorter stride length compared to shod running which encourages a rear foot strike pattern.[13,241] Current research suggests that the effect on injury risk largely depends on an individual's ability to transition to a forefoot strike in barefoot running, however there is currently no strong evidence that barefoot running reduces injury risk in general, or AT in particular.[13,241]

Type of footwear is also often listed as a risk factor for developing AT.[132,239] Rowson et al. (2010)[245] recently tested the effects of footwear on tendon loading in cadaveric human Achilles tendon. Wearing high-top athletic shoes was found to decrease the peak tension in the Achilles tendon by 9.9% compared to low-top athletic shoes (P=0.0002), and well-tied laces decreased peak tension a further 12.8% (P=0.0006). In addition, wearing well-tied high-top athletic shoes decreased the peak dorsiflexion angle by 7.2% compared to low-top shoes. It is suggested that well-tied laces ensure that the shoe cannot move independently of the foot and, together with the high-top shoes, allow more force to be transferred to the shoe rather than being transferred into the foot and tendons.

Whilst there is limited evidence that footwear, either the lack thereof or different types thereof, will affect the biomechanics and force distribution across the foot, there is insufficient evidence to include footwear as an independent risk factor for AT. From a

University of Cape Town

different perspective, footwear orthotics and in-soles can be used to correct biomechanical faults or malalignments which may predispose to injury.[244]

1.5.1 (VI) SMOKING

A retrospective study of 1115 male Royal Marine recruits found a significant difference in smoking status prior to the training programme in injured and non-injured recruits (RR 1.7, 95% CI:1.2-2.8; P<0.01).[202] This study did not distinguish between injury mechanism or diagnosis but reported only on overall injury. A similar study on 2002 army recruits, however, reported that recruits who had smoked at least one cigarette in the month prior to the start of basic training subsequently sustained more overuse injuries (P<0.01).[12] Interestingly, the proportion of recruits who suffered from "tendinitis" was higher in the smokers when compared to the non-smokers, although this observation was not statistically significant. Contrary to this result, a cross-sectional study of medical risk factors for AT reported less smokers in patients with chronic AT when compared to healthy athletes (OR 0.2, 95% CI:0.1-0.4; P=0.001).[157]

There are substantial confounding factors to consider when investigating the influence of smoking on injury risk. In particular, runners and active individuals are more likely to adhere to a healthy lifestyle which does not include smoking, whereas smoking may be an indication of a less healthy lifestyle which could be associated with other potential risk factors such as inactivity or risk taking behaviour.[12] In summary, there is limited evidence of low certainty that smoking, or a history of smoking, is a risk factor for injury, but not specifically AT. There is no evidence that directly supports smoking as an independent risk factor for AT.

1.5.1 (VII) MEDICATIONS

(a) Statins

Statins are 3-hydroxy-3-methylglutaryl coenzyme-A reductase inhibitors used as a therapy for lowering serum cholesterol and low density lipoprotein levels.[187] A 2008 retrospective meta-analysis reported 96 cases of statin-attributed tendinous manifestations which accounted for 2.09% of statin related adverse effects over a 15 year period.[187] Cases were

University of Cape Town

attributed to statin use based on a combination of (i) a temporal relationship between the onset of symptoms and initiation of statin therapy, (ii) improvement in symptoms after cessation of statin therapy and (iii) recurrence of symptoms after statin therapy was reinstated. Specific diagnoses were "tendinitis" (65.6%) and tendon rupture (34.4%), and the Achilles tendon was the most frequent site of injury accounting for 52.1% of the 96 cases. Furthermore, a 2001 case report presented four patients who specifically developed tendinopathy after statin treatment.[48]

The mechanisms for statin toxicity in tendons are still not clear although several hypotheses have been suggested. As cholesterol is an important component of cell membranes, Marie et al. (2008)[187] suggest that statins may result in the destabilisation of cell membranes. In addition, statins may decrease levels of regulatory proteins within the tendon, and increase levels of apoptosis.[187] It has subsequently been reported that statins inhibit MMP-9 and enhance the activity of TIMP-1 in macrophages which would influence tendon remodelling and repair.[29,229,262] Furthermore, the effects of statins on prostaglandin E_2, which is involved in the development of tendinopathy due to repetitive loading, and leukotriene B_4 activity have been implicated in statin-induced tendinopathy.[55,262] Although a rare complication, there is limited evidence to suggest statin use as a risk factor for AT with a moderate certainty.

(b) *Anabolic steroids*

Anabolic androgenic steroids (AAS) are prohibited, performance enhancing drugs that are widely used in certain sporting codes.[188] They have been linked to increased risk of musculotendinous injury, particularly tendon ruptures, in a number of reviews and case studies.[132,167,171,292] In particular, case reports of very rare simultaneous bilateral quadriceps tendon rupture[171] and bilateral avulsions of the biceps tendon[292] identified AAS abuse as a predisposing cause. Interestingly, another case report presented a 35-year old male with a history of over-the-counter oral androstenediol supplementation who suffered asynchronous bilateral Achilles tendon ruptures in the month following supplementation.[25] The association of AAS use with tendon injury was, however, not supported by a questionnaire based study of retired NFL players which found that AAS use was not associated with self-reported, medically diagnosed muscle/tendon injuries.[118]

University of Cape Town

AAS use combined with exercise in rats has been shown to result in large pathological changes in tendon tissue including changes in contractility and crimp morphology, collagen dysplasia, decreases in collagen degradation, increased synthesis of type I collagen and decreased tensile strength.[124,167,188,218] Whilst there is substantial evidence that AAS use is pathological to tendon tissue in rats, and to a lesser extent humans, there is currently only weak evidence of low certainty to support it as a risk factor for AT. The conduction of randomised controlled trials (RCT) and prospective studies on the effects of AAS use on injury risk is complicated by the ethical and legal issues around its use.

(c) Oral Contraceptives and HRT

The use of oral contraceptives (OC) and hormone/oestradiol replacement therapy (HRT/ERT) in women has been suggested as a risk factor for developing tendinopathy.[116,306] A cross-sectional study of 82 patients with symptomatic AT compared to two control groups reported a higher prevalence of both HRT (P=0.01) and OC (P=0.001) use in women with AT.[116] Several case-control studies of healthy pre-menopausal and post-menopausal women have shown that circulating oestradiol (the predominant oestrogen) has significant effects on collagen synthesis in tendon.[41,102,103,104] In particular, use of OC attenuated the exercise induced increase in collagen synthesis in the tendon either as a direct effect of circulating oestradiol, or an indirect effect of reduced levels of free IGF-1.[103,104] The long term attenuation of endogenous plasma oestrogen in users of the monophasic OC pill was found to significantly decrease compliance of the Achilles tendon which may have implications for risk of injury.[41] In addition, long term HRT in post-menopausal women stimulates collagen turnover at rest in a dose dependant manner and results in decreased stiffness of tendon during mechanical loading.[102] There are currently, however, no prospective studies on the effect of either OC or HRT on AT incidence or risk, therefore the level of evidence for OC and HRT use as risk factors for AT is limited and of low certainty.

(d) Fluoroquinolone antibiotics

Fluoroquinolones are a class of synthetic broad spectrum antimicrobial drugs commonly used to treat infections in adults.[46,172] They inhibit bacterial deoxyribonucleic acid (DNA) replication and have a significant ability to act against anaerobic bacteria.[53] These drugs are widely used to treat a variety of infections including soft tissue, bone and joint infections,

University of Cape Town

common urinary and respiratory tract infections, as well as complicated skin infections.[148,172,303] Fluoroquinolone associated tendinopathies were identified in the early 1980's and have been shown to occur at an incidence of 15-20 per 100 000 patients, predominantly in males.[46,53,100] It should however be noted that females may be at a higher risk of sustaining quinolone induced tendon ruptures.[303] In a nested case-control study by van der Linden et al. (2002),[286] the adjusted relative risk of Achilles tendon disorders with use of fluoroquinolones was 1.9 (95% CI:1.3 to 2.6). In patients aged 60 or over this risk increased to 3.2 (95% CI:2.1 to 4.9) and concomitant use of corticosteroids in this age group increased the risk to 6.2 (95% CI:3.0 to 12.8). This supports the concept of "synergistic toxicity" in patients undergoing simultaneous fluoroquinolone and corticosteroid therapy.[148] In a review of the literature describing fluoroquinolone associated tendinopathy, Khaliq et al. (2003)[148] found that 37.0% and 25.5% of cases followed treatment by pefloxacin and ciprofloxacin respectively. The most commonly affected tendon was the Achilles, accounting for 89.8% of cases, 44.3% of which were bilateral injuries. The mean latency period till the onset of symptoms was 17.6 ± 19.5 days, with 50% of cases occurring within 6 days and several reports described prolonged recovery periods for the tendon injury. Prior use of corticosteroids was found in 32.7% of tendinopathy cases and 52.5% of ruptures. These findings are consistent with initial reports.[46,100,172,213]. There is consistent limited evidence to suggest with high certainty that fluoroquinolone treatment is a risk factor for AT and ATR.

The mechanism underlying fluoroquinolone associated tendon injury is not clear but influences on ischaemia and vascularity, proteoglycan synthesis and matrix degradation have been implicated, as well as the generation of reactive oxygen species and increases in cytotoxicity and necrosis.[46,53,148] Toxicity may be related to molecular structure as the most toxic fluoroquinolones share a similar structure.[148] Evidence suggests that renal transplant patients and patients suffering from other metabolic conditions which affect renal drug clearance may be further predisposed to fluoroquinolone associated tendon injuries.[148]

(e) <u>Corticosteroids</u>

Corticosteroid, particularly glucocorticoid, injections are commonly used to treat tendon lesions. There is however lack of consensus on both the efficacy and safety of

University of Cape Town

intratendinous corticosteroid injections.[63,192,206,263,273] Systematic reviews have consistently concluded that there is insufficient evidence to support or refute the efficacy of corticosteroid injections in the treatment of tendon lesions given the potential for complications, including tendon rupture.[63,192,206,263,273] In particular, a 2010 systematic review of RCTs of corticosteroid treatment for tendinopathy concluded that, although there was a large beneficial short-term effect, the long-term outcome was worse than that achieved with more conservative treatments.[63] The efficacy of glucocorticoid treatment is in fact influenced by the underlying pathology of the tendon injury, and it will likely only be effective if inflammation, not degeneration, is the underlying cause.[192] In addition, there is lack of consistency regarding the effects of glucocorticoid treatment on the biomechanical properties of tendon in both human and animal studies.[206,254] A systematic review by Nichols (2005)[206] reported some evidence that local injection of corticosteroids acutely reduced tendon strength to failure, reduced tendon stiffness and induced histological evidence of tendinopathy. However, these findings were not consistent, with some studies reporting no deleterious effects of corticosteroid injection. In a study on cultured rat tenocytes, corticosteroid (dexamethasone) treatment was found to reduce tenocyte number and collagen synthesis in a dose-dependent manner.[254] In addition, the concomitant use of corticosteroids increases the risk of suffering fluoroquinolone induced tendinopathy (reviewed above).[148,286,303] In summary, corticosteroids are catabolic, inhibit the synthesis of collagen and are consistently linked to tendon injury.[53,132,206,208] There is however limited evidence of moderate certainty linking corticosteroid treatment to AT and ATR.

University of Cape Town

1.5.2. INTRINSIC RISK FACTORS

1.5.2 (I) AGE

As medical care improves and physicians encourage cardiovascular fitness at all ages, the elderly population is remaining active for longer and therefore at risk of sustaining physical activity related soft tissue injuries.[74,75] Age is often listed as an intrinsic risk factor for AT and ATR.[53,116] However, not only the elderly are at higher risk of tendon injury, certain age groups are associated with a higher prevalence of AT and/or ATR. In a retrospective study of ATR in a Danish population, 62% of ATR occurred in the 30-49 year age group.[121] Sports

related ruptures in particular occurred almost exclusively between 30 and 49 years of age, while non-sport related ruptures peaked between 50 and 59 years of age. These results are similar to those found in a Canadian population.[275] The mechanism of rupture therefore appears to be different between younger and older patients.[121,275] In a cross-sectional study, the highest incidence of AT in the general Dutch population was found in the 21-40 and 41-60 year age groups.[68] In particular, the incidence rate of AT was highest for females in the 21-40 year age group and highest for males in the 41-60 year age group. Interestingly, the incidence rate was substantially lower for adults over the age of 60 years, than for the younger age groups. There is therefore limited evidence of low certainty that older adults are at higher risk of ATR and AT.

Age related changes in tendon properties may explain the higher incidence of ATR and AT in older adults. An early, but very informative, review of the ageing tendon states that the rate of ageing is highly individual but that tendon tissue is particularly susceptible to early ageing.[282] Age-related cellular changes within the tendon include decreased density and activity of tenoblasts, although it should be noted that rabbit tendon stem cells do not lose their ability to repair tendon with age.[74,282] Within the ECM and collagen network, ageing results in decreased extracellular water and muccopolysaccharide content, decreased GAG content and an increase in the total collagen content.[282] There is a decrease in collagen turnover and collagen fibres increase in diameter to varying degrees.[282] In particular, ageing affects the cross-linking profiles of collagen.[78,282] Collagen cross-links are subject to non-enzymatic glycation by the reduction of circulating sugars.[23] This results in permanent modifications to the protein cross-links which are known as advanced glycation end products (AGEs) such as pentosidine. After maturity, AGEs accumulate with age which is indicative of decreased collagen turnover. The amount of permanent cross-linking increases with age up till maturity (approximately 30 years of age) and then gradually decreases.[23,66,78] Cross-linking affects the physical properties of collagen fibres, and therefore tendon, and an increase in permanent cross-links results in increased stiffness of the tendon as well as a decrease in maximum strain and load.[23,75,282] In addition, the Achilles tendon has little ability to adapt to mechanical loading after maturity and therefore accumulates microtrauma.[269] Considered in conjunction with the age-related decline in mechanical properties, this results in a tendon that is more susceptible to rupture and injury. It should

University of Cape Town

however be noted that it has been suggested that these changes are not intrinsically age-related, but rather a secondary response to tendon injury and microtrauma.[23]

1.5.2 (II) SEX

In two retrospective studies on the epidemiology of ATR, males were three to four times more likely to suffer ATR than females.[121,275] Suchak et al. (2005)[275] reviewed the incidence of ATR and found the male to female ratio to be between 1.7 and 6.3, but predominantly higher than 3.0, and male sex is repeatedly reported as a risk factor for both ATR and AT.[116,131,132,157] In a cross-sectional study of 298 healthy participants, asymptomatic Achilles tendon pathology was found at a higher prevalence in males (13%) compared to females (5%)($P=0.007$, χ^2 7.189).[94] However, a retrospective study of AT in the general Dutch population found that total incidence rates for males and females were not different,[68] and a prospective clinical trial of 648 runners found no association between gender and risk of developing AT.[113] Females have, however, been found to be at higher risk for other connective tissue injuries.[153,201] In particular, females were found to have a higher incidence of sports related ACL ruptures when injuries from male-only sports were excluded.[201] Nevertheless, there is limited evidence of low certainty that male sex is a risk factor for AT.

Oestrogen is suggested to play an important role in the homeostasis of female connective tissue and oestrogen receptors have been identified in human tendons.[106] A cross-sectional study of six males and six females found that oestradiol attenuates the exercise induced increase in collagen synthesis, and that collagen fibres from male tendon had a higher ultimate stress ($P<0.05$).[181] As previously described, several studies on the effects of oral contraceptives and HRT on collagen and tendon properties suggested that synthetic oestradiol inhibits collagen synthesis in response to exercise, but increases collagen synthesis at rest in post-menopausal women.[41,102,103,104] The seemingly contradictory evidence for the effects of oestrogen (particularly oestradiol) on tendon properties has been succinctly summarised by Kjaer and Hansen (2008).[153] They suggest that endogenous oestrogen has a homeostatic effect on collagen synthesis by stimulating collagen synthesis at rest but attenuating any exercise-induced increase. This is corroborated by evidence that the Achilles tendons of trained women have a smaller cross-sectional area when compared to trained men ($P<0.01$).[298] Oral contraceptives (synthetic oestradiol) however, result

University of Cape Town

predominantly in overall suppression of collagen synthesis. This results in women having a slower adaptive response to mechanical loading but may allow them to more robustly withstand periods of inactivity and detrimental perturbations.

1.5.2 (III) ADIPOSITY

A cross-sectional study of 82 patients with asymptomatic AT compared to two control groups reported a higher prevalence of obesity in both males (P=0.001) and females (P=0.025) with AT.[116] In a systematic review of the relationship between adiposity and AT, exactly half of the included studies (n=28) showed a positive association while the other 14 studies reported no significant association between adiposity and AT.[96] Two possible mechanisms for this tentative association are suggested, (i) a mechanical and (ii) a systemic hypothesis.[96] The mechanical hypothesis proposes that tendons, particularly weight bearing tendons, are exposed to higher loads in heavier individuals and this results in tendinopathy. It should, however, be noted that similar distributions of positive and negative associations with adiposity were found in studies on upper and lower extremity tendinopathy when one would expect this relationship to be more evident in weight-bearing tendons of the lower extremity. It is therefore clear that mechanical loading does not fully account for the association between adiposity and tendinopathy. The systemic hypothesis proposes that bioactive peptides released from adipose tissue may influence the integrity of tendon directly, or indirectly by altering the metabolic milieu. This review did however report a significantly higher number of positive associations found in studies on clinical patients and case-control studies in which patients are recruited from specialist centres after referral, often because they are not responding to treatment. It is therefore possible that the association between adiposity and AT observed in these studies may actually reflect increased adiposity as a consequence of tendinopathy, or adiposity as a factor limiting recovery. In order to address the confounding issue of pain leading to inactivity and weight gain, Gaida et al. (2010)[94] recently investigated fat distribution in asymptomatic tendon pathology in a large cross-sectional study of otherwise healthy Swedish and Australian participants. Men with asymptomatic tendon pathology were found to have a more central distribution of fat, while women with asymptomatic tendon pathology were found to have a more peripheral distribution of fat. The highest prevalence of asymptomatic tendon pathology was found in males over the age of 40 years old with a waist circumference larger

University of Cape Town

than 83cm. There is therefore limited evidence of moderate certainty that adiposity, and particularly fat distribution, is a risk factor for AT.

1.5.2 (IV) PREVIOUS INJURY

A retrospective follow up of 168 patients with ATR reported an increased risk of an acute rupture of the contralateral tendon within four years compared to the general population (OR 176, 95% CI:70-282; P<0.001).[16] In a cross sectional matched pair analysis, 37% of patients with an acute ATR were reported to have a positive history of chronic AT before the rupture,[157] and 41% of patients diagnosed with acute/subchronic AT suffered from symptoms on the contralateral side within eight years.[215] A recent prospective clinical trial on the prognostic value of power Doppler ultrasound in 634 asymptomatic runners found that pre-existing healed Achilles tendon disorders were a significant risk factor for the development of mid-portion AT within one year (OR 3.8, 95% CI:1.7-8.5; P=0.0014).[113] In addition, a prospective study of 54 Danish soccer players showed a significantly higher prevalence (n=5; 45%) of symptomatic AT at the end of the season in those who had asymptomatic tendinopathic features on ultrasound at the beginning of the season (n=11) (P<0.05).[88] Although the subject numbers may have been small in this study, considered in conjunction with the other evidence, there is strong evidence of a high certainty that previous injury of the Achilles tendon is an independent risk factor for AT, and probably ATR.

1.5.2 (V) DECREASED BLOOD SUPPLY TO TENDON

Ischaemia and hypoxia have previously been suggested as risk factors for AT and ATR.[132,145] However, a case-control study of laser Doppler flowmetry in chronic AT reported a higher blood flow in symptomatic tendons than in controls, with both groups displaying a similar reduction in blood flow to the Achilles tendon in response to exercise.[18] In addition, a recent prospective clinical trial screened 634 asymptomatic runners with power Doppler ultrasound and found that the presence of intratendinous microvessels, indicative of neovascularisation, was a significant risk factor for developing mid-portion AT within a year (OR 6.9, 95% CI:2.6-18.8; P=0.0001).[113] Blood flow in the tendon is decreased during exercise and certain areas are therefore exposed to hypoxia which may result in reperfusion injury during the subsequent hyperemia.[19,145] It has also been suggested that reperfusion

University of Cape Town

injury sustained following contraction may enhance the production of reactive oxygen species and lead to oxidative damage in the tendon.[145,176] The Achilles tendon may be particularly prone to this as the tendon spirals by approximately 90° during use, causing a "wringing" of the tendon and vascular constriction.[280] Although several case studies have reported the occurrence of ATR following ischaemia,[126,264] there is no further evidence that ischaemia or hypoxia are risk factors for AT or ATR. There is however strong evidence of moderate certainty that neovascularisation observed during power Doppler ultrasound is an independent risk factor for AT.

1.5.2 (VI) TENDON TEMPERATURE

Tendons are not perfectly elastic and some energy is lost as heat (hysteresis) during exercise. Blood supply to the tendon is not sufficient to dissipate the heat generated and the temperature of the tendon therefore increases.[31] Wilson et al. (1994)[302] recorded peak intra-tendinous temperatures of 43-45°C in equine SDFT during a sustained gallop and it was suggested that this exercise induced hyperthermia may result in tendon degeneration.[302] However, these temperatures were found to be unlikely to result in cell death in cultured equine SDFT fibroblasts.[31] Tendon fibroblasts were in fact found to be more resistant to hyperthermia than other fibroblast lines, although repeated exposure to hyperthermic conditions may compromise cell functioning.[31] Farris et al. (2011)[80] collected ultrasound, kinetic and kinematic data from 12 males and used a mathematical model similar to that used by Wilson et al. (1994)[302] to predict the core temperature of human Achilles tendon during treadmill running. A conservative estimate of a core temperature of 41.4°C during running was calculated. It should, however, be noted that hysteresis and stiffness, both variables that influence the calculation, were highly variable between individuals and were significantly negatively correlated. It is therefore possible that differences in mechanical tendon properties may predispose some individuals to greater thermal damage.[80] However, there are currently no studies directly linking tendon hyperthermia with AT.

1.5.2 (VII) BIOMECHANICAL FACTORS

The list of lower limb malalignments and biomechanical factors that have been suggested as risk factors for AT and other musculoskeletal disorders is long. The malalignments include forefoot varus and valgus, hindfoot varus and valgus, pes planus and cavus, hyper- or

University of Cape Town

hypopronation, tibia vara, genu valgum and varum, a high- or low-riding patella and femoral neck anteversion.[145,293] Other factors include leg length discrepancy, muscle weakness and imbalances, decreased flexibility and joint laxity.[145]

During gait, the foot strikes the ground in a supinated position, pronates and then supinates again at toe-off. During pronation, the foot is everted, abducted and dorsiflexed, while during supination the foot is inverted, adducted and plantar-flexed.[127] This results in a "bowstring" or "whipping" action which creates shear forces across the Achilles tendon during gait.[191,250] Hyperpronation may result from forefoot varus, leg length discrepancy, lax ligaments or muscular imbalances in the gastrocnemius and soleus muscles, and results in an exacerbated "whipping" action during gait.[145] At midstance the foot is pronated and the knee extended which results in contradictory rotational forces across the Achilles tendon.[250] If the foot then remains in a pronated position it may result in wringing of the tendon.[191] This theory was however, not supported by a recent systematic review of prospective cohorts and case-control studies investigating the biomechanical factors associated with AT.[203]

The subtalar joint in the ankle is responsible for both the inversion/eversion and pronation/supination movements of the foot, while the ankle joint is responsible for plantar- and dorsiflexion. It has been suggested that excessive loading of the Achilles tendon with inversion or eversion of the subtalar joint, such as in hyperpronation, may result in injury.[169] In support of this, a prospective study of 449 US Navy SEALs reported that a tight gastrocnemius muscle (RR 3.57, 95% CI:1.01-12.68; $P<0.05$) and increased hindfoot inversion (RR 2.79, 95% CI:0.91-8.55; $P<0.001$) were significantly associated with risk of developing AT.[147] Furthermore, a case-control study showed that runners with AT have a higher ankle eversion than seen in controls ($P<0.05$).[246] As the Achilles tendon antagonises pronation at heel strike and is the predominant plantar-flexor during the toe-off phase, stresses across the Achilles tendon are correlated to the orientation of the subtalar joint axis.[235,250] A retrospective analysis of the spatial orientation of the subtalar joint axis in runners showed that runners with chronic AT had a more oblique deviation angle of the subtalar joint axis at maximal dorsiflexion ($P=0.002$), and this may result in more wringing of the tendon during pronation.[235] Waldecker et al. (2012)[293] recently performed a cross-

University of Cape Town

sectional investigation of the tibiocalcaneal axis in 1394 feet and found that varus malalignment (-14° to 0°) of the hindfoot was present in 78.3% of feet with AT compared to 34.3% of control feet (P<0.001). This corroborates the retrospective findings of Kvist et al. (1994)[162] in which forefoot varus was reported to be associated with Achilles tendon overuse injuries.[162]

Flexibility and range of motion (ROM) have been associated with patellar tendinopathy,[185,287] and decreased ROM during passive dorsiflexion of the ankle and subtalar joints is also implicated as a predisposing factor for AT.[162] In a prospective study of 69 well matched military recruits, lower plantar flexor muscle strength and a higher ROM during dorsiflexion were found to be significant risk factors for AT overuse injury.[182] In contrast, athletes with a lower ROM in dorsiflexion were found to absorb landing loads with a more everted and extended plantarflexor muscle-tendon unit, which may increase the risk of overuse injuries such as AT.[299] A cross sectional study of biomechanical variables in runners with AT found a decreased ROM at knee flexion from heel strike to midstance.[20]

Recently, Wyndow et al. (2012)[305] have suggested that altered neuromotor control of the triceps surae is a feature of AT. They reported significant differences in the activation of the triceps surae in runners with AT compared to controls, with earlier soleus muscle offset relative to lateral gastrocnemius offset.[305] This corroborates a cross sectional study of biomechanical variables in runners with AT which found decreased muscle activity in the lower extremity of these runners.[20]

A 10 week prospective cohort study on well-matched novice runners showed that a decrease in the anterior displacement of the centre of force (P=0.015) and a more lateral force distribution at forefoot flat (P=0.016) were significant risk factors for Achilles tendon overuse injuries.[289] The authors suggest that the decrease in forward transfer of the centre of force results in less propulsion being generated during the late stance and therefore higher tensile forces are required during the toe-off stance.

A 2011 systematic review of two prospective cohorts and seven case-control investigations of biomechanical factors specifically associated with AT concluded that individuals with AT often have: (i) increased eversion ROM of the hindfoot, (ii) a reduced maximum lower leg

University of Cape Town

abduction, (iii) reduced ankle joint dorsiflexion velocity, (iv) reduced knee flexion during the gait cycle, (v) altered plantar pressures and ground reaction forces, (vi) a reduced tibial external rotation moment and (vii) altered timing and amplitude of activity in lower limb muscles.[203] In summary, the studies reviewed above present strong evidence of low certainty that a more lateral force distribution across the foot is a risk factor for AT, limited evidence of moderate certainty that varus malalignment is a risk factor for AT and strong evidence of moderate certainty that increased eversion and inversion of the foot is a risk factor for AT. There is limited evidence of low certainty that reductions in muscle activity in the lower extremity, and changes in the relative timing of lower limb muscle activity are risk factors for AT. There is also strong evidence of low certainty that ROM and decreased flexibility of the gastrocnemius muscle are risk factors for AT.

1.5.2 (VIII) SYSTEMIC DISEASE

It has been stated that approximately 2% of Achilles tendon disorders arise as complications of systemic disease.[132] Several systemic diseases have been suggested as risk factors for AT and ATR including diabetes mellitus, compromised renal function, hyperlipidemia, hypercholesterolemia, hyperuricaemia, Alkaptonuria, gout, rheumatoid arthritis, adrenal disorders, thyroid disorders (hyperparathyroidism), amyloidosis and chronic lung disease.[4,53,172,197,233,239] In addition, there are several severe inherited disorders of connective tissue such as Ehlers-Danlos syndrome, Marfans syndrome and osteogenesis imperfecta which have joint hypermobility and other musculoskeletal manifestations as symptoms.[184]

Diabetes mellitus has been shown to result in degeneration and decreased vascularity in tendons.[4] The high availability of plasma glucose in diabetes results in the increased accumulation of AGEs and permanent cross-links in collagen fibres.[4,239] A case-control study of induced diabetes in rats showed that sustained hyperglycaemia resulted in altered structural properties of tendon and an increased incidence of mid-portion patellar tendon rupture.[87]

Patients with chronic painful AT also show evidence of dyslipidemia with lipid profiles characteristic of insulin resistance and the metabolic syndrome.[95] Hypercholesterolemia

University of Cape Town

was reported to result in increased stiffness in the supraspinatus tendon in several species, and chronic hypercholesterolemia in rats is associated with detrimental changes in tendon biomechanics.[26,27] A case-control study showed an association between hyperlipidemia and Achilles tendon pathology such as tendon xanthomas and "tendinitis".[155] Tendon xanthomas are common in familial hypercholesterolemia and are attributed to the accumulation and oxidisation of low density lipoprotein in tendon.[4] However, tendon degeneration in non-familial hypercholesterolemia is not well understood and it has been suggested that microscopic cholesterol accumulation in tendon may initiate a chronic low-grade inflammatory response.[4] In addition, a case-control study showed that patients with ATR had significantly more unfavourable lipid profiles, and higher serum lipid concentrations than a control group.[214]

The evidence for diabetes mellitus and hypercholesterolemia as risk factors for AT is therefore limited, and of low and moderate certainty respectively. When considered with the evidence for adiposity and fat distribution as risk factors for AT, this indicates that tendinopathy may be a significant feature of the metabolic syndrome.

1.5.2 (IX) BLOOD GROUP

In 1989 Jozsa et al.[139] reported an apparent association between ATR and the O blood group in a Hungarian population. This result was corroborated by some[146,158] but not all studies.[16,170,179] The ABO blood group system is the most commonly used blood grouping system as nearly all individuals express the H surface antigen on their red blood cells. The two alleles of the *ABO* gene, situated on the telomeric end of the long arm of chromosome 9, encode the A- and B-glycosyltransferases.[307] These enzymes catalyse the final step in the biosynthesis of the A and B blood group antigens by transferring different sugars to the acceptor H antigen.[307] It has been suggested that these transferases may not only determine the structure of glycoprotein antigens, but also the structure of glycoproteins within the ECM.[139,158] However, it has also been proposed that other genes closely linked to the *ABO* gene on chromosome 9 may be more likely to be associated with AT.[146,158] In summary, there is limited evidence of low certainty to link the ABO blood groups with AT.

University of Cape Town

1.5.2 (X) GENETIC PREDISPOSITION

It has long been hypothesized that genetic predisposition may influence risk of developing chronic tendon disorders.[133,145] Subsequent findings have confirmed that a familial predisposition, and therefore genetic predisposition, exists in the risk of developing soft tissue injuries. In particular, the relative risk of sustaining rotator cuff tears in siblings of patients with full thickness tears was higher than that in the control population of patients spouses (RR 2.42, 95% CI:1.77-3.31).[107] More recently, in a cross-sectional match paired analysis of medical risk factors for AT and ATR, the only significant risk factor to emerge was a positive family history of chronic AT.[157]

In addition to family history, several DNA sequence variants have been associated with risk of developing AT and ATR. Mokone et al. (2006)[199] and September et al. (2009)[255] investigated the association of a polymorphism within the *COL5A1* gene, which encodes the α1 chain of type V collagen and is mapped to the same locus as the *ABO* gene on chromosome 9 (9q34), with Achilles tendinopathy in South African and Australian populations. A *BstUI* restriction fragment length polymorphism (RFLP) within the 3'-untranslated region (UTR) of the *COL5A1* gene (rs12722, C>T) was significantly associated with AT in both of these groups. Specifically, individuals with a CC genotype at this locus were at a significantly lower risk of developing chronic AT than individuals with either a CT or TT genotype in both the South African (OR 0.38, 95% CI:0.18-0.77; P=0.008) and Australian (OR 0.42, 95% CI:0.20-0.86; P=0.017) group. The rs12722 single nucleotide polymorphism (SNP) has also been associated with ACL tears in females, ultra-endurance running performance, exercise associated muscle cramping and ROM, a possible risk factor for tendinopathy.[40,56,209,227] Recently, three more SNPs in the 3'-UTR of the *COL5A1* gene, rs71746744, rs16399 and rs1134170, have also been independently associated with chronic AT in South African and Australian participants.[6] The associated polymorphisms within *COL5A1* have been mapped to a functional region of the 3'-UTR of this gene and are suggested to play a role in the stability of *COL5A1* mRNA.[164] In particular, two functional forms of *COL5A1* 3'-UTRs cloned from controls and AT subjects were identified.[164] The C- and T-forms were predominantly identified in the controls and AT participants respectively and the luciferase activity of the C-form was significantly lower than that of the T-form (69.0 ± 22.0% vs 90.6 ± 13.7%; P<0.001). This suggests that mRNA from the T-form of the *COL5A1*

University of Cape Town

3'-UTR is significantly more stable than that of the C-form. In addition, there is a functional micro RNA (miRNA) binding site for Hsa-miR-608 within the *COL5A1* 3'-UTR and the rs4919510 SNP within the *MIR608* gene, encoding Hsa-miR-608, has also been shown to be independently associated with chronic AT.[6]

Collins and Posthumus (2011)[57] hypothesized that the *COL5A1* rs12722 TT genotype results in increased type V collagen based on the increased stability of *COL5A1* mRNA in the presence of the T-allele (Figure 1.9). Increased type V collagen results in a decreased mean fibril diameter and an increased fibril density, which leads to changes in the biomechanical properties of the tissues. They further hypothesize that these changes result in: (i) a reduced tensile strength and therefore higher risk of musculotendinous injuries, and (ii) increased creep inhibition and/or stiffness which improves running economy and endurance running performance.[57]

The α-chains of the type V collagen fibril are encoded by the *COL5A1*, *COL5A2* and *COL5A3* genes. In addition, type III collagen (encoded by the *COL3A1* gene) interacts with type I and type V collagen in the ECM. Although the association of other collagen genes with AT has been investigated,[228,259] to date genetic variation in the *COL5A2*, *COL5A3* and *COL3A1* genes has not been investigated in AT. Polymorphisms within the *COL1A1*, *COL12A1* and *COL14A1* genes were shown not to associate with chronic AT.[228,259] The Sp1 binding site polymorphism, rs1800012, within the *COL1A1* gene and an *AluI* RFLP within the *COL12A1* gene were, however, reported to be associated with ACL ruptures.[225,226]

In addition to the association of several collagen gene polymorphisms with AT, polymorphisms within genes encoding glycoproteins and matrix associated enzymes have also been implicated in AT. The *TNC* gene, encoding the glycoprotein tenascin-C (TN-C), is also situated on the long arm of chromosome nine and contains a GT dinucleotide repeat polymorphism in intron 17.[197] Alleles with 12 or 14 repeats were found to be significantly over represented, and alleles with 13 or 17 repeats were under represented, in patients with Achilles tendon injury compared to healthy controls. Polymorphisms within genes encoding other glycoproteins of the ECM, including the thrombospondin family of five glycoproteins, have not been investigated for association with AT.

University of Cape Town

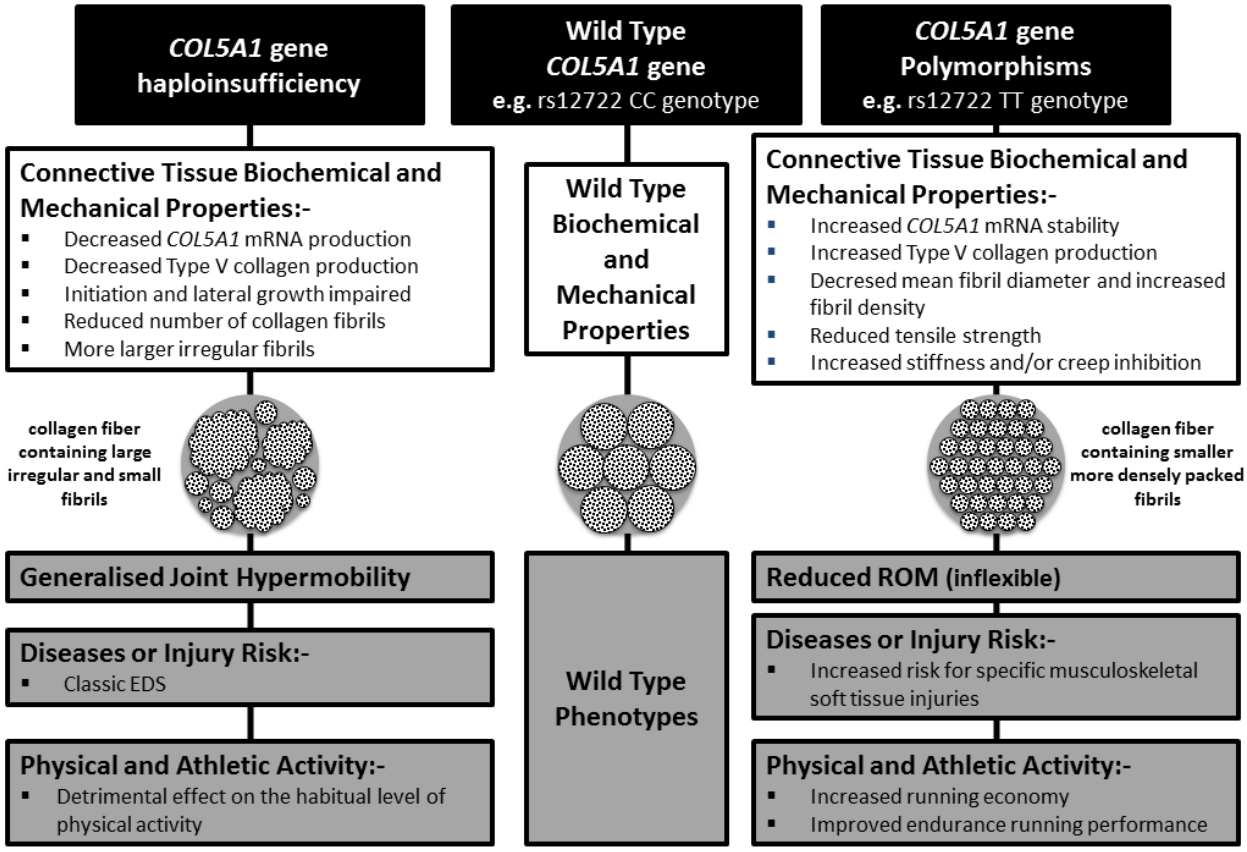

Figure 1.9: Collins and Posthumus' hypothesis of the relationship between *COL5A1* rs12722 genotype and exercise related phenotypes

A schematic summary of the relationship between (i) COL5A1 genotype (black boxes), (ii) connective tissue biochemical and mechanical properties (white boxes), (iii) flexibility, (iv) disease or injury risk, and (v) physical activity. The left panel illustrates the effects of disease-causing COL5A1 mutations on decreased type V collagen production, abnormal fibrillogenesis, and generalized joint hypermobility. These mutations cause Ehlers-Danlos syndrome (EDS), which has been shown to have a detrimental effect on the habitual level of physical activity within these patients. A mixture of large and small irregular fibrils in EDS is shown. The scenarios illustrated in the middle and right panels described the normal interindividual biological variation. The middle panel represents the wild-type COL5A1 gene and phenotypes. It is proposed that larger regularly shaped fibrils are produced from the wild-type gene, which is stronger and more compliant. These fibrils are associated with increased joint range of motion (ROM) decreased risk for specific musculoskeletal soft tissue injuries, and slower endurance running performance. The right panel illustrates the effect of functional common polymorphisms within the COL5A1 gene on increased type V collagen production. Smaller regularly shaped weaker fibrils are produced during fibrillogenesis. These fibrils, which are proposed to have an increased stiffness and/or creep inhibition, are associated with reduced joint ROM, increased risk for specific musculoskeletal soft tissue injuries, and faster endurance running performance

[Figure and caption reprinted from Collins and Posthumus (2011)[57] with permission from Wolters Kluwer Health]

In this same South African population, three polymorphisms within the *MMP3* gene for matrix metalloproteinase-3 have been associated with AT.[231] The GG genotype of rs679620, CC genotype of rs591058 and the AA genotype of rs650108 were all associated with AT, and the ATG haplotype for these three polymorphisms was significantly under-represented in the AT group (P=0.038). Furthermore, rs679620 within this *MMP3* gene interacts with the *COL5A1* rs12722 polymorphism to modify the risk of tendinopathy. In addition, the CC genotype of rs4789932 within the *TIMP2* gene, which encodes tissue inhibitor metalloproteinase-2, has been reported to be under-represented in a control group when compared to participants with AT (P=0.016).[76]

Several polymorphisms within genes encoding components of cellular pathways have also been shown to be associated with AT. In particular, individuals with a TT genotype of the rs143383 polymorphism within the growth/differentiation factor-5 gene, *GDF5*, were found to have double the risk of developing AT compared to controls in a South African and Australian population.[224] Several sequence variants within caspase (*CASP8*) and interleukin (*IL-6*, *IL1β*, *IL-1RN*) genes involved in the apoptosis and cell signalling pathways have also been implicated in AT.[205,257]

As the studies outlined above are all case-control genetic association studies, they constitute only limited evidence. However, given the repeatability and number of studies, the evidence for genetic variation as a risk factor for AT is of high certainty. It is interesting to note that there are no genome-wide association studies (GWAS) on tendinopathy. This may reflect the relative immaturity of the field of tendinopathy genomics as GWAS require very large cohorts in order to be successful.[186,313] As the field and, subsequently, collaborations and research groups grow, larger cohorts may allow for a GWAS approach to tendinopathy genomics. This will be discussed further in chapter 6. It is however clear that genetic predisposition is a risk factor in the multi-factorial condition of AT. It is likely that tendinopathy occurs at the milder end of a spectrum of connective tissue disorders ranging from complex, multifactorial conditions in which the genetic contribution is small, to severe Mendelian diseases in which genetic factors are the major determinants.[260]

University of Cape Town

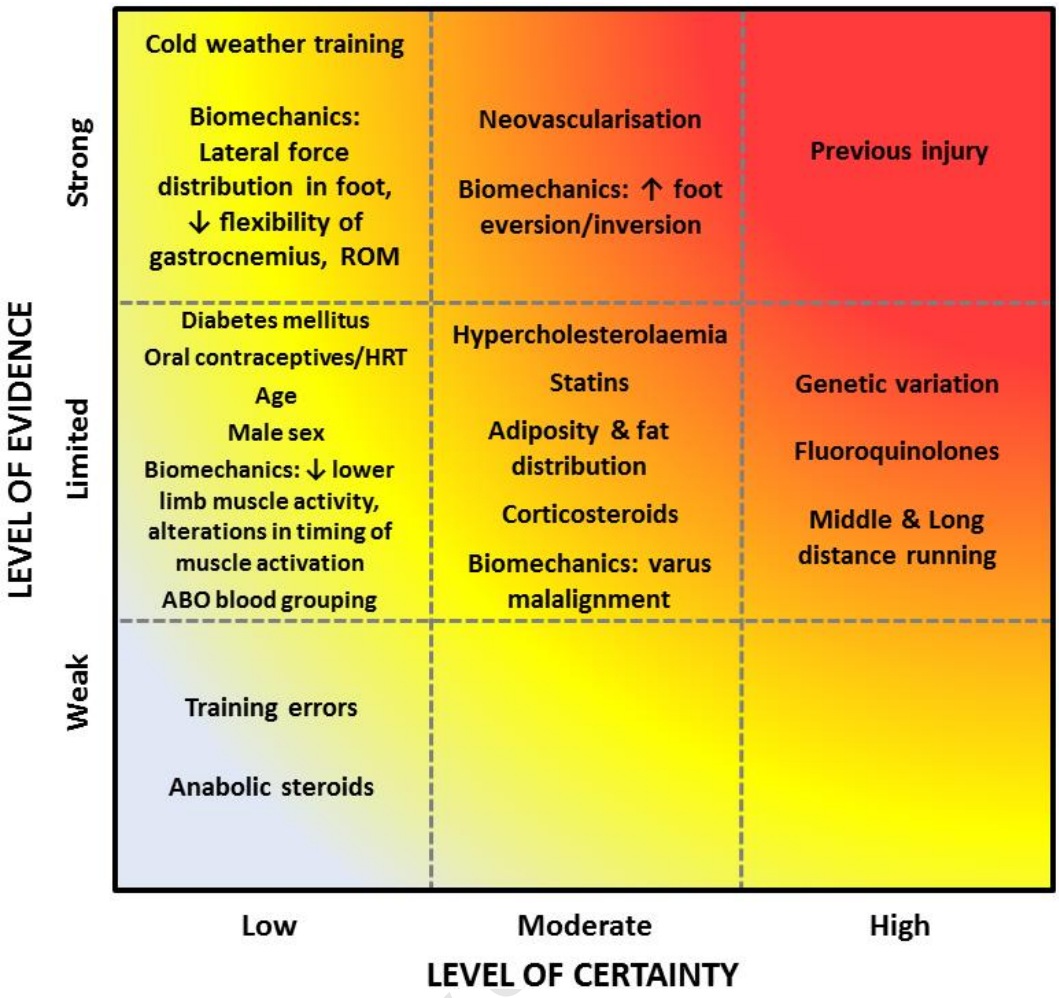

Figure 1.10: Level of evidence and level of certainty of risk factors for Achilles tendinopathy

University

1.6. SUMMARY AND CONCLUSIONS OF THE LITERATURE REVIEW

The incidence rates of Achilles tendon rupture and Achilles tendinopathy are both increasing. Although there are a number of models to explain the aetiology and mechanism of AT, the exact aetiology has not been elucidated. There is however a trend towards more integrated models and hypotheses which better explain the symptoms associated with AT, and the progression of the condition. The intrinsic and extrinsic risk factors associated with AT have been reviewed in this chapter and it is clear that AT is a multifactorial condition. From the evidence based review of each risk factor, it is also clear that more research is required to fully elucidate the risk factors which influence risk of AT. In particular, more prospective cohort studies are required in order to confirm initial findings of risk factors

associated with AT with more certainty. The risk factors which emerged from this review with the strongest evidence and highest degree of certainty were previous Achilles tendon injury, neovascularisation of the tendon, genetic variation, treatment with fluoroquinolone antibiotics and middle- and long-distance running.

The certainty for the existence of a genetic predisposition towards AT is increasing as more studies are published reporting the association of gene sequence variants with risk of AT. In addition, the increasing number of genes that are reported to be associated with AT emphasises the polygenic nature of genetic susceptibility to AT. It is therefore vital that the full spectrum of polymorphisms which influence risk of AT is uncovered.

1.7. AIMS AND OBJECTIVES OF THE THESIS

The primary aim of this thesis was to further identify specific genetic elements predisposing individuals to risk of AT using a candidate gene, case-control genetic association approach, and to propose the biological mechanisms underlying this genetic risk. Candidate genes (*COMP*, *THBS2*, *COL27A1*, *TNC*, *COL5A2*, *COL5A3* and *COL3A1*) were selected based on their chromosomal location and/or the biological function of their encoded proteins within the ECM.

The objectives of the specific studies which addressed this primary aim of the thesis were:

- To test the association of sequence variants within the *COMP* and *THBS2* candidate genes with risk of AT in participants from South Africa and Australia (Chapter 2)
- To test the association of several sequence variants spanning a 780.9kb region which includes the *COL27A1* and *TNC* genes, and a potential haplotype consisting of two or more of these SNPs, with AT in participants from South Africa and Australia (Chapter 3)
- To investigate the association of polymorphisms within the *COL5A2*, *COL5A3* and *COL3A1* genes with AT in participants from South Africa and Australia (Chapter 4)
- To investigate interactions between polymorphisms within the *COL5A2*, *COL5A3* and *COL3A1* genes and the *BstUI* RFLP within the *COL5A1* gene in modulating risk of AT (Chapter 4)

University of Cape Town

- To test the gene-gene interactions between polymorphisms within genes encoding components of the collagen fibril and components of cell-signalling pathways within the ECM in the modulation of risk of AT (Chapter 5)

The secondary aim of this thesis was to develop and test preliminary models assessing genetic risk of developing AT. The objectives of the specific studies which addressed this secondary aim of the thesis were:

- To investigate the relative contribution of polymorphisms within genes encoding components of the collagen fibril and components of cell-signalling pathways within the ECM to overall genetic risk in a polygenic risk model (Chapter 5)
- To develop and evaluate a preliminary, clinically relevant genetic risk assessment model for AT using (i) polymorphisms that have previously been independently associated with AT in two populations, and (ii) polymorphisms with a definitive AT risk genotype that were implicated in this thesis in a haplotype or gene-gene interaction (Chapter 5)

University of Cape Town

CHAPTER 2: VARIANTS WITHIN THE *COMP* AND *THBS2* GENES ARE NOT ASSOCIATED WITH ACHILLES TENDINOPATHY

This chapter has been submitted for publication in a condensed format as:

Saunders, C.J., Van Der Merwe, L., Cook, J., Handley, C.J., Collins, M., & September, A.V. Variants within the *COMP* and *THBS2* genes are not associated with Achilles tendinopathy in a case-control study of South African and Australian populations. Journal of Sports Sciences. *In press.*

2.1. INTRODUCTION

As reviewed in chapter one, the Achilles tendon is exposed to high mechanical forces during exercise and is often a site of injury in both athletes and the general population.[68,132] The pathology of AT has been well characterised with clearly documented clinical and imaging diagnostic criteria (Table 1.2),[69,128,141] however the exact aetiology of the condition remains undefined. The heritable factors contributing to risk of developing AT have been extensively reviewed in chapter two, and include polymorphisms in several genes coding for structural and associated proteins of the extracellular matrix.[197,199,205,224,231,247,255,257]

In this study we investigated two further candidate genes, *THBS2* and *COMP*, which code for thrombospondin 2 (TSP-2) and cartilage oligomeric matrix protein/thrombospondin 5 (COMP; TSP-5) respectively. The thrombospondins are calcium binding glycoproteins that can be divided into two subfamilies with distinct biological roles.[7,36] TSP-1 and TSP-2 (Group A) are matricellular proteins that form homotrimers and modulate cell functions and cell-matrix interactions,[36] while TSP-3 to TSP-5 (Group B) are structural proteins of the extracellular matrix and form homopentamers.[7,36,37] The thrombospondin proteins have five domains (Figure 2.1).[44] The NH_2-terminal domain is a globular β-sandwich present in all thrombospondins except COMP. It provides a structural framework and contains several binding sites for various ligands. These binding sites may allow the protein to anchor to proteoglycans at this NH_2-terminal while leaving the rest of the protein free to interact with proteins within the cell membrane. The oligomerisation domain is a highly flexible domain

University of Cape Town

present in all thrombospondins and is susceptible to proteolytic activity. The von Willebrand factor type C (vWFC) domain is present only in group A thrombospondins, and has anti-angiogenic activity. Three thrombospondin type-1 repeats (TSR) are also only found in group A thrombospondins. These repeats are normally extracellular and function in cell attachment, the inhibition of angiogenesis and protein-glycosaminoglycan interactions. All thrombospondins contain a highly conserved signature protein domain at the COOH-terminal and a number of disease associated polymorphisms have been mapped to this domain.[44] This domain includes three or more epidermal growth factor (EGF) like repeats, a calcium-binding wire of 13 calcium binding repeats, and a COOH-terminal lectin-like module.

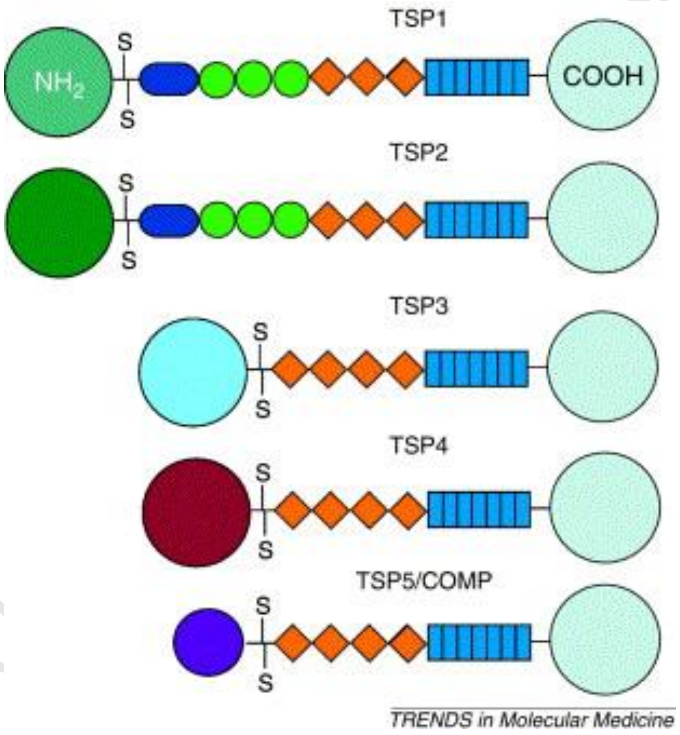

Figure 2.1: Protein structure of the thrombospondin family
Members of the thrombospondin (TSP) family have a multimodular structure. TSP1 and TSP2 are composed of an N-terminal heparin-binding domain (large green circle), a procollagen homology domain (dark-blue oval), three type I repeats (green circles), three epidermal growth factor (EGF)-like type II repeats (orange diamonds), seven calcium-binding type III repeats (pale blue rectangles) and a C-terminal globular domain (blue circle). By contrast, TSP3, TSP4 and TSP5/COMP contain distinct N-terminal domains (light-blue, red, purple circles), four EGF-like repeats, seven type III repeats and a similar C-terminal globular domain. 'S' indicates the position of the cysteine residues involved in multimerization. Sequence homology increases from the N- to the C-termini of these proteins
[Figure and caption reprinted from de Fraipont et al. (2001)[67] with permission from Elsevier]

TSP-2, in particular, is involved in the healing response in connective tissues and plays a significant role in cell-matrix interactions.[7,36,37] It has an anti-angiogenic effect during the late-proliferative and remodelling phase of wound healing.[36] In addition, investigation of *THBS2* null mice has implicated this protein in the regulation of collagen fibrillogenesis and fibril diameter. *THBS2* null mice display a highly variable phenotype which includes skin fragility, lax tendons and increased vascularity.[163] It has been postulated that TSP-2 is a clearance factor for matrix metalloproteinase-2 (MMP-2) and that the increased levels of MMP-2 in *THBS2* null mice may contribute to the observed defects in collagen fibrillogenesis.[36,37] Two SNP's within the *THBS2* gene, rs6422747 and rs9283850, have previously been associated with lumbar disc herniation, a multifactorial musculoskeletal condition, in a Japanese population.[112]

COMP interacts with type I and type III collagens and plays an important role in matrix assembly and the repair of injured tissues.[108,223,243,272] It is expressed in bovine and equine tendons, particularly in growing tendons and in response to mechanical load.[71,269,271] Recent investigation has identified a mechano-responsive element in the 3kb proximal promoter of the human *COMP* gene.[14] In numerous studies of equine tendons, Smith et al.[269,270,271] have shown that COMP levels are correlated to ultimate tensile stress and stiffness of tendon tissue and suggest that the appropriate expression of COMP is necessary for the formation of a functional ECM and the structural integrity of tendons. Additionally, serum COMP has been inversely associated with joint hypermobility,[50] a suggested risk factor for AT.[20,185,287] Genetic variants clustered in the conserved calcium binding domains and carboxy terminal of the COMP protein have repeatedly been associated with skeletal dysplasia's.[7,108,122] In addition, the C-allele of the rs730079 SNP within the 5'UTR of this gene was previously associated with osteoarthritis of the knee in a Caucasian male population.[284]

The observations above, together with the biological roles of these proteins, suggest that both *COMP* and *THSB2* are good candidate genes for association with risk of developing AT. The aim of this study was, therefore, to test the association of sequence variants within the *COMP* and *THBS2* candidate genes with risk of AT in participants from South Africa and Australia.

University of Cape Town

2.2. MATERIALS AND METHODS

2.2.1. PARTICIPANTS

The two groups described below were used in the studies presented in chapter's two to five. Only participants who were of self-reported European Caucasian ancestry were included in order to avoid the effects of population stratification. Participants were given an information sheet regarding the study and then required to sign an Informed Consent Form in accordance with the Declaration of Helsinki [Appendix A]. Once recruited, participants were required to complete a detailed questionnaire recording personal particulars, medical history and injury history [Appendix A]. Participants were allocated a unique coding number to ensure anonymity of DNA samples and confidentiality of all recorded information.[79] Sample sizes were determined using QUANTO version 0.5 (Figure 2.2).[97,98] The South African group included 131 physically active control (SA-CON) participants without any history of tendon or ligament pathology, as well as 94 physically active affected participants (SA-TEN) clinically diagnosed with chronic Achilles tendinopathy (AT). The SA-TEN participants were invited to participate through the medical practice at the Sports Science Institute of South Africa, Cape Town, South Africa. The presence of gradual progressive pain over the posterior lower leg for more than six months (n=94) together with at least one of the following clinical criteria was used to diagnose chronic AT: (i) early morning pain (n=32), (ii) early morning stiffness (n=53), (iii) history of swelling over the Achilles tendon area (n=36), (iv) tenderness to palpation over the Achilles tendon (n=74), (v) palpable nodular thickening over the affected Achilles (n=28), and/or (vi) a positive shift test (movement of the nodular area with plantar-/dorsi-flexion)(n=22).[141,250] The diagnosis of each participant was reviewed by the same clinician (Appendix A), and was confirmed using soft tissue ultrasound in a sub-group of 36 participants.

The Australian group included 209 apparently healthy control (AUS-CON) participants without any history of tendon or ligament pathology, as well as 85 affected participants (AUS-TEN) clinically diagnosed with chronic AT. The AUS-TEN participants were recruited by the Musculoskeletal Research Centre at La Trobe University in Melbourne, Australia. The diagnosis was made using similar clinical criteria to those described for the SA-TEN group, and was confirmed by soft tissue ultrasound examination. The Achilles tendons of the AUS-

University of Cape Town

CON participants were also examined under ultrasound and were confirmed not to have any evidence of asymptomatic AT. The same sports physiotherapist confirmed the diagnosis for all AUS participants.

Participants who had a self-reported history of treatment with either fluoroquinolone antibiotics or local corticosteroid injections were excluded from the study to reduce confounding. In addition, participants who had previously suffered, or were currently suffering from any connective tissue disorders or other systemic diseases associated with AT were also excluded from the study. These included EDS, rheumatoid arthritis, benign joint hypermobility syndrome, diabetes mellitus, familial hypercholesterolaemia, systemic lupus erythematosus, hyperparathyroidism and renal insufficiency.

e Town

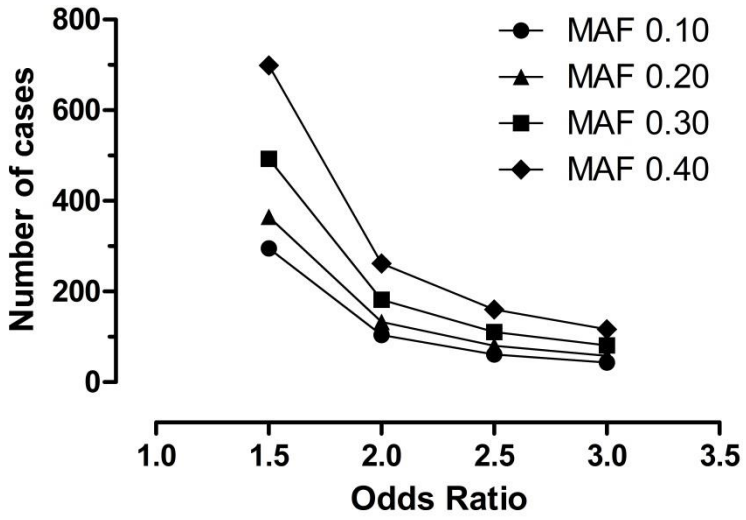

Figure 2.2: QUANTO determined sample sizes
Sample sizes were determined at a 10% prevalence of AT, with 80% statistical power. Number of controls was determined in a 2:1 ratio with cases. MAF: Minor allele frequency

2.2.2. DNA EXTRACTION

Approximately 4.5ml of venous blood was collected from each participant into an ethylenediaminetetraacetic acid (EDTA) vacutainer tube by venipuncture of a forearm vein. The SA samples were stored at 4˚C until DNA extraction as described by Lahiri et al.

(1999)[165] and modified by Mokone et al. (2005)[197] Briefly, blood samples were transferred to sterile 15ml polypropylene tubes to which 2 volumes of TKM-1 buffer (10mM Tris-HCl pH 7.6, 10mM KCl, 10mM $MgCl_2$, 2mM EDTA) containing 2.5% Nonidet P-40 was added to lyse the red blood cells. After incubation at room temperature for 10 minutes, the white blood cells were pelleted by centrifugation at 1200Xg and washed at least twice with one volume of TKM-1 buffer. The pellets were resuspended in 800µl of TKM-2 buffer (10mM Tris-HCl pH 7.6, 10mM KCl, 10mM $MgCl_2$, 0.4M $NaCl_2$, 2mM EDTA) containing 50µl of 10% SDS solution, and incubated at 55°C for 10 minutes or until the pellet was dissolved. One hundred and fifty (150) µl of 5M $NaClO_4$ and 500µl of chloroform were added to each sample and mixed by vortexing for 15-20 seconds. The samples were then transferred to sterile 1.5ml microfuge tubes and centrifuged at 2000Xg for 5 minutes at room temperature to precipitate the protein. The supernatants were transferred to new, sterile microfuge tubes containing 1ml of absolute ethanol and centrifuged at 2000Xg for 5 minutes at room temperature to precipitate the DNA. The supernatants were removed and the pellets were air dried and resuspended in 100µl of 1X TE buffer (10mM Tris-Cl, 1mM EDTA, pH 8.0) before incubation at 65°C for 15 minutes. DNA from the Australian participants was extracted as per the manufacturer's recommendations using a sequenced extraction technique (FlexiGene DNA Kit, *Qiagen P/L*, Valencia, California, USA).[255] The DNA was stored at 4°C until subsequent genotyping.

2.2.3. VARIANT SELECTION

Lists of SNPs within the *COMP* and *THBS2* genes were generated using the National Centre for Biotechnology Information's (NCBI) SNP database available at http://www.ncbi.nlm.nih.gov/snp/ and were interrogated using the Ensembl Genome Browser (http://www.ensembl.org/index.html) and Genome Variation Server (GVS)(http://gvs.gs.washington.edu/GVS137/).[283] Three SNPs within the *THBS2* gene (chr6q27) were selected for investigation (Figure 2.3). SNPs rs9283850 (Intron 9), rs6422747 (Intron 14) and rs9505888 (Intron 17) are all A>G substitutions and were selected based on their heterozygosity score (>35%), minor allele frequency (>20%), linkage and previous association with other clinical phenotypes.[112] The rs9505888 polymorphism was selected based on its linkage disequilibrium (LD) with rs6422747. In addition, all three SNP's were identified as Tag SNP's.[283] Tag SNPs are representative SNPs in genomic regions that

University of Cape Town

are in high LD and these SNPs should therefore provide moderate coverage of this 38,3kb gene. Two SNPs within the *COMP* gene (chr19p13.11), rs730079 (-1417C>G) and rs28494505 (Intron 18, A>G), were investigated based on their heterozygosity score (>35%), minor allele frequency (>20%) and previous association with clinical phenotypes (Figure 2.4).[284] The rs28494505 polymorphism was selected based on its location within the highly conserved COOH-terminal domain of the COMP protein and its identification as a Tag SNP.[283]

2.2.4. GENOTYPING

Polymorphisms rs9505888, rs6422747 and rs28494505 were amplified using standard polymerase chain reaction (PCR) and genotyped using restriction fragment length polymorphism (RFLP) techniques. Unless otherwise specified, PCR reactions were carried out in a solution containing 200ng DNA template, 1X reaction buffer, 2.0mM Mg^{2+}, 2.5mM each of dATP, dCTP, dTTP and dGTP, 20pmol each of the forward and reverse primers and 0.5 units of either *Taq* DNA polymerase (*New England Biolabs® Inc.,* Ipswich, Massachusetts, USA) or SuperTherm *Taq* DNA polymerase (*Hoffmann-La-Roche*, US) using an XP Thermal Cycler Block (*Bioer Technology Co.*, Middlesex, UK). Reaction buffer was specific to the *Taq* used. Thermo-Pol® reaction buffer (*New England Biolabs® Inc.,* Ipswich, Massachusetts, USA) consisted of 10mM KCl, 10mM $(NH_4)SO_4$, 20mM Tris-HCl, 2mM $MgSO_4$ and 0.1% Triton X-100 at pH8.8. SuperTherm reaction buffer (*Hoffmann-La-Roche*, US) was provided with the SuperTherm *Taq* DNA polymerase and used as per manufacturer's instructions. PCR reactions were initiated with a five minute denaturing cycle at 94˚C, followed by 35 cycles consisting of 30 seconds denaturation at 94˚C, 30 seconds annealing at variant specific temperatures and 40 seconds extension at 72˚C with a final five minute extension step at 72˚C. Variant specific oligonucleotide primer sequences, reaction conditions and amplicon sizes are listed in Appendix B. PCR amplicons of polymorphisms rs9505888, rs6422747 and rs28494505 were digested overnight with *BstUI, MfeI* and *MspI* restriction endonucleases respectively. Restriction fragments were resolved, together with a 100 base pairs (bp) size standard (*Promega Corporation*, Madison, Wisconsin, USA), by 6% polyacrylamide gel electrophoresis (PAGE) at 120V for 2 hours and visualised under UV light after staining with SYBR® Gold nucleic acid gel stain (*Invitrogen Molecular Probes*™, Oregon, USA) [Appendix B].

University of Cape Town

Polymorphisms rs9283850 and rs730079, as well as a subset of samples for rs28494505, were genotyped using inventoried TaqMan® SNP Genotyping Assays (*Applied Biosystems™*) that were amplified and distinguished using the StepOnePlus™ Real-Time PCR System (*Applied Biosystems™*). Briefly, real-time PCR reactions were performed in a final volume of 8µl and included approximately 20ng DNA, TaqMan® Assay Mix and TaqMan® Genotyping Master Mix (2X). TaqMan® Assay Mix consisted of sequence specific forward and reverse primers to amplify the polymorphic sequence, and two MGB probes labelled with VIC® and FAM™ dyes to detect allele 1 and allele 2 respectively. Real-time PCR cycling started with an initial hold step at 95°C for 10 minutes followed by 40 cycles of 92°C for 15 seconds, and 60°C for 1 minute. Fluorescence signal for each MGB probe was used to discriminate alleles and genotypes using the StepOnePlus™ Real-Time PCR System software [Appendix B]. For quality control purposes, a number of positive controls of known genotype and DNA-free controls were randomly included on every 96-well PCR plate. In addition, a subset of samples (approximately 12%) was genotyped twice using the same methodology to ensure genotyping was consistent.

It should be noted that the DNA from the AUS and SA samples was extracted in different laboratories and therefore suspended in different buffers. It is suspected that a reagent used either in the DNA extraction process, or in the suspension buffer of the AUS DNA samples interferes with PCR in certain samples. In order to avoid genotyping errors, samples that failed twice to amplify during PCR for a particular SNP were considered as unsuccessfully genotyped and no further attempts were made to genotype them at that SNP locus. Samples that consistently failed to amplify during PCR were excluded in further studies. The number of successfully genotyped samples for each SNP investigated in this study is shown in Table 2.1 with similar tables in subsequent chapters.

University of Cape Town

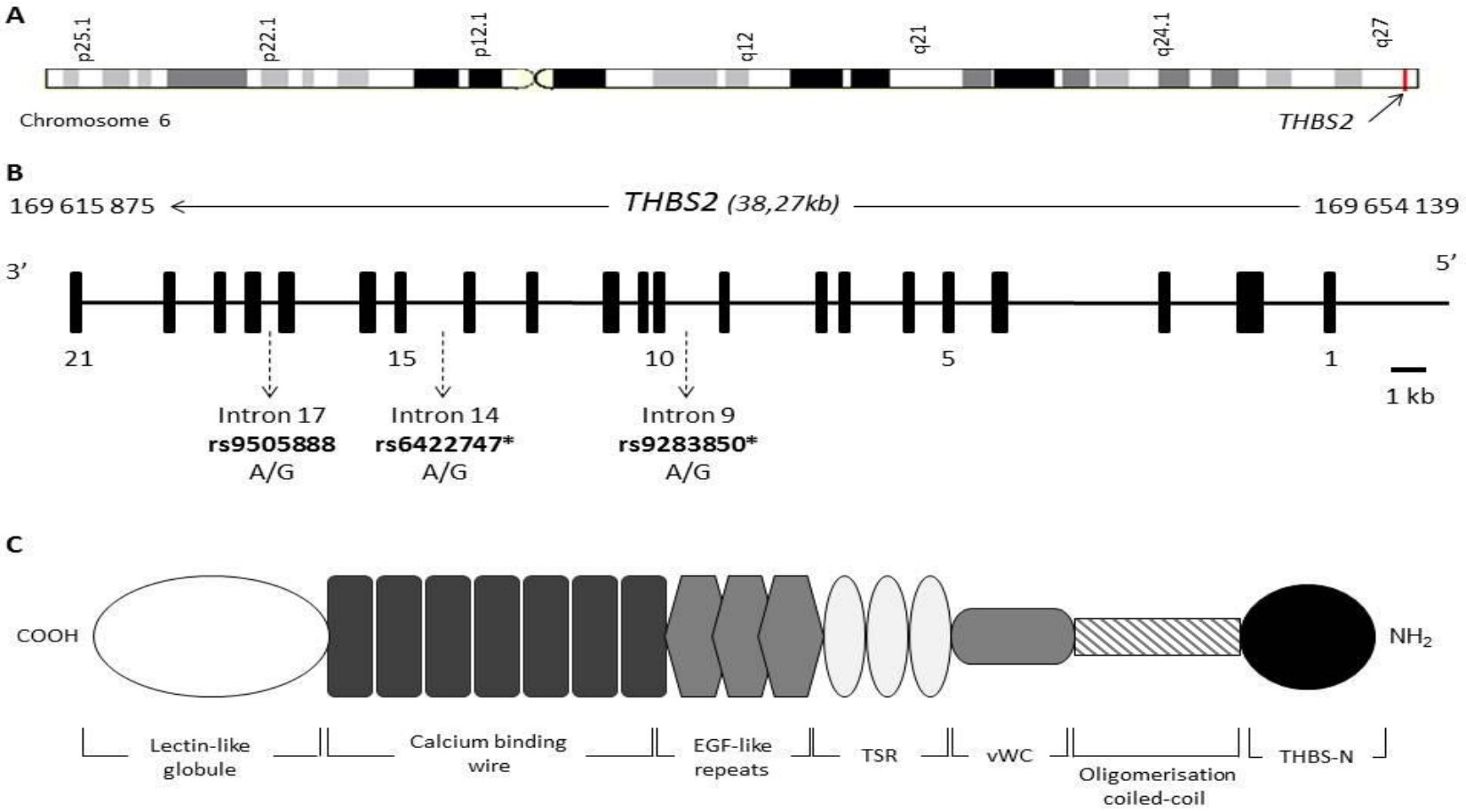

Figure 2.3: Thrombospondin 2

(A) *Shaded bands indicate cytogenetic banding patterns for chromosome 6. Chromosomal banding position of THBS2 gene is indicated with an arrow* ***(B)*** *Genomic location and intron/exon structure of THBS2 including previously cited variants #Significant association with lumbar disc herniation [Hirose et al. (2008)[112]] Solid bars represent numbered exons separated by introns (solid lines). Variants in bold are included in the present study.* ***(C)*** *TSP-2 protein domain architecture. EGF: Epidermal growth factor; TSR: Thrombospondin repeat; vWC: von Willebrand factor type-C; THBS-N: Thrombospondin amino terminal [Compiled from* www.ensembl.org *and Carlson et al. (2008)[44]]*

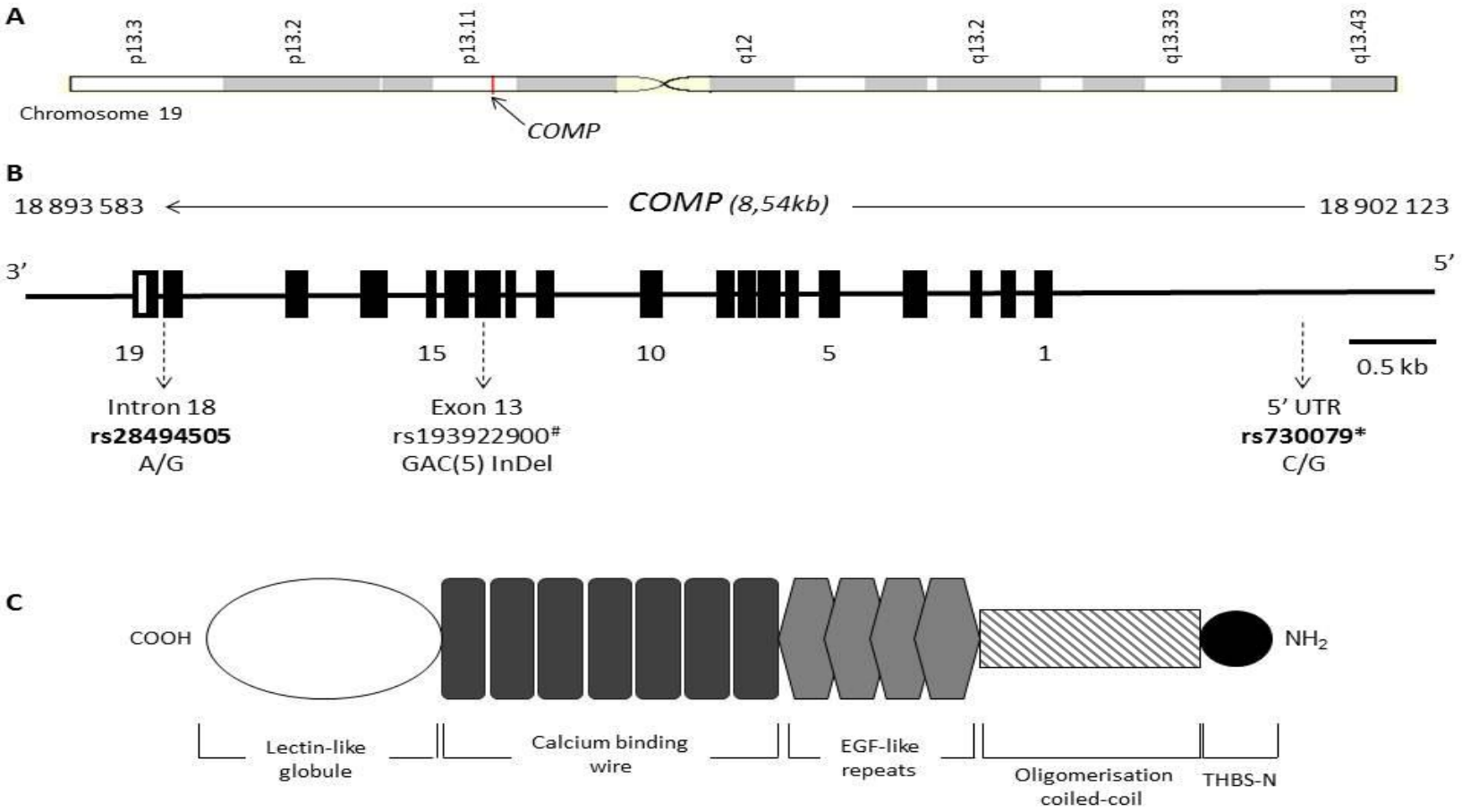

Figure 2.4: Cartilage oligomeric matrix protein

(A) *Shaded bands indicate cytogenetic banding patterns for chromosome 19. Chromosomal banding position of COMP gene is indicated with an arrow* ***(B)*** *Genomic location & intron/exon structure of COMP *Associated with osteoarthritis (Valdes et al. (2007))[284] #Associated with PSACH & MED [Delot et al. (1999)[70]] Solid bars represent numbered exons separated by introns (solid lines). Variants in bold are included in the present study* ***(C)*** *COMP protein domain architecture. UTR: Untranslated region; EGF: Epidermal growth factor; THBS-N: Thrombospondin N-terminal domain [Compiled from* www.ensembl.org *and Carlson et al. (2008)[44]]*

Table 2.1: The number and proportion of recruited participants successfully genotyped for each of the *THBS2* and *COMP* polymorphisms

Total number of participants		Number and proportion of participants genotyped									
		rs28494505		rs730079		rs9505888		rs6422747		rs9283850	
AUS-CON	**209**	165	79%	205	98%	185	89%	200	96%	203	97%
AUS-TEN	**85**	62	73%	83	98%	75	88%	82	96%	79	93%
TOTAL	**294**	**227**	**77%**	**288**	**98%**	**260**	**88%**	**282**	**96%**	**282**	**96%**
SA-CON	**131**	131	100%	131	100%	130	99%	130	99%	130	99%
SA-TEN	**94**	94	100%	93	99%	93	99%	93	99%	93	99%
TOTAL	**225**	**225**	**100%**	**224**	**99%**	**223**	**99%**	**223**	**99%**	**223**	**99%**

University of Cape Town

2.2.5. STATISTICAL ANALYSES

The programming environment R,[230] and R package, genetics,[296] was used to estimate genotype and allele frequencies and exact Hardy-Weinberg equilibrium (HWE) probabilities. Age and sex were considered potential confounders and were, therefore, corrected for in the logistic regression model of the association of genotypes and alleles with AT. No adjustments were made for multiple testing because it has been suggested that these corrections, such as Bonferroni, markedly overcorrect for an inflated false-positive rate and unnecessarily reduce power in genetic association studies where background LD exists between SNPs.[207,220]

Genotype effects on age, height, weight, body mass index (BMI) and sex were determined by a one-way analysis of variance (ANOVA) using STATISTICA version 10 (*StatSoft Inc.*, Tulsa, OK, USA). Statistical significance was accepted when $P<0.05$.

2.3. RESULTS

2.3.1. DESCRIPTIVE CHARACTERISTICS

The physiological characteristics of the SA-CON, SA-TEN, AUS-CON, AUS-TEN and the combined groups are given in Table 2.2. There were no significant differences between age at recruitment (CON) and age at injury (TEN) in either the AUS (P=0.268), SA (P=0.167) or combined (SA & AUS) (P=0.099) groups. There was no significant difference in the sex distribution between SA-CON and SA-TEN (P=0.163), however the AUS-CON and combined CON groups contained proportionally less males than the AUS-TEN (P<0.001) and combined TEN (P<0.001) groups respectively. There was no difference in height between the SA-CON and SA-TEN groups (P=0.212), however the AUS-TEN and combined TEN groups were significantly taller than the AUS-CON (P=0.046) and combined CON (P=0.006) groups respectively. When adjusted for sex, this difference in height was no longer significant between either the AUS (P=0.468) or combined (P=0.337) groups. The AUS-TEN, SA-TEN and combined TEN participants weighed more at recruitment than the AUS-CON (P<0.001), SA-CON (P<0.001) and combined CON (P<0.001) participants respectively. Consequently, the AUS-TEN, SA-TEN and combined TEN participants had a higher BMI than the AUS-CON (P<0.001), SA-CON (P<0.001) and combined CON (P<0.001) participants respectively. When adjusted for age at recruitment, height and sex, there was no significant difference in weight between the AUS-CON and AUS-TEN (P=0.682), nor the combined CON and combined TEN (P=0.077) groups. However, the SA-TEN group was still heavier than the SA-CON after adjustment (P<0.001). It should be noted that the TEN groups could not be adjusted for age at injury as weight was measured at recruitment and not at time of injury. After adjusting for age at recruitment and sex, BMI was no longer different between the two AUS (P=0.661) or combined (P=0.095) groups, but was still different between SA-CON and SA-TEN (P<0.001). The proportion of participants born in Australia (born “here”) was significantly different between the AUS-CON (84.9%) and AUS-TEN (78.6%) groups (P=0.047), but there was no difference in the proportion of participants born in South Africa (born “here”) between the SA-CON (81.7%) and SA-TEN (82.2%) groups (P=0.304).

University of Cape Town

The sporting and medical history of the SA participants has been analysed and reported previously.[197,198,199,255] Briefly, the majority of both SA-CON (75.9%) and SA-TEN (58.7%) participants were active runners. SA-CON and SA-TEN participants had a similar number of years running history (SA-CON: 8.0 ± 8.0 years (n=129) vs SA-TEN: 9.0 ± 10.3 years (n=75); P=0.423), but SA-CON participants ran for more hours per week in the two years prior to recruitment (SA-CON: 3.3±2.9 hours, n=125) than the SA-TEN participants ran in the two years prior to injury (SA-TEN: 2.5±2.8 hours, n=66)(P=0.046). There was a significant difference in the number of years spent playing high-impact sports between the SA-CON and SA-TEN groups (SA-CON: 16.5±16.1 years (n=129) vs SA-TEN: 31.1±29.5 years (n=75); P<0.001) but not for the number of hours per week spent playing high-impact sports in the two years prior to injury (SA-TEN: 4.4±5.2 hours, n=75) or recruitment (SA-CON: 5.1±4.8 hours, n=129)(P=0.321). The proportion of current (SA-CON: 3.2% vs SA-TEN: 5.4%) and ex-smokers (SA-CON: 21.4% vs SA-TEN: 25.7%) was not significantly different between the two groups (P=0.480). The physical activity and smoking history of the Australian participants was not documented and this is recognised as a limitation to this study.

The majority of SA-TEN participants suffered from mid-portion AT (71.0%) with a small proportion suffering from AT in the distal (27.4%) and upper third of the tendon (1.7%). This distribution was similar in the AUS-TEN participants in which 91.8% of participants suffered from mid-portion AT and only 5.9% of participants suffered from AT in the distal tendon. There was no difference in the distribution of AT between right (AUS: 24.7%; SA: 26.7%) or left (AUS: 23.5%; SA: 27.9%) legs, however, a large proportion of TEN participants suffered from bilateral AT (AUS: 51.8%; SA: 45.6%).

2.3.2. GENOTYPE AND ALLELE FREQUENCIES

Genotype and minor allele frequency distributions for each of the polymorphisms, together with the HWE P-values, are shown in Table 2.3. The frequency distributions of the SNPs tested in this study were found to differ significantly between the South African and Australian participant groups at the rs28494505 (P<0.001) and rs6422747 loci (P=0.024). The participants from the two groups were therefore summarised separately. All genotype distributions were in HWE except at the rs28494505 locus that deviated significantly in the AUS-CON group (P=0.043). It was however noted that a lower genotype call rate was

University of Cape Town

observed for this polymorphism in the AUS samples (Table 2.1). There were no significant genotype effects on age, weight, height, BMI or sex (Table 2.4). After adjusting for the potential confounder's age, sex and born "here", there were no significant differences in either the genotype or allele frequency distributions between the CON and TEN groups.

Table 2.2: Descriptive characteristics of the Australian and South African participants

	AUSTRALIA			SOUTH AFRICA			COMBINED		
	CON	**TEN**	**P-VALUE**	**CON**	**TEN**	**P-VALUE**	**CON**	**TEN**	**P-VALUE**
AGE (years)	38.5±12.4 (205)	40.4±14.2 (84)	0.268	37.3±10.6 (128)	39.7±14.1 (84)	0.167	38.1±11.7 (333)	40.0±14.1 (168)	0.099
SEX (% male)	40.2 (84)	72.9 (62)	**<0.001**	63.2 (84)	72.0 (67)	0.163	49.1 (168)	72.5 (129)	**<0.001**
HEIGHT (m)	171.5±9.2 (207)	174.0±9.9 (82)	**0.046** (0.468)[a]	174.8±9.3 (127)	176.4±9.0 (84)	0.212 (0.289)[a]	172.8±9.3 (334)	175.2±9.5 (166)	**0.006** (0.337)[a]
WEIGHT (kg)	72.7±14.2 (208)	80.5±14.0 (85)	**<0.001** (0.682)[b]	71.3±12.0 (131)	77.7±13.5 (87)	**<0.001** (**<0.001**)[b]	72.2±13.4 (339)	79.1±13.8 (172)	**<0.001** (0.077)[b]
BMI ($kg.m^{-2}$)	24.6±3.9 (207)	26.5±3.8 (82)	**<0.001** (0.661)[c]	23.3±2.7 (127)	24.8±3.3 (84)	**<0.001** (**<0.001**)[c]	24.1±3.5 (334)	25.6±3.7 (166)	**<0.001** (0.095)[c]
COUNTRY OF BIRTH (% Born "here")	84.9 (174)	78.6 (66)	**0.047**	81.7 (107)	82.2 (74)	0.304	n/a	n/a	n/a

Values are means ± SD (age, height, weight, BMI) or frequencies (sex, country of birth) with n in parentheses. Height, weight & BMI measured at recruitment. Combined: AUS and SA participants; Age: Age at injury (TEN) or recruitment (CON); Country of birth: proportion of participants born in Australia (AUS) or South Africa (SA)

P-values in parentheses are co-varied for [a]sex, [b]age, sex and height, and [c]age and sex. Bold P-values are <0.05

University of Cape Town

Table 2.3: Genotype and minor allele frequency distributions of the sequence variants investigated within the *COMP* and *THBS2* genes in control (CON) and Achilles tendinopathy (TEN) groups of Australian (AUS) and South African (SA) populations

	CON		TEN		P-Values	
	AUS	**SA**	**AUS**	**SA**	**Country**	**Group**
COMP:						
rs28494505						
N	165	131	62	94		
A/A	0.80	0.54	0.68	0.62	**<0.001**	0.994
A/G	0.17	0.43	0.26	0.36		
G/G	0.03	0.03	0.06	0.02		
G	0.12	0.24	0.19	0.2	**<0.001**	0.919
HWE	**0.043**	0.097	0.212	0.345		
rs730079						
N	205	131	83	93		
G/G	0.34	0.38	0.39	0.35	0.687	0.727
G/C	0.50	0.43	0.45	0.47		
C/C	0.16	0.19	0.17	0.17		
C	0.41	0.40	0.39	0.41	0.978	0.455
HWE	0.773	0.206	0.644	0.832		
***THBS2*:**						
rs9505888						
N	185	130	75	93		
G/G	0.26	0.33	0.32	0.32	0.602	0.244
G/A	0.55	0.52	0.44	0.47		
A/A	0.19	0.15	0.24	0.20		
A	0.46	0.41	0.46	0.44	0.315	0.533
HWE	0.236	0.467	0.354	0.679		
rs6422747						
N	200	130	82	93		
A/A	0.42	0.34	0.34	0.35	**0.024**	0.645
A/G	0.51	0.43	0.50	0.47		
G/G	0.08	0.23	0.16	0.17		
G	0.33	0.45	0.41	0.41	0.055	0.480
HWE	0.058	0.156	0.823	0.832		
rs9283850						
N	203	130	79	93		
A/A	0.28	0.28	0.25	0.25	0.965	0.604
A/G	0.43	0.48	0.53	0.46		
G/G	0.29	0.23	0.22	0.29		
G	0.50	0.47	0.48	0.52	0.889	0.632
HWE	0.068	0.728	0.655	0.533		

P-values are for the difference between countries and between diagnostic groups respectively, adjusted for each other, age at injury/recruitment, sex and whether or not a person was investigated in his/her country of birth. The genotype P-value is calculated using a 2-degrees of freedom test, with genotypes as categories and the allelic P-value is calculated using an additive allelic model. HWE gives exact P-values from tests of Hardy–Weinberg equilibrium. N is number of samples genotyped Bold P-values are <0.05.

University of Cape Town

Table 2.4: Genotype effects of *COMP* and *THBS2* polymorphisms on physiological characteristics of participants

POLYMORPHISM	P-VALUES					
	AGE	AGE INJURED	HEIGHT	WEIGHT	BMI	SEX
rs28494505	0.177	0.450	0.445	0.855	0.722	0.446
rs730079	0.457	0.370	0.996	0.992	0.857	0.663
rs9505888	0.486	0.296	0.343	0.094	0.073	0.121
rs6422747	0.233	0.098	0.442	0.711	0.778	0.061
rs9283850	0.939	0.934	0.844	0.381	0.313	0.285

Age: Age at recruitment (CON & TEN); Age Injured: Age at recruitment (CON) & age at diagnosis (TEN); P-values for age, age injured, height, weight and BMI are determined by one-way ANOVA; P-value for sex is determined by Pearsons χ^2 analysis

University of Cape Town

Pairwise linkage disequilibrium (LD) using D' was calculated between the two *COMP* polymorphisms, and between the three *THBS2* polymorphisms, and is presented in Table 2.5. LD is the non-random inheritance of alleles at two or more genomic loci. D' is calculated by comparing the frequency of haplotypes of polymorphisms with the product of the individual allele frequencies. If they are equal, D' will be equal to zero and will indicate that these polymorphisms are not in LD. If they differ, D' will not equal zero and will indicate that these polymorphisms are in LD. The closer the D' value is to one, the tighter the LD. Pairwise LD analysis between the two *COMP* polymorphisms indicates that rs730079 and rs28494505 are in tight LD within this Australian group, but only moderately so in this South African group. The same analysis between the three *THBS2* polymorphisms indicates that the rs6422747 and rs9505888 polymorphisms are in moderate LD in both groups.

Table 2.5: Pairwise linkage disequilibrium (LD) between polymorphisms using D'

GROUP	*COMP* Polymorphisms		*THBS2* Polymorphisms		
		rs730079		rs6422747	rs9283850
AUS CON	rs28494505	0.77	rs9505888	0.59	0.36
			rs6422747	n/a	0.20
AUS TEN	rs28494505	1.00	rs9505888	0.32	0.06
			rs6422747	n/a	0.08
SA CON	rs28494505	0.38	rs9505888	0.30	0.43
			rs6422747	n/a	0.19
SA TEN	rs28494505	0.52	rs9505888	0.52	0.23
			rs6422747	n/a	0.21

University of Cape Town

2.4. DISCUSSION

The primary finding of this study is that no association of the selected polymorphisms within the *COMP* and *THBS2* candidate genes with Achilles tendinopathy (AT) was detected in either of the two groups studied. Although this study was only sufficiently powered to detect large effects on risk of developing AT (OR>2),[97,98] and was therefore unlikely to detect any risk variant with a small effect, it should be noted that there was no tendency towards association with AT for any of the polymorphisms investigated. We cannot exclude the possibility that other polymorphisms within these genes are associated with AT, however, all the SNPs investigated in this study were identified as Tag SNPs and it is, therefore, unlikely that any SNPs tightly linked to these would show a significant effect on risk of AT. Both *THBS2* (38,3kb) and *COMP* (8,5kb) are small genes and the Tag SNPs selected span 11,6kb and 9,9kb respectively, therefore giving moderate coverage of both genes.

Previous work in this area has shown that genetic variants in genes encoding structural components of the collagen fibril often have large effects on risk of developing AT (OR>2).[197,199,231] Furthermore, rare disease-causing mutations within genes encoding these structural components have been implicated in a range of severe connective tissue

disorders including pseudoachondroplasia, multiple epiphyseal dysplasia, osteogenesis imperfecta and Ehlers-Danlos syndrome.[108,122,183,193,222,268] Based on these observations it has been suggested that there is limited redundancy within the biological mechanisms leading to collagen fibrillogenesis, and that common polymorphisms in the associated genes may have large effects on risk of developing mild connective tissue pathology, such as AT.[236]

In particular, the carboxy terminal signature domain of all thrombospondins, which contains the EGF-like repeats, type III calcium binding repeats and a carboxy terminal lectin-like globule, is highly conserved between the five members of this protein family.[7,44] Within the *COMP* gene, mutations in this signature domain result in a spectrum of clinical manifestations from mild to severe skeletal dysplasia's.[44,122] There is evidence to suggest that variations in this domain alter protein structure and folding, and disrupt calcium binding. [49,108,120,161,177] As suggested above, it is therefore possible that the biological role of COMP is non-redundant, and that sequence variants which have even a minor effect on transcription, translation or protein structure and function will result in at least a mild inherited clinical dysplasia. Taking into account the high sequence conservation of both the *COMP* and *THBS2* genes between species, it remains possible that rare sequence variants (minor allele frequency 0.1-3.0%) within these genes may be associated with a large effect on risk of AT.[34,212] Such rare polymorphisms often have a larger effect on risk as well as greater penetrance, making them good targets for prophylactic treatments.[34] For example, the rare TT genotype (<5% in Caucasian populations) of the functional Sp1-binding site polymorphism (rs1800012; G>T) within intron 1 of the *COL1A1* gene is associated with reduced risk of acute soft tissue ruptures.[58,81,152,225] Associations of common multifactorial diseases with rare polymorphisms are unlikely to be found by either genome wide association studies or case-control association studies such as the current study, and further work would therefore be necessary to test this hypothesis. [34]

This study, and those presented in chapters four to six, have a number of limitations. Firstly, the TEN participants were significantly heavier at recruitment than the CON participants. Given the possible association between adiposity and AT (reviewed in chapter one), this is recognised as a limitation to this study, as well as those presented in chapter's three to five. Although weight of TEN participants at time of injury could not be obtained, several TEN

University of Cape Town

participants reported weight gain due to a decrease in physical activity after developing AT. In support of this, Gaida et al.[94] state that 40% of participants with AT reported an average of 5kg weight gain after onset of symptoms. In the AUS participants, this difference may also be explained by the higher frequency of males in the AUS-TEN group as the difference was no longer significant when adjusted for age at recruitment, sex and height. Secondly, another limitation to the studies presented in this thesis is the lack of physical activity and smoking data for the Australian participants. Lastly, the low genotype call rate for some polymorphisms in the AUS participants is a limitation. In this particular study, the genotype call rate was low for rs28494505 in the AUS participants (77%). As discussed in the description of the genotyping techniques, this may be due to chemical interference with PCR in the AUS DNA samples. This low genotype call rate is reflected in the deviation from HWE in the AUS-CON participants, and the difference in genotype frequencies between the SA and AUS groups at this locus. HWE is a measure of Mendelian inheritance and in a large, randomly mating population, alleles which are not subjected to any selective pressures will distribute according to the principles of HWE.[105] Deviations from HWE at any particular locus result from selective pressure and non-random mating within a population, but may also indicate technical problems such as genotyping errors and assay non-specificity.[119,168] In the studies presented in this thesis, deviations from HWE may be explained by (i) the selection of participants and, (ii) possible genotyping errors. In the first instance, the inclusion and exclusion criteria adhered to during recruitment of participant's results in a study sample that is not a true reflection of the general population as a whole, but rather represents two small and specific samples of this population. The TEN participants, in particular, are a highly selected group but the exclusion criteria also result in selection of the CON participants. In the second instance, it is recognised that genotyping errors may have occurred. However, several steps were taken in all studies included in this thesis to ensure that potential contamination and genotyping errors would be detected. These steps are described in the methods section of each study but include repeat genotyping of subsets of samples as well as the inclusion of positive (known genotype) and DNA-free controls. It can, therefore, be reasonably assumed that genotyping errors are not the cause of deviations from HWE observed in these studies.

University of Cape Town

2.5. CONCLUSION

In conclusion, this study did not identify an association between the common sequence variants investigated and increased risk of AT. However, it remains possible that other sequence variants (rare or common) within the *COMP* and *THBS2* genes may be associated with risk of AT.

University of Cape Town

CHAPTER 3: HAPLOTYPE ANALYSIS OF *COL27A1* & *TNC* CANDIDATE GENES AS RISK FACTORS FOR ACHILLES TENDINOPATHY

This chapter has been published in a condensed format as:

Saunders, C.J., Van Der Merwe, L., Posthumus, M., Cook, J., Handley, C.J., Collins, M., & September, A.V. Investigation of variants within the *COL27A1* and *TNC* genes and Achilles tendinopathy in two populations. Journal of Orthopaedic Research. 2013. 31(4), 632–637

3.1. INTRODUCTION

As reviewed in chapter one of this thesis, genetic susceptibility has been shown to be a significant intrinsic risk factor for the development of AT. This thesis aims to identify particular candidate genes that may harbour variants contributing to this genetic susceptibility. Initial research in this area linked the *ABO* gene, involved in determining blood type, on the long arm of chromosome 9 (9q34.2) with risk of developing AT[139,146,158] but this association was not repeated in all studies.[170,179] This suggested that another gene in close proximity to the *ABO* gene could be associated with AT instead. As discussed in chapter one, Mokone et al. (2009)[199] and September et al. (2009)[255] reported an association between the rs12722 SNP, also known as the *BstUI* RFLP, within the 3'-UTR of the *COL5A1* gene with chronic AT in participants from South Africa and Australia. The *COL5A1* gene encodes the α1 chain of type V collagen and is mapped to the same locus as the *ABO* gene on the long arm of chromosome 9 (9q34)(Figure 3.1). This rs12722 polymorphism has recently been fine mapped to a functional region within the 3'-UTR of *COL5A1* and is implicated in maintaining the stability of *COL5A1* mRNA.[6,164] In addition, three more polymorphisms within the 3'-UTR and a polymorphism within a microRNA gene (*MIR608*) which binds to this region have been independently associated with chronic AT, and are predicted to alter the secondary structure of the 3'-UTR and *COL5A1* mRNA stability.[6]

University of Cape Town

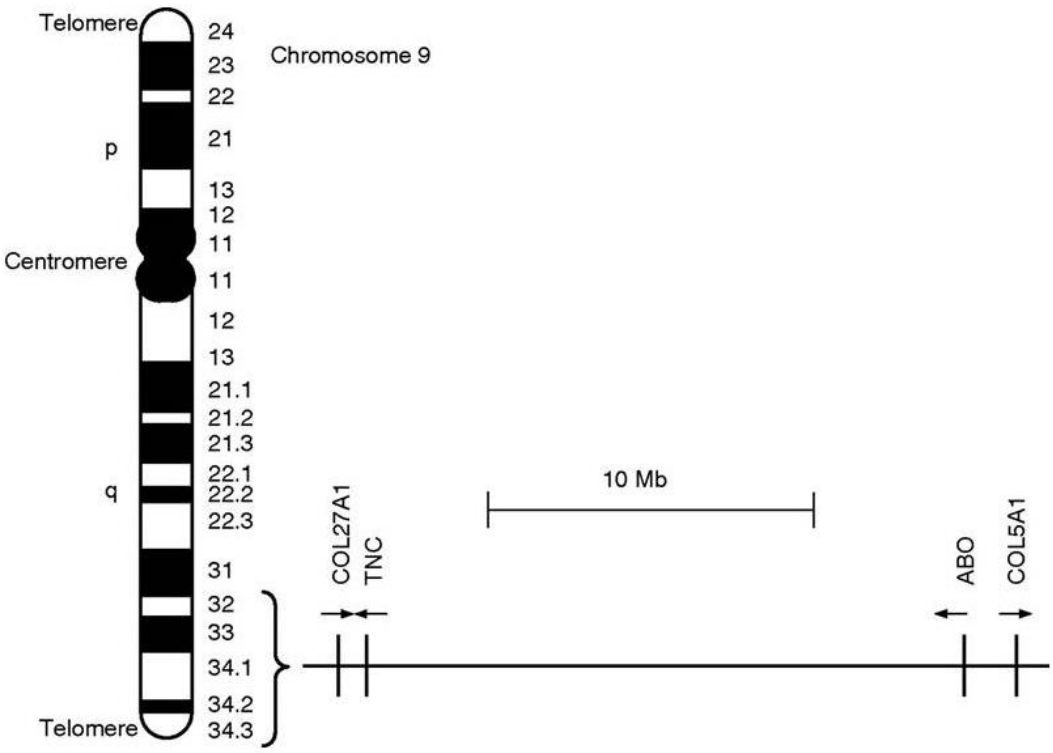

Figure 3.1: Ideogram schematic representation of human chromosome 9

The centromere is the structural region of the chromosome that binds with the nuclear spindle at mitosis and meiosis. The telomere refers to each end region of the chromosome, which consists of tandem repeats of simple DNA sequences. The chromosome is comprised of a short arm (p) and a long arm (q). The dark and light areas reflect the unique banding pattern of chromosome 9 when stained using cytogenetic techniques. The nomenclature of each band is indicated on the right side of the chromosome. A segment corresponding to the telomeric end of the long arm of chromosome 9 (9q32–q34), which encompasses the 155 kb COL27A1, 20 kb tenascin C (TN-C), 20 kb ABO and 201 kb COL5A1 genes, is also shown. The vertical lines represent the relative position of each gene, while the arrows indicate the direction of their transcription. COL27A1 encodes type XXVII collagen; TNC encodes TN-C; ABO encodes the ABO transferases; and COL5A1 encodes the pro-α1(V) chain of type V collagen. Mb, megabase

[Figure and caption reprinted from September et al. (2007)[260] with permission from BMJ Publishing Group Ltd.]

University of Cape

Mokone et al. (2005)[197] also investigated a polymorphism within the *TNC* gene which encodes tenascin-C (TN-C) and is mapped to a region 19,6Mbp upstream of *COL5A1*. A guanine-thymine (GT) dinucleotide repeat polymorphism within intron 17 of the *TNC* gene was shown to be significantly associated with Achilles tendon injuries (ATI). In particular, alleles containing 12 and 14 GT repeats were over-represented in the ATI group (χ^2=21.6, P<0.001), and alleles containing 13 and 17 GT repeats were under-represented in the ATI group (χ^2=42.4, P<0.001). Participants who were heterozygous or homozygous for alleles

containing 13 and 17 GT repeats were 6.2 fold less likely to develop ATI. However, the possibility that this original association, between the GT dinucleotide repeat polymorphism within the *TNC* gene and AT, was actually due to the association of a polymorphism within a neighbouring gene could not be excluded. Given the large genomic distance between the *TNC* and *COL5A1* genes (Figure 3.1), it is unlikely that the association observed with the GT dinucleotide repeat polymorphism in *TNC* reflects that observed in the *COL5A1* gene. Interestingly, the *COL27A1* gene, which encodes the α1 chain of type XXVII collagen, has been mapped to chromosome 9q32-33, 708kbp upstream of *TNC*.[35,216] It was therefore decided to further explore the chromosomal region which included the *TNC* and *COL27A1* genes in order to potentially define a smaller genetic interval associated with AT.

The tenascins (TN) are a family of six glycoproteins – TN-N, TN-R, TN-X, TN-W, TN-Y and TN-C.[136,204] Tenascin-C (TN-C) is a symmetrical hexameric ECM glycoprotein (Figure 3.2 and 3.4) and its structure, function and regulation is expertly reviewed by Jones and Jones (2000).[136] Briefly, each TN-C subunit ranges between 190 and 300kDa in size and contains an NH_2-terminal tenascin assembly (TA) domain at which subunit polypeptide chains are disulphide-linked to form a six-armed hexabrachion. Adjacent to the TA domain is a thin and rigid portion of the arm containing several EGF-like repeats. These repeats are approximately 31 amino acids in length and are important in facilitating intrachain disulphide bonds. This region also has counter-adhesive properties that are particularly important during neuronal development. The distal portion of the arm is thick and flexible and contains a variable number of fibronectin type III (FNIII) domains. These FNIII domains are highly elastic, bind several ECM proteins, glycosaminoglycans as well as cell surface receptors, and are particularly susceptible to proteolytic degradation.[10,136,210] The variable number of these FNIII domains in tenascin-C subunits is a result of alternative mRNA splicing and this may play an important role in the diversity of functions fulfilled by the tenascin-C protein.[136,137] At the COOH-terminus, several polypeptide loops form a fibrinogen globe which binds Ca^{2+} and influences interactions with other proteins.[135] These domains are highly conserved and are therefore likely to fulfil essential roles.[77] TN-C is expressed transiently and in a highly constrained manner in embryonic tissues, and during wound healing and tissue remodelling in adult tissues.[136,137] Its expression is regulated by a number of factors, including several growth factors and cytokines.[130,137] It is particularly expressed in musculoskeletal regions

University of Cape Town

that are exposed to high tensile and compressive stress during the transmission of mechanical forces and there is evidence that TN-C expression is mechanosensitive.[129,130,136] In particular, Jarvinen et al.[130] showed that TN-C expression was down regulated in the OTJ of the rat quadriceps tendon following three weeks of immobilisation, increased slightly after free remobilisation and increased significantly after intensified remobilisation. This suggests that TN-C expression in the OTJ is regulated by mechanical strain in a dose-dependent manner. Similar results were found in the MTJ of the rat Achilles tendon and gastrocnemius muscle, as well as the Achilles tendon itself, where mechanical loading was again shown to regulate TN-C expression in a dose-dependent manner.[129] TN-C expression has also been reported to increase significantly in tendinopathy.[133,221]

In close proximity to the *TNC* gene, is the *COL27A1* gene (chr9q32) (Figure 3.3) encoding the highly conserved α1 chain of type XXVII collagen. Boot-Handford et al. (2003)[35] hypothesized that mutations in the conserved regions of this gene are likely to be the cause of heritable conditions in man although the authors identified no candidate diseases based on the chromosomal location of this gene. The human *COL27A1* gene is approximately 156kbp in length and consists of 61 exons where exons 1 and 2 code for the signal peptide, exon 3 codes for the remainder of the N-terminal non-collagenous domain, exons 4 through 55 encode the highly conserved collagenous domains while the C-terminal noncollagenous domain is encoded by exons 56-61.[35,216] Type XXVII collagen is expressed in many tissues but particularly in developing cartilage,[35,216] and is up-regulated in tendinopathy (2.5 fold, $P<0.005$).[133]

The aim of this study was, therefore, to test the association of several sequence variants spanning a 780.9kbp region which includes the *COL27A1* and *TNC* genes, and a potential haplotype consisting of two or more of these SNPs, with Achilles tendinopathy in two separate populations.

University of Cape Town

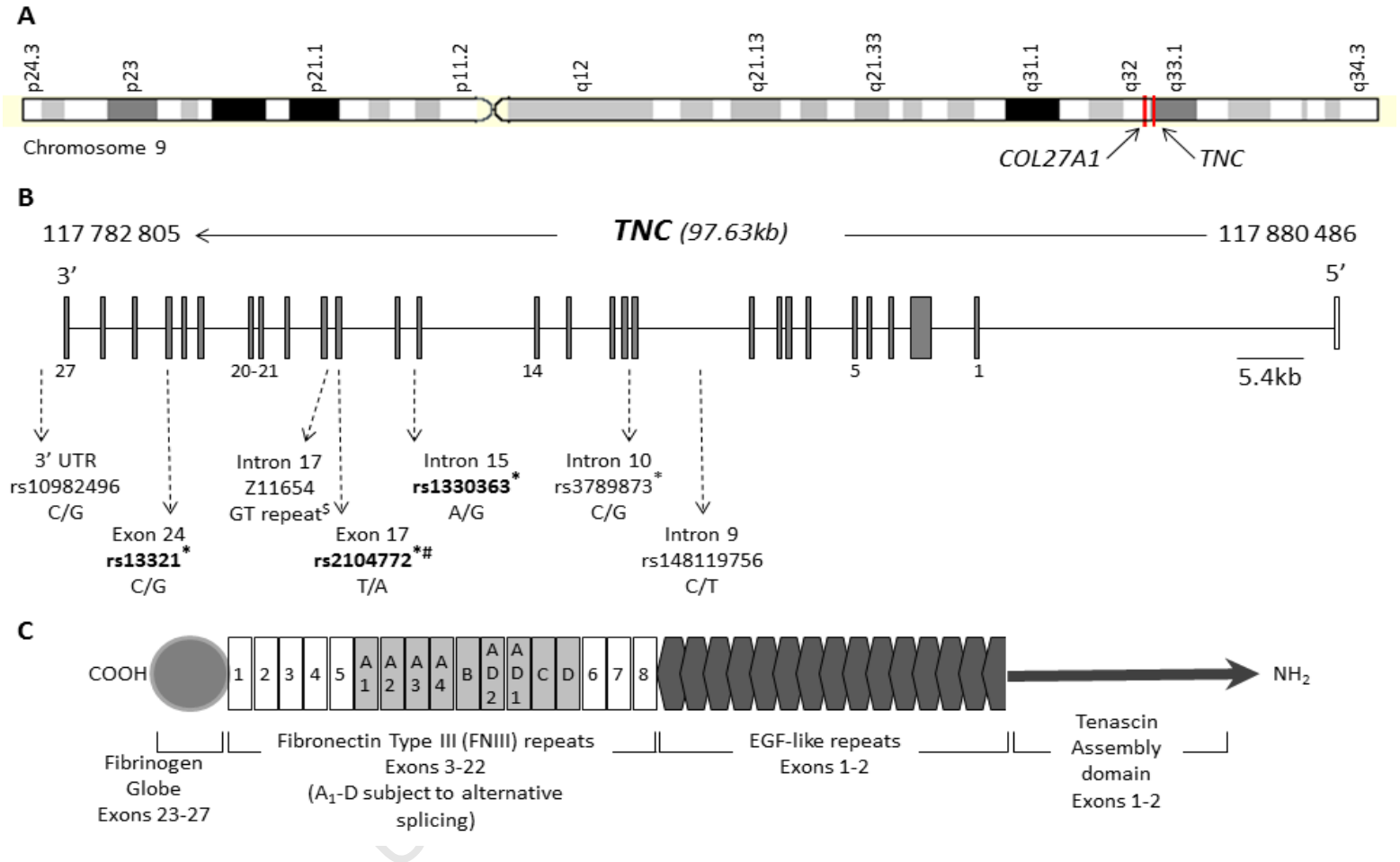

Figure 3.2: Tenascin C

(A) *Shaded bands indicate cytogenetic banding patterns for chromosome 9. Chromosomal banding position of COL27A1 and TNC genes are indicated with arrows* ***(B)*** *Genomic location and intron/exon structure of TNC including previously cited variants.* **Associated with rhinoconjunctivitis [Orsmarck-Pietras et al. (2008)*[211]*]* $^{\#}$*Included in haplotype associated with adult asthma [Matsuda et al. (2005)*[190]*]* $^{\$}$*Associated with Achilles tendon injuries [Mokone et al. (2005)*[197]*] Solid bars represent numbered exons separated by introns (solid lines). Variants in bold are included in the present study* ***(C)*** *TN-C protein domain architecture [Compiled from www.ensembl.org and Jones et al. (2000)*[136]*; Reprinted from Saunders et al. (2012)*[247]*]*

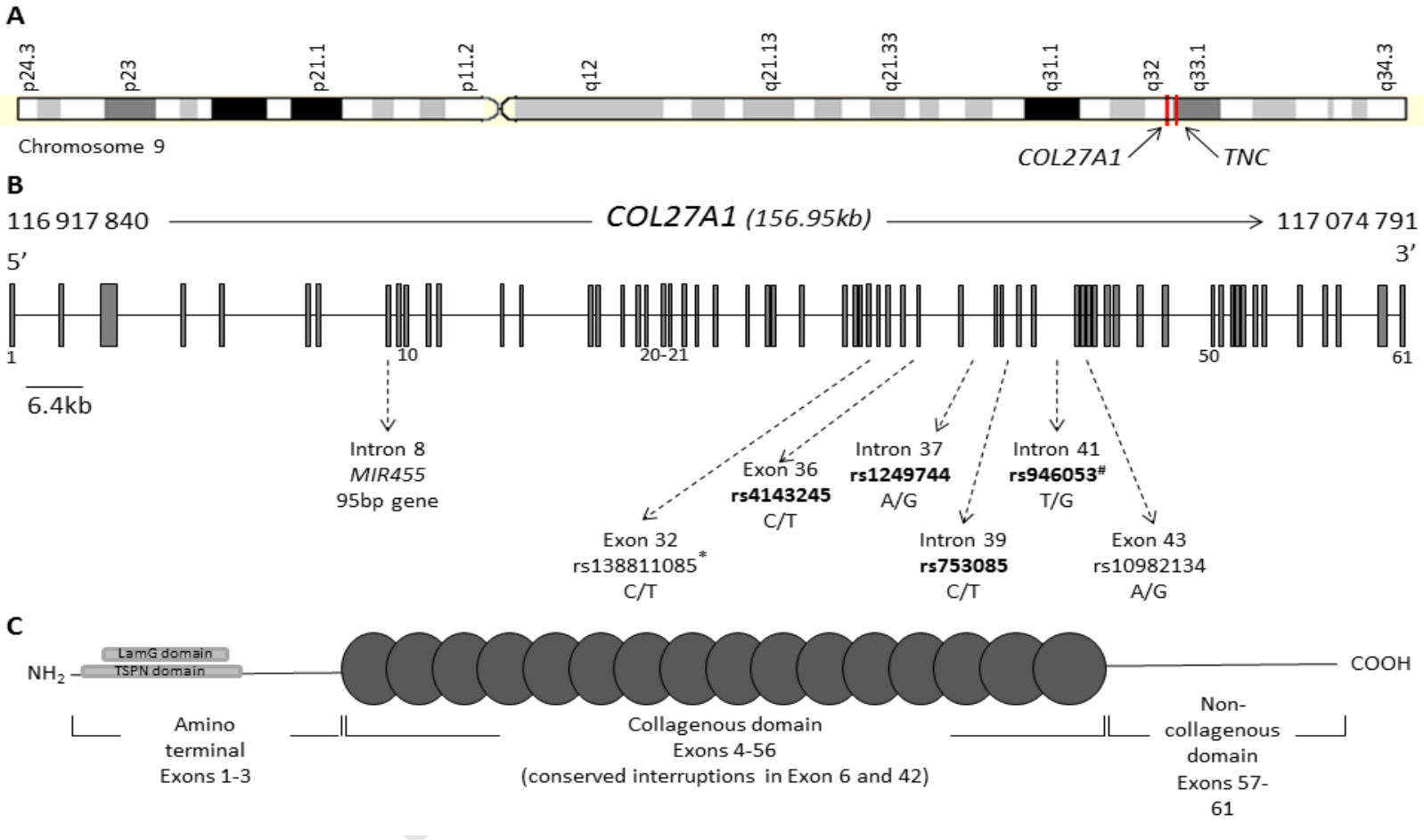

Figure 3.3: The α1 chain of type XXVII collagen

(A) *Shaded bands indicate cytogenetic banding patterns for chromosome 9. Chromosomal banding position of COL27A1 and TNC genes are indicated with arrows* ***(B)*** *Genomic location & intron/exon structure of COL27A1 including previously cited variants.* [*] *Potential clinical source* [#] *Associated with height [Gudbjartsson et al. (2008)[101]] Solid bars represent numbered exons separated by introns (solid lines). Variants in bold are included in the present study* ***(C)*** *Collagen XVII α1 chain protein domain architecture [Compiled from www.ensembl.org, Boot-Handford et al. (2003)[35] and Pace et al. (2003)[216]; Reprinted from Saunders et al. (2012)[247]]*

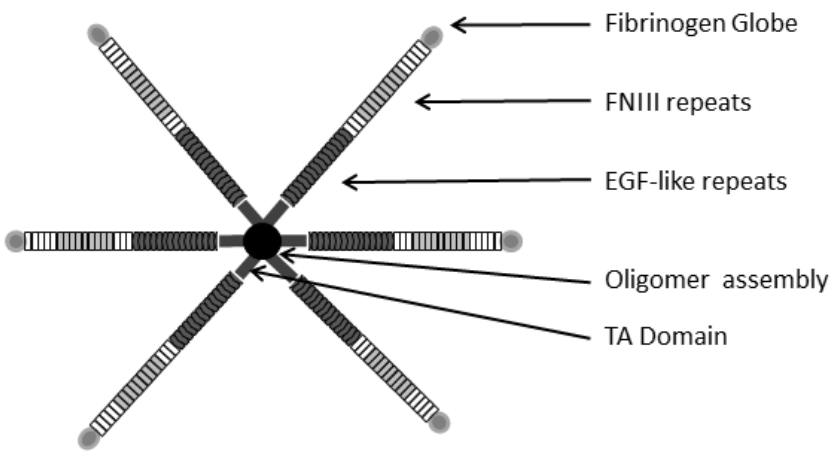

Figure 3.4: Tenascin-C hexabrachion
FNIII: Fibronectin type III; EGF: Epidermal growth factor; TA: Tenascin Assembly
[Adapted from Jones et al. (2000)[136]]

University of Cape Town

3.2. MATERIALS AND METHODS

3.2.1. PARTICIPANTS

The SA-TEN (n=94), SA-CON (n=131), AUS-TEN (n=85) and AUS-CON (n=209) groups described in chapter two were used in this study. Participant characteristics are presented in Table 2.2 and described in 2.3.1.

3.2.2. VARIANT SELECTION

Lists of SNPs within the *COL27A1* and *TNC* genes were generated using the NCBI SNP database available at http://www.ncbi.nlm.nih.gov/snp/ and were interrogated using the Ensembl Genome Browser (http://www.ensembl.org/index.html) and Genome Variation Server (GVS)(http://gvs.gs.washington.edu/GVS137/).[283] SNPs were selected for analysis based on their location, reported minor allele frequency (>0.20), reported heterozygosity (>0.35), biological significance and previously reported associations with other multi-factorial conditions. Four SNPs within the highly conserved collagenous domains of the *COL27A1* gene were therefore selected for investigation. Rs4143245 is a synonymous T>C substitution in exon 36 while rs1249744 (A>G), rs753085 (G>A) and rs946053 (T>G) are all intronic. The latter SNP, rs946053, is in intron 41 and has previously been associated with human height in a genome wide association study.[101]

Three SNPs in close proximity to the GT dinucleotide repeat polymorphism within intron 17 of the *TNC* gene were also selected for investigation. A non-synonymous G>C substitution in exon 24 of the *TNC* gene, rs13321, results in a change from a glutamic acid residue to a glutamine residue at position 2008 in the TN-C protein (Glu2008Gln), and has been associated with rhinoconjunctivitis in children.[211] Another non-synonymous SNP in exon 17, rs2104772 (T>A), results in a change from a leucine to an isoleucine residue at position 1677 (Leu1677Ile) and is associated with both adult asthma and childhood rhinoconjunctivitis, while rs1330363 (C>T) is in intron 15 and has previously been associated with childhood rhinoconjunctivitis[190,211]

3.2.3. ALLELIC DISCRIMINATION

SA-CON, SA-TEN, AUS-CON and AUS-TEN participants were genotyped for all seven polymorphisms. Standard PCR and RFLP techniques were used to amplify and genotype polymorphisms rs4143245, rs946053, rs13321, rs2104772 and rs1330363. Unless otherwise specified, PCR reactions were carried out in a solution containing 200ng DNA template, 1X reaction buffer, 2.0mM Mg^{2+}, 2.5mM each of dATP, dCTP, dTTP and dGTP, 20pmol each of the forward and reverse primers and 0.5 units of either *Taq* DNA polymerase (*New England Biolabs® Inc.,* Ipswich, Massachusetts, USA) or SuperTherm *Taq* DNA polymerase (*Hoffmann-La-Roche*, US) using an XP Thermal Cycler Block (*Bioer Technology Co.*, Middlesex, UK). Reaction buffer was specific to the *Taq* used. Thermo-Pol® reaction buffer (*New England Biolabs® Inc.,* Ipswich, Massachusetts, USA) consisted of 10mM KCl, 10mM $(NH_4)SO_4$, 20mM Tris-HCl, 2mM $MgSO_4$ and 0.1% Triton X-100 at pH8.8. SuperTherm reaction buffer (*Hoffmann-La-Roche*, US) was provided with the SuperTherm *Taq* DNA polymerase and used as per manufacturer's instructions. PCR reactions were initiated with a five minute denaturing cycle at 94˚C, followed by 35 cycles consisting of 30 seconds denaturation at 94˚C, 30 seconds annealing at variant specific temperatures and 40 seconds extension at 72˚C with a final five minute extension step at 72˚C. Variant specific oligonucleotide primer sequences, reaction conditions and amplicon sizes are listed in Appendix B. PCR amplicons of polymorphisms rs4143245, rs946053, rs13321, rs2104772 and rs1330363 were digested overnight for 15 hours with *XhoI, BglI, BbvCI, MseI* and *HpyCH4III* restriction endonucleases respectively as per manufacturer's guidelines.

University of Cape Town

Tetra-primer amplification refractory mutation system (T-ARMS) PCR as described by Ye et al[308] was used to genotype polymorphisms rs1249744 and rs753085. Unless otherwise specified, PCR reactions were carried out in a solution containing 200ng DNA template, 1X reaction buffer, 2.0mM Mg^{2+}, 2.5mM each of dATP, dCTP, dTTP and dGTP, 20pmol each of the forward and reverse primers and 0.5 units of either *Taq* DNA polymerase (*New England Biolabs® Inc.,* Ipswich, Massachusetts, USA) or SuperTherm *Taq* DNA polymerase (*Hoffmann-La-Roche*, US) using an XP Thermal Cycler Block (*Bioer Technology Co.*, Middlesex, UK). Reaction buffer was specific to the *Taq* used. Thermo-Pol® reaction buffer (*New England Biolabs® Inc.,* Ipswich, Massachusetts, USA) consisted of 10mM KCl, 10mM $(NH_4)SO_4$, 20mM Tris-HCl, 2mM $MgSO_4$ and 0.1% Triton X-100 at pH8.8. SuperTherm reaction buffer (*Hoffmann-La-Roche*, US) was provided with the SuperTherm *Taq* DNA polymerase and used as per manufacturer's instructions. PCR reactions were initiated with a five minute denaturing cycle at 95˚C, followed by an initial five cycles consisting of 25 seconds denaturation at 95˚C, 45 seconds annealing at 70˚C and 30 seconds extension at 72˚C. This was followed by 27 cycles consisting of 25 seconds denaturation at 95˚C, 45 seconds annealing at variant specific temperatures and 30 seconds extension at 72˚C with a final 10 minute extension step at 72˚C. Variant specific oligonucleotide primer sequences, reaction conditions and amplicon sizes are listed in Appendix B. Genotyping accuracy was confirmed by allele specific PCR in a subset of samples, and a small subset of samples were genotyped twice. A number of known genotype positive controls and DNA-free controls were included on each 96-well PCR plate. The number and proportion of successfully genotyped samples is presented in Table 3.1.

All PCR amplicons and restriction fragments were mixed with *SYBER® Gold* nucleic acid gel stain (*Invitrogen Molecular Probes*™, Oregon, USA) and, together with a 100bp DNA size standard (*Promega Corporation*, Madison, Wisconsin, USA), separated by electrophoresis on 2% agarose (100V for 35 minutes) and 6% PAGE (120V for 2 hours) gels respectively. All gels were photographed under UV light using a photodocumentation system (*Uvitec Limited*, Cambridge, UK) [Appendix B].

University of Cape Town

3.2.4. STATISTICAL ANALYSIS

The programming environment R[230] and R packages were used for all analyses. The R package, Genetics,[296] was used to estimate genotype and allele frequencies and HWE probabilities. Logistic regression was used to compare genotype, allele and allele-combination frequencies between the TEN and CON groups, as well as SA and AUS groups. Results with a P-value of less than 0.05 were accepted as significant. As the genotype frequencies did not differ significantly between either AUS-TEN and SA-TEN or AUS-CON and SA-CON, the two populations were combined for further analyses. All analyses were corrected for potential confounding by including age, sex, and group (SA or AUS) in the models as fixed effects. Haplotype frequencies were inferred for both the SA and AUS groups using the R package, haplo.stats.,[249,266] and were based on the genotype data of all seven polymorphisms. No adjustments were made for multiple testing because it has been suggested that these corrections, such as Bonferroni, markedly overcorrect for an inflated false-positive rate and unnecessarily reduce power in genetic association studies where background LD exists between SNPs.[207,220]

Genotype effects on age, height, weight, BMI and sex were determined by a one-way analysis of variance (ANOVA) using STATISTICA version 10 (*StatSoft Inc.*, Tulsa, OK, USA). Statistical significance was accepted when $P<0.05$.

3.2.5. PRE-MRNA STRUCTURE MODELLING

Sfold software[72,73] (http://sfold.wadsworth.org accessed May 2012) was used to predict differences in pre-mRNA secondary structure in ancestral versus minor alleles for rs946053, rs13321 and rs2104772. Briefly, the Sfold algorithm generates statistically representative samples from the ensemble of secondary structures in proportion to their Boltzmann weights.[72,73] The centroid structure is that which has the minimum base-pair distance to all the structures in a given set and most accurately represents the tendency of a set of structures, either in a free energy cluster (cluster centroid) or the entire Boltzmann weighted ensemble of structures (ensemble centroid). In this study, the ten cluster centroids and the ensemble centroid were examined for differences between the ancestral and minor allelic forms of these three polymorphisms. In particular, structures were examined for changes in accessibility to splicing regulators and transcription factors by

University of Cape Town

Table 3.1: Number and proportion of recruited participants successfully genotyped for the seven polymorphisms investigated in this study

Total number of participants		Number and proportion of participants genotyped													
		rs4143245		rs1249744		rs753085		rs946053		rs13321		rs2104772		rs1330363	
AUS-CON	**209**	194	93%	205	98%	206	99%	198	95%	200	96%	195	93%	158	76%
AUS-TEN	**85**	77	91%	84	99%	83	98%	80	94%	79	93%	84	99%	64	75%
TOTAL	**294**	**271**	**92%**	**289**	**98%**	**289**	**98%**	**278**	**95%**	**279**	**95%**	**279**	**95%**	**222**	**76%**
SA-CON	**131**	130	99%	130	99%	128	98%	131	100%	129	98%	127	97%	129	98%
SA-TEN	**94**	93	99%	94	100%	93	99%	91	97%	94	100%	90	96%	91	97%
TOTAL	**225**	**223**	**99%**	**224**	**99%**	**221**	**98%**	**222**	**99%**	**223**	**99%**	**217**	**96%**	**220**	**98%**

University of Cape Town

comparing folding and structural elements (e.g. hairpin loops), and single-strandedness. In general, proteins regulating splice site selection preferentially recognise single-stranded rather than double-stranded pre-mRNA,[111] and folding of pre-mRNA, based on base-pairing and hydrophobic interactions, results in occlusion or exposure of cis-acting regulatory elements which effect accessibility of regulatory proteins.[43] Secondary pre-mRNA structures are, therefore, important in determining splicing outcome.[43,111] Increased accessibility is assumed where pre-mRNA is single-stranded and/or presented in a protruding hairpin loop. Decreased accessibility is assumed where pre-mRNA is double stranded and/or folded in such a way as to be hidden by other secondary structures.

3.2.6. FUNCTIONAL ANALYSIS OF VARIANTS

The online software programmes SIFT,[160] PolyPhen-2[8] and FastSNP[310] were used to predict functional effects of the three SNPs implicated by the haplotype analysis (rs946053, rs13321 and rs2104772). The FASTSNP[310] analysis tool was used to identify any putative transcription factor binding sites and regulatory elements in close proximity to the SNPs and predict a risk for deleterious effects of that SNP.

3.3. RESULTS

3.3.1. GENOTYPE AND ALLELE FREQUENCIES

Genotype and allele frequency distributions for each of the polymorphisms, together with HWE probabilities, are shown in Tables 3.2 and 3.3. The relationship between the genotypes and AT was tested and found not to differ significantly between the SA and AUS groups. Data from the population groups was, therefore, combined for all further analyses. Age, sex and country group were considered confounders (Table 2.2) and were adjusted for in all analyses by including them in the models.

After adjusting for the confounders age, sex and country group there were no significant differences in either genotype or allele frequencies between the CON and TEN groups at the rs4143245, rs1249744, rs753085, rs946053 and rs13321 loci. However, a significant difference in the allele distribution was observed for rs2104772 with the T-allele occurring

University of Cape Town

Table 3.2: Genotype and minor allele frequency distributions of the sequence variants investigated within the *COL27A1* gene in control (CON) and Achilles tendinopathy (TEN) groups of South African (SA) and Australian (AUS) populations

	CON		TEN		P-Values	
	AUS	**SA**	**AUS**	**SA**	**Country**	**Group**
COL27A1:						
rs4143245						
N	194	130	77	93		
T/T	0.39	0.44	0.44	0.42	0.601	0.610
T/C	0.49	0.45	0.38	0.47		
C/C	0.12	0.12	0.18	0.11		
C	0.37	0.34	0.37	0.34	0.314	0.933
HWE	0.538	1.000	0.092	0.818		
rs1249744						
N	205	130	84	94		
A/A	0.55	0.54	0.46	0.40	0.340	0.106
A/G	0.37	0.36	0.45	0.47		
G/G	0.08	0.10	0.08	0.13		
G	0.27	0.28	0.31	0.36	0.159	0.084
HWE	0.476	0.276	0.798	1.000		
rs753085						
N	206	128	83	91		
C/C	0.56	0.54	0.54	0.43	0.186	0.199
C/T	0.33	0.34	0.39	0.47		
T/T	0.11	0.13	0.07	0.10		
T	0.27	0.29	0.27	0.34	0.083	0.440
HWE	**0.013**	**0.034**	1.000	0.644		
rs946053						
N	198	131	80	93		
G/G	0.23	0.24	0.23	0.38	0.149	0.248
G/T	0.56	0.56	0.58	0.48		
T/T	0.22	0.21	0.20	0.14		
T	0.49	0.48	0.49	0.38	0.074	0.096
HWE	0.155	0.223	0.262	1.000		

P-values are for the difference between countries and between diagnostic groups respectively, adjusted for each other, age, sex and whether or not a person was investigated in his/her country of birth. The genotype P-value is calculated using a 2-degrees of freedom test, with genotypes as categories and the allelic P-value is calculated using an additive allelic model. HWE gives exact P-values from tests of Hardy–Weinberg equilibrium. N is number of samples genotyped. Bold P-values are <0.05

University of Cape Town

significantly less frequently in TEN than CON (OR 0.70, 95% CI: 0.52-0.94; P=0.017). For rs1330363, the G-allele was significantly more prevalent in TEN than in CON after adjusting for the three known confounders (OR 1.46, 95% CI: 1.06-2.03; P=0.020), although any association found with rs1330363 should be interpreted with caution due to the smaller number of Australian participants genotyped for this SNP (76%)(Table 3.1). At the rs753085 locus, the genotype frequency for both the AUS CON (P=0.013) and SA CON (P=0.034) deviated from HWE with too few heterozygotes observed (Table 3.2). There were no genotype effects on age, weight, height, BMI or sex (Table 3.4).

Table 3.3: Genotype and minor allele frequency distributions of the sequence variants investigated within the *TNC* gene in control (CON) and Achilles tendinopathy (TEN) groups of South African (SA) and Australian (AUS) populations

	CON		TEN		P-Values		OR (95% CI)
	AUS	SA	AUS	SA	Country	Group	
***TNC*:**							
rs13321							
N	200	129	79	94			
C/C	0.48	0.52	0.52	0.39	0.430	0.291	
C/G	0.46	0.40	0.41	0.48			
G/G	0.06	0.08	0.08	0.13			
G	0.29	0.28	0.28	0.37	0.418	0.166	
HWE	0.123	1.000	1.000	0.828			
rs2104772							
N	195	127	84	90			
A/A	0.33	0.31	0.36	0.43	0.141	0.056	
A/T	0.49	0.45	0.54	0.40			
T/T	0.18	0.24	0.11	0.17			
T	0.43	0.46	0.38	0.37	0.398	**0.017**	0.70 (0.52-0.94)
HWE	1.000	0.286	0.248	0.181			
rs1330363							
N	158	129	64	91			
A/A	0.34	0.39	0.22	0.27	0.220	0.065	
A/G	0.46	0.46	0.52	0.53			
G/G	0.20	0.16	0.27	0.20			
G	0.43	0.38	0.52	0.46	0.082	**0.020**	1.46 (1.06-2.03)
HWE	0.516	0.712	1.000	0.674			

P-values are for the difference between countries and between diagnostic groups respectively, adjusted for each other, age, sex and whether or not a person was investigated in his/her country of birth. The genotype P-value is calculated using a 2-degrees of freedom test, with genotypes as categories and the allelic P-value is calculated using an additive allelic model. HWE gives exact P-values from tests of Hardy–Weinberg equilibrium. N is number of samples genotyped. Bold P-values are <0.05

University of Cape Town

Table 3.4: Genotype effects of *COL27A1* and *TNC* SNPs on physiological parameters of participants

POLYMORPHISM	P-VALUES					
	AGE	AGE INJURED	HEIGHT	WEIGHT	BMI	SEX
rs4143245	0.546	0.879	0.914	0.435	0.317	0.554
rs1249744	0.235	0.262	0.399	0.934	0.441	0.578
rs753085	0.062	0.265	0.903	0.192	0.070	0.500
rs946053	0.646	0.988	0.599	0.932	0.905	0.975
rs13321	0.457	0.724	0.755	0.372	0.525	0.605
rs2104772	0.558	0.125	0.986	0.616	0.246	0.569
rs1330363	0.154	0.502	0.668	0.592	0.712	0.709

Age: Age at recruitment; Age Injured: Age at recruitment (CON) & age at diagnosis (TEN); P-values for age, age injured, height, weight and BMI are determined by one-way ANOVA; P-value for sex is determined by Pearsons χ^2 analysis

University of Cape Town

3.3.2. HAPLOTYPE ANALYSIS

Haplotypes consisting of various combinations of the seven SNPs investigated were inferred and are presented in Table 3.5. There was no significant difference between the combined CON and combined TEN groups in the frequency of inferred haplotypes constructed from all seven polymorphisms (P=0.511). However, significant differences in haplotype frequencies were noted between the CON and TEN groups, after adjusting for the confounders, when the inferred haplotypes were constructed from the following six polymorphisms: rs4143245; rs1249744; rs753085; rs946953; rs13321 and rs2104772 (P=0.029). Similarly, significant differences in the haplotype distributions were noted when haplotypes were constructed using the most distal *COL27A1* polymorphism, rs946053, and the two proximal *TNC* polymorphisms, rs13321 and rs2104772 (P=0.026).

Table 3.6, Table 3.7 and Figure 3.5. show the inferred frequencies of each combination of alleles in the implicated six and three SNP haplotypes. In the six SNP haplotype, the TACTCT haplotype occurs at the highest combined frequency and is therefore used as the reference haplotype. In the three SNP haplotype the GCA haplotype occurs at the highest combined frequency and is therefore used as the reference haplotype. This GCA haplotype was more

frequently observed in the TEN group (27%) compared to the CON group (18%) and was therefore significantly associated with an increased risk of AT (P=0.019). Three haplotypes showed significantly lower odds of TEN versus the GCA haplotype [(GCT: OR 0.35, 95% CI: 0.17-0.70)(TCA: OR 0.38, 95% CI: 0.18-0.81)(TGT: OR 0.09, 95% CI: 0.01-0.85)]. This association is reflected in the association of a six SNP haplotype starting with rs4143245 with AT, in which the TACTCT haplotype allele occurs at a significantly higher frequency amongst controls (CON: 0.11 vs TEN: 0.08, P=0.024)(Table 3.6).

Table 3.5: The probability of differences in the frequency of inferred haplotypes between the CON (AUS & SA) and TEN (AUS & SA) participants

GENE	POLYMORPHISM	CHROMOSOMAL POSITION	NUMBER OF POLYMORPHISMS (P)					
			2	3	4	5	6	7
COL27A1	**rs4143245**	9:117033022T>C	0.276	0.293	0.387	0.245	**0.029**	0.511
	rs1249744	9:117043352A>G	0.208	0.434	0.797	0.175	0.513	
	rs753085	9:117045447G>A	0.398	0.633	0.122	0.448		
	rs946053	9:117049891T>G	0.185	**0.026**	0.134			
	rs13321	9:117792583C>G	0.091	0.323				
TNC	**rs2104772**	9:117808785T>A	0.114					
	rs1330363	9:117813990C>T						

Polymorphisms listed indicate the first polymorphism of the haplotype and 'number of polymorphisms' indicates the total number of adjacent polymorphisms included in that haplotype analysis. Pairs of haplotypes were inferred and tested for association with group (CON vs. TEN) adjusting for age, sex and group (P)

University of Cape Town

Table 3.6: Inferred allele frequencies for a six SNP haplotype including rs4143245-rs1249744-rs753085-rs946053-rs13321-rs2104772

ALLELE	CON	TEN	P	ODDS RATIO (95% CI)	
TACTCT	0.11	0.08	**0.024**	1	P
TGTGCT	0.08	0.06	0.781	1.55 (0.49-4.89)	0.459
TACGCA	0.07	0.11	0.733	2.22 (0.69-7.16)	0.183
TGTGCA	0.07	0.09	0.242	2.13 (0.78-5.83)	0.141
CACTCT	0.06	0.09	0.462	2.04 (0.65-6.40)	0.222
CACTCA	0.07	0.05	0.151	0.79 (0.22-2.76)	0.707
TACTCA	0.06	0.04	0.815	1.88 (0.47-7.49)	0.373
TACGGA	0.07	0.03	0.959	1.39 (0.44-4.41)	0.579
CACTGA	0.06	0.03	0.770	1.32 (0.34-5.06)	0.686
TGTGGA	0.05	0.05	0.337	2.23 (0.66-7.51)	0.196
TACTGA	0.02	0.07	0.605	2.72 (0.75-9.94)	0.131
CACGCA	0.04	0.03	0.147	2.87 (0.86-9.57)	0.087
CACGCT	0.04	0.03	0.314	0.69 (0.07-6.80)	0.751
TACGCT	0.04	0.02	0.106	0.83 (0.15-4.60)	0.827
CACTGT	0.02	0.02	0.977	1.09 (0.01-103.79)	0.970
TACTGT	0.02	0.00	0.227	0.37 (0.01-16.24)	0.607
CACGGA	0.00	0.05	**0.005**	67.01 (0.15-30687.31)	0.179

Values are inferred haplotype frequencies and P-value for a test of equality of those frequencies. Odds ratios of AT are calculated versus the TACTCT reference allele. Bold P-values are <0.05

University of Cape Town

Table 3.7: Inferred allele frequencies for a three SNP haplotype including rs946053-rs13321-rs2104772

ALLELE	CON	TEN	P	ODDS RATIO (95% CI)	
GCA	0.18	0.27	**0.019**	1	P
TCT	0.19	0.20	0.083	0.64 (0.37-1.10)	0.106
GCT	0.18	0.10	0.061	0.35 (0.17-0.70)	**0.003**
TCA	0.15	0.10	0.324	0.38 (0.18-0.81)	**0.013**
GGA	0.12	0.14	0.107	0.73 (0.36-1.50)	0.390
TGA	0.10	0.12	0.394	0.92 (0.49-1.73)	0.792
TGT	0.04	0.01	0.246	0.09 (0.01-0.85)	**0.036**
GGT	0.02	0.06	0.403	1.65 (0.52-5.22)	0.394

Values are inferred haplotype frequencies and P-value for a test of equality of those frequencies. Odds ratios of AT are calculated versus the GCA reference allele. Bold P-values are <0.05

University of Cape Town

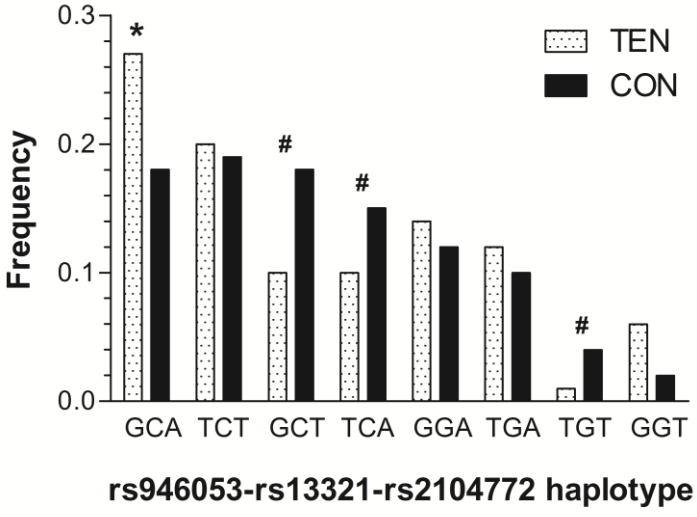

Figure 3.5: Inferred allele frequencies for a three SNP haplotype containing rs946053-rs13321-rs2104772

[*]CON vs TEN (P=0.019); [#]OR vs GCA reference allele (P<0.05)

3.3.3. PREDICTED FUNCTIONAL EFFECTS OF SELECTED SNPS

Bioinformatics tools were used to investigate the functional effects of the SNPs implicated by the three SNP haplotype (rs946053; rs13321; rs2104772) and revealed several potentially functional elements that overlap this region. These are summarised in Table 3.8 and described in detail below.

Rs946053 (G>T) is intronic and therefore not listed on either the SIFT[160] or PolyPhen-2[8] database. FASTSNP[310] reports a very low to low predicted risk associated with this SNP, despite the elimination of a c-Myc transcription factor binding site in the presence of the minor T-allele. In addition, for the majority of Sfold[72,73] predicted pre-mRNA secondary structures generated from the 457bp DNA sequence surrounding this SNP, including the ensemble centroid, the G-allele is predicted to result in a more single stranded pre-mRNA conformational structure immediately surrounding the SNP compared to the ancestral T-allele (Figure 3.6). This suggests that the c-Myc binding site, which is abolished in the presence of the minor T-allele, is normally highly exposed.

Rs13321 (G>C) involves a non-synonymous amino acid substitution (Glu2008Gln) which is predicted to be tolerated (SIFT[160]) and benign (PolyPhen-2[8]). Furthermore, FASTSNP[310] predicts a low to medium risk associated with this sequence variant which falls within, but does not abolish, a GATA transcription factor binding site. The Sfold[72,73] predicted pre-mRNA secondary structure for this polymorphism suggests that this transcription factor binding site is exposed by unfolding when the G-allele is present. FASTSNP[310] also predicts a splicing regulation effect with the different allelic forms of this SNP. An exonic splicing enhancer (ESE) and an exonic splicing silencer (ESS) occur 7bp and 15bp downstream of the SNP respectively. Analysis of the Sfold predicted centroid structures suggests that the ESE is more accessible, and the ESS is less accessible, in the secondary pre-mRNA structure of the C-allele compared to the G-allele. This is, however, not evident in the representative ensemble centroid (Figure 3.7). Although this configuration is not mirrored in the ensemble centroid structure, the majority of cluster centroid structures generated using the G-allele sequence predict the ESS to be more exposed in a single-stranded hairpin loop, while the ESE is occluded by folding. This SNP lays 39bp upstream of the exon/intron boundary for exon 24.

University of Cape Town

Rs2104772 (T>A) also involves a non-synonymous amino acid substitution (Leu1677Ile) which is predicted to be tolerated (SIFT[160]) and benign (PolyPhen-2[8]). FASTSNP reports a low to medium risk associated with this missense SNP, but does not predict any change in transcription factor binding sites or ESE and ESS elements. In addition, analysis of the Sfold predicted pre-mRNA secondary structures reveals very similar centroid structures when comparing the A- and T-alleles (Figure 3.8).

In summary, (i) rs946053 (G>T) lays within a putative c-Myc transcription factor binding site which is eliminated in the presence of the T-allele; (ii) rs13321 (G>C) lays within a putative GATA transcription factor binding site, and putative splicing regulatory elements are found 7bp and 15bp downstream of this SNP; but (iii) no predicted functional effects were identified for the sequence region overlapping rs2104772 (T>A).

University of Cape Town

Table 3.8: Summary of the predicted functional effects of polymorphisms implicated in the haplotype

dbSNP rsID	Genomic location	Amino acid change	Functional effect predicted by: SIFT	PolyPhen-2	FASTSNP
rs946053 (T>G)	*COL27A1*: intron 41	n/a	n/a	n/a	Very Low to Low risk - *Abolishes c-Myc binding site*
rs13321 (G>C)	*TNC*: exon 24	Glu2008Gln	Tolerated	Benign	Low to Medium risk - *within GATA binding site* - *7bp upstream of ESE* - *15bp upstream of ESS*
rs2104772 (T>A)	*TNC*: exon 17	Leu1677Ile	Tolerated	Benign	Low to Medium risk

ESE = Exonic splicing enhancer; ESS = Exonic splicing silencer

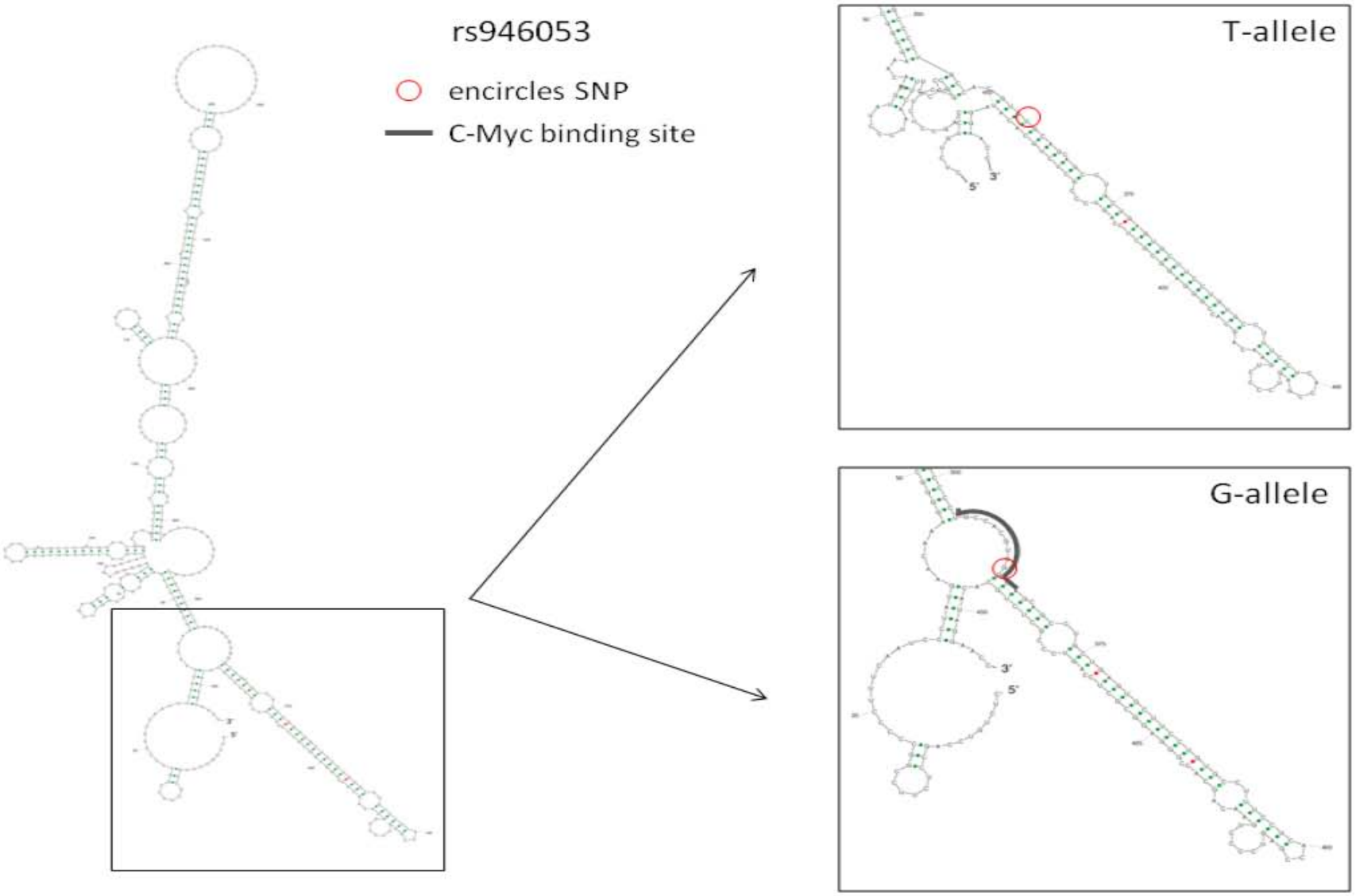

Figure 3.6: The ensemble centroid of Sfold predicted pre-mRNA secondary structures for the sequence surrounding rs946053 in both the T- and G-alleles

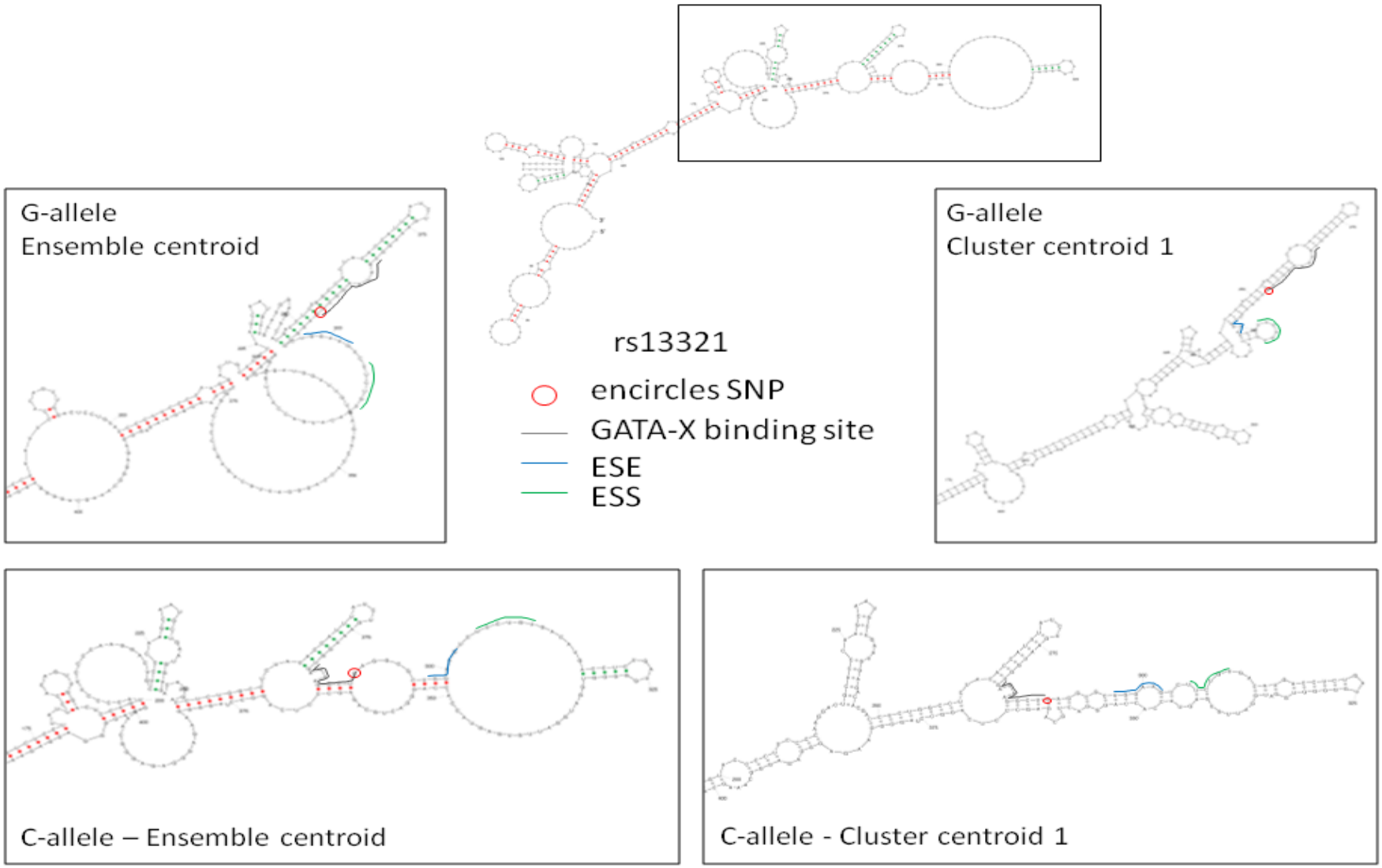

Figure 3.7: The ensemble (left) and cluster 1 (right) centroids of Sfold predicted pre-mRNA secondary structures for the sequence surrounding rs13321 is both the G- and C-alleles

ESE: Exonic splicing enhancer; ESS: Exonic splicing silencer

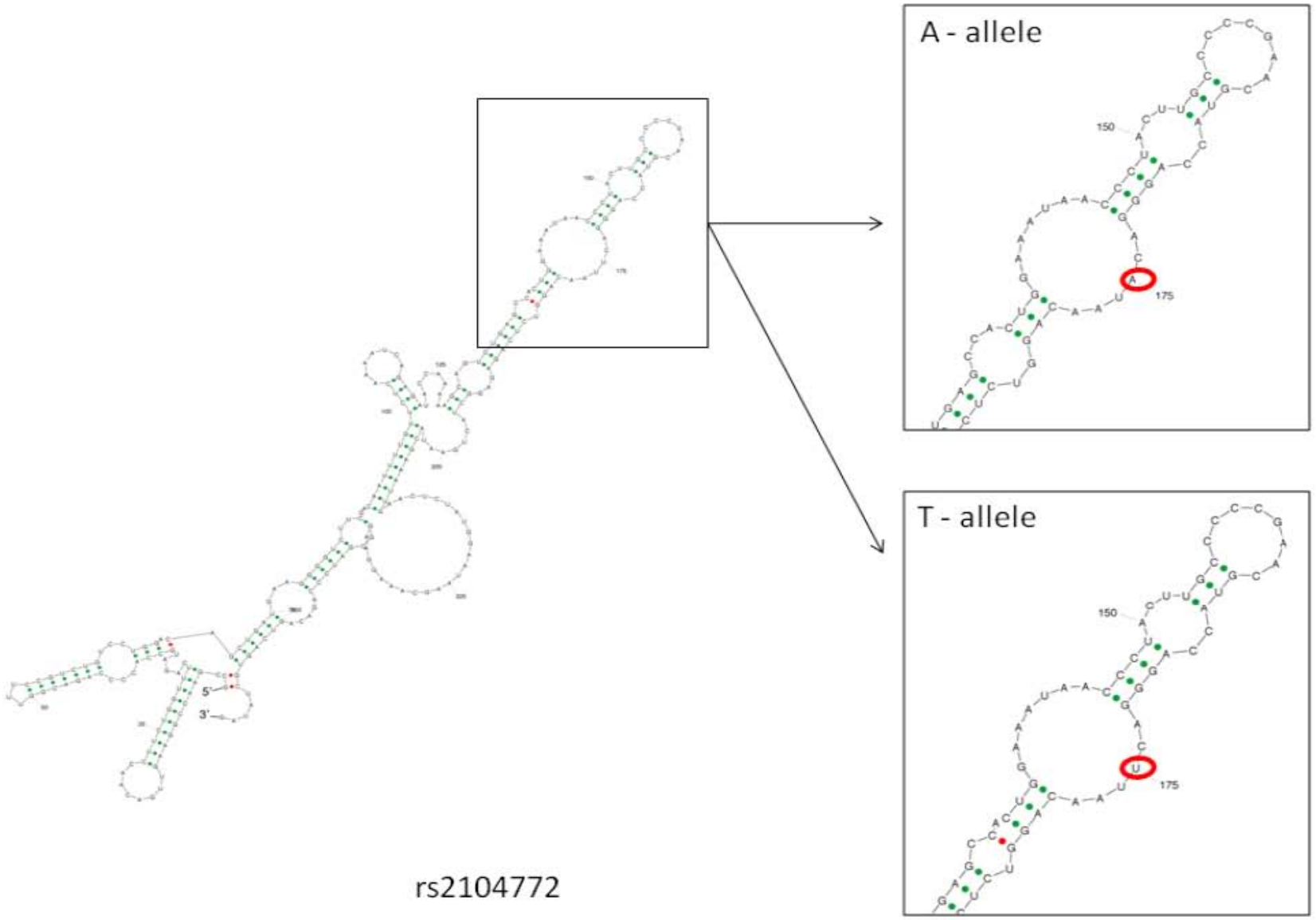

Figure 3.8: The ensemble centroid Sfold predicted pre-mRNA secondary structures for the sequence surrounding rs2104772 in both the A- and T-alleles

3.3.4. LINKAGE DISEQUILIBRIUM

Figure 3.9 shows the D' and r estimates for LD between all seven SNPs included in this study, based on the genotype data. Variant pairs that are in tight LD are indicated in light yellow while pairs that are not in LD are indicated with dark red. The *TNC* SNPs rs13321, rs2104772 and rs1330363 are in tight LD in the TEN groups and, to a lesser degree, in the CON groups. Similarly, the *COL27A1* SNPs rs4143245, rs1249744, rs753085 and rs946053 are in tight LD in both TEN groups and AUS-CON, and moderate LD in the SA-CON group.

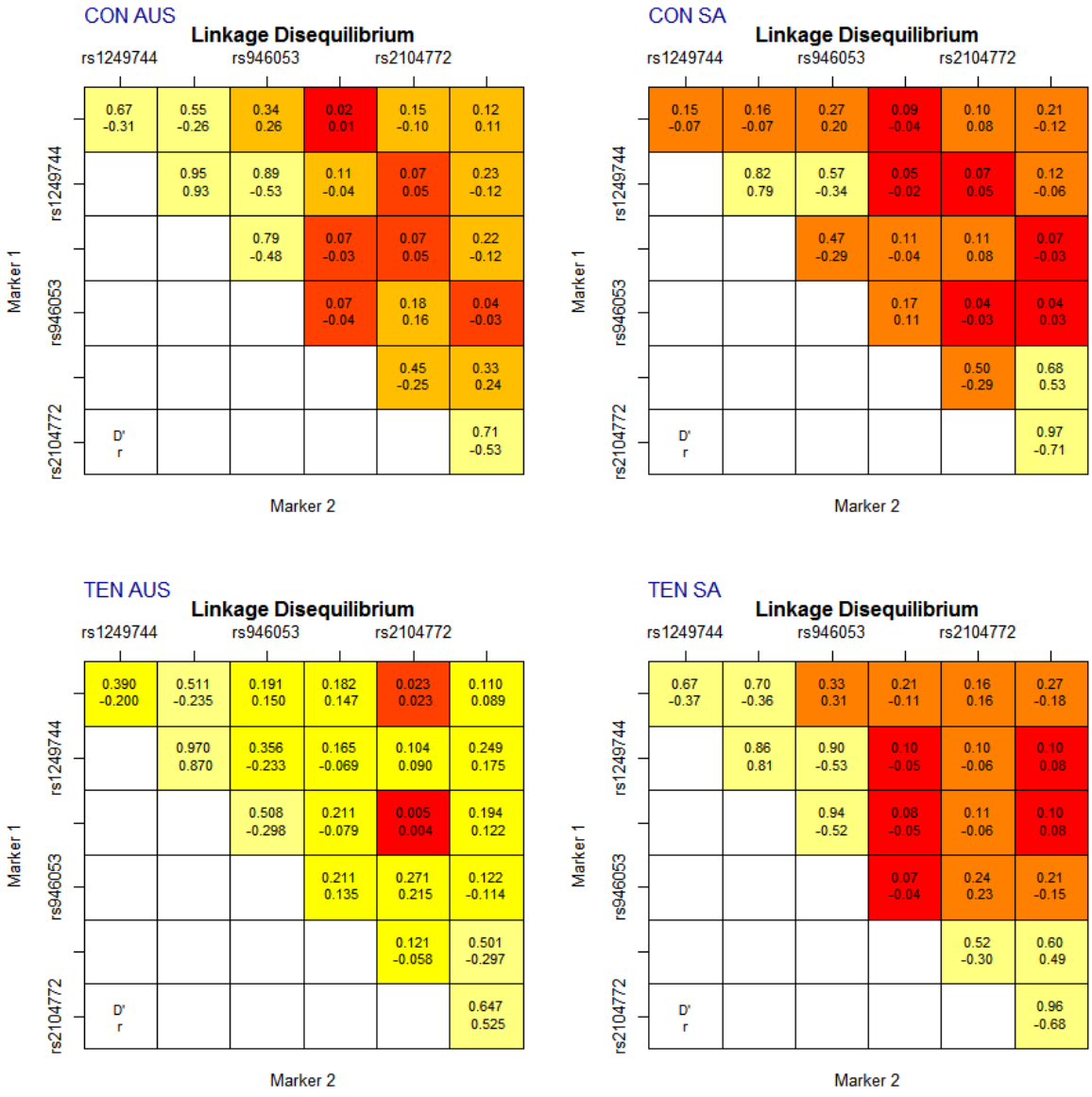

Figure 3.9: Heatmaps showing linkage disequilibrium (LD) estimates between seven polymorphisms within the *COL27A1* and *TNC* genes in each of the groups investigated
LD based on D' and r estimates. Light yellow indicates SNP pairs in tight LD, red indicates SNP pairs that are not in LD

3.4. DISCUSSION

The primary finding of this study is that the GCA haplotype, consisting of the *COL27A1* rs946053 (T>G), *TNC* rs13321 (G>C) and *TNC* rs2104772 (T>A) SNPs, is significantly associated with AT in a South African and Australian population. Mokone et al. (2005)[197] previously described the association between a GT dinucleotide microsatellite marker within the *TNC* gene and ATI. As microsatellite markers have a long range effect, the possibility that this polymorphism was in LD with another neighbouring risk polymorphism could not be excluded. The aims of this case-control genetic association study were therefore, (i) to test this association in samples representative of two populations (SA and AUS) and, (ii) to narrow the interval containing the potential risk allele by conducting genetic analysis of the region containing both candidate genes, *COL27A1* and *TNC,* using SNPs instead of the long range microsatellite marker used in the original study.

The results of this genetic analysis of the combined SA and AUS groups further implicates the *TNC* gene region, as well as the *COL27A1* gene region, in increased risk of developing AT. In particular, the T-allele of rs2104772 was associated with a 0.7 fold reduced risk of developing AT (95% CI:0.52-0.94) and the G-allele of rs1330363 was associated with a 1.46 fold reduced risk of developing AT (95% CI:1.06-2.03). The identification of a risk haplotype (GCA) implicates the genetic interval spanning SNPs rs946053, rs13321 and rs2104772 with increased risk of AT. It is interesting to note that the GT dinucleotide microsatellite (GenBank accession number: Z11654), previously associated with AT in the initial South African ATI group,[197] localises to intron 17 while rs13321 and rs2104772 localise to exons 24 and 17 respectively. Although rs1330363 was independently associated with AT, it was not implicated in the haplotype analysis. This is most likely due to the lower number of AUS participants successfully genotyped for this SNP, which is a limitation to the study. The reason for the lower genotype call rate (76%) for this SNP is the same as that described for rs28494505 in chapter two.

The observed allele and haplotype associations with AT suggest that either (i) the SNPs implicated are in LD with a true risk polymorphism within, close to or between the *TNC* or *COL27A1* genes, or (ii) one or a combination of these SNPs exerts a biological effect and is a potential risk allele for AT. All three SNPs implicated in the haplotype have previously been

University of Cape Town

shown to be associated with other multi-factorial phenotypes. Although there were no genotype effects of rs946053, nor any of the other SNPs investigated in this study, on height, the *COL27A1* rs946053 SNP has previously been associated with human height.[101] The *TNC* SNPs rs13321 and rs2104772 were associated with rhinoconjunctivitis in children from the PARSIFAL study,[211] and rs2104772 is associated with adult asthma in a Japanese population.[190] Considered in conjunction with the bioinformatics analysis conducted in this study, this suggests that the implicated haplotype harbours potentially functional regulatory elements.

The T-allele of rs946053 is predicted to abolish a c-Myc binding site. This oncogenic protein is a global regulator of transcription with an extensive target gene network.[65] In particular, it represses transcription of many collagen genes as well as genes involved in cell adhesion and cell growth arrest, possibly by interfering with activators of gene expression.[65] In the ancestral G-allele, this binding site is situated in a stretch of pre-mRNA that is predicted to be single stranded and, therefore, easily accessible for binding. Despite the location of this binding site in the distal region of the *COL27A1* gene, its abolishment when the T-allele is present may result in less transcriptional repression and inappropriate expression of the α1 chain of type XXVII collagen

Of particular interest is the functional modelling of the sequence surrounding rs13321. This polymorphism occurs in the region encoding the fibrinogen globe domain of the TN-C protein. This highly conserved globe is formed by several polypeptide loops and binds Ca^{2+} thereby influencing TN-Cs interaction with other proteins.[77,136] The G>C substitution at rs13321 results in an amino acid change from a Glutamic acid (Glu) residue in the G-allele to its uncharged amide derivative Glutamine (Gln) at position 2008. This may affect the folding of the protein in this domain and, subsequently, the interaction of TN-C with its associated proteins.[211] In addition, rs13321 falls within, but does not abolish, a GATA transcription factor binding site. The GATA family of transcription factors bind to the (G/A)GATA(A/T) sequence motif and interact with other transcriptional regulators to influence haematopoiesis, cell differentiation and cell proliferation.[99] In particular, GATA-6 has been shown to repress *TNC* expression through response elements in the promoter,[99] and this additional binding site may play a role in the transcriptional regulation of *TNC*. The Sfold

University of Cape Town

predicted pre-mRNA secondary structure for the sequence surrounding this binding site and rs13321 suggests that the GATA transcription factor binding site is more accessible in the C-allele. In addition, this polymorphism results in considerable changes in the predicted secondary structure of the pre-mRNA molecule containing both the ESE and ESS elements and may, therefore, have an effect on splicing regulation in two ways: (i) by affecting binding of splicing regulators directly through exposure or occlusion of these elements and, (ii) by altering the relative distance between these elements which leads to alterations in splice site usage and efficiency.[43,111] This is noteworthy because intron/exon boundaries occur 51bp and 102bp up- and downstream of this polymorphism respectively. It is therefore possible that the rs13321 polymorphism may subtly affect the function of the TN-C protein in one, or all, of three ways: (i) altering its' interaction with other proteins at the fibrinogen globe, (ii) altering transcriptional regulation, and subsequent expression, by a GATA transcription factor, or (iii) altering splicing regulation in the region of the SNP.

Rs2104772 and the GT dinucleotide repeat polymorphism both occur in the DNA sequence coding for the fibronectin type III (FnIII) repeat domains of the TN-C protein. These domains confer elasticity and are highly susceptible to protein degradation during remodelling of TN-C containing matrices.[136] The T>A substitution in rs2104772 does not result in any noteworthy changes in the predicted pre-mRNA secondary structure, however it does result in an amino acid change from Leucine to Isoleucine (Leu1677Ile) within the alternatively spliced FnIII-D domain. Matsuda et al.[190] previously suggested that Ile1677 results in steric hindrance with Phe1636 and, consequently, instability of the beta-sheet and alterations in the molecular elasticity of this domain. As TN-C is upregulated in wound healing, tissue remodelling,[137] tendinopathy[133,221] as well as in response to increased load,[129] it is reasonable to hypothesize that it plays a vital role in the adaptation of tendon tissue during the initial reactive and the tendon disrepair phases of tendinopathy.[62] Subtle alterations in its protein interactions, transcription levels and elastic properties may result in an inappropriate healing response during the early stages of reactive tendinopathy and predispose the tendon to early progression towards advanced stages of tendinopathy.

The majority of limitations to this study have already been discussed in chapter two. These include the heavier weight of the TEN participants when compared to the CON participants

University of Cape Town

and the lack of physical activity and smoking data for the Australian participants. As with rs28494505 in chapter two, there was a low genotype call rate (76%) for rs1330363 within the *TNC* gene in this study which is recognised as a limitation to this study. For this reason, the finding that the G-allele of this SNP was significantly associated with AT should be interpreted with caution. Although not necessarily a limitation to this study, it should be noted that the S-fold predicted changes in pre-mRNA secondary structure are generated using a short stretch of DNA (<500bp) surrounding each SNP and do not take the cellular milieu into account. Notwithstanding, it is of value to note that a simple base-pair substitution can result in considerable changes in pre-mRNA secondary structures.

3.5. CONCLUSION

In conclusion, the main finding of this study is the further implication of the genomic region containing the *TNC* and *COL27A1* genes in influencing risk of Achilles tendinopathy. This study maps the potential risk allele to a 759kbp region containing the GCA haplotype (rs946053; rs13321; rs2104772), as well as the 3'-end of the *COL27A1* gene (exons 42-61, 3'UTR) and the 5'-end of the *TNC* gene (5'UTR, exons 1-17). The region implicated by this haplotype may have functional effects on the transcription, structure and properties of TN-C and the α1 chain of type XXVII collagen. One can further hypothesize that these effects on protein properties may modify subsequent protein-protein interactions, or other protein-environment interactions thereby compromising wound healing, tissue remodelling and adaptation. Further research is required to replicate this finding in an independent study and to explore the functional significance of this interval spanning the two genes. Understanding the functional mechanisms underlying genetic associations is important in elucidating disease mechanisms underlying AT. Only once we fully understand the disease processes can we start to design effective methodologies to prevent and treat these injuries.

University of Cape Town

CHAPTER 4: COLLAGENS OF THE EXTRACELLULAR MATRIX AND ACHILLES TENDINOPATHY

4.1. INTRODUCTION

The predominant protein in the extracellular matrix (ECM) of tendons is collagen. In particular, type I collagen fibrils are the basis of the hierarchical structure of tendons and determine the mechanical strength of tendon tissue.[78,115,234] Type V collagen is a minor fibrillar collagen that is co-expressed with type I collagen fibrils and regulates both the size and shape of these heterotypic fibrils during development and growth.[32,123,297] The triple helical domain of type V collagen monomers lays within the type I collagen fibrils with the NH_2-terminal domain exposed, and this exposed NH_2-terminal domain is responsible for the regulation of type I collagen fibril diameter by type V collagen.[32] In particular, the presence of increasing amounts of type V collagen results in progressively decreasing diameters of type I collagen fibrils, while reductions in type V collagen result in abnormally large fibrils being deposited into the matrix.[32] Type V collagen predominantly exists as heterotrimers of two α1(V) and one α2(V) chain of type V collagen, however homotrimers of the α3(V) chain of type V collagen have also been reported.[32,123] As described in chapter one, Mokone et al. (2006)[199] and September et al. (2009)[255] previously reported an association of the *BstUI* RFLP (rs12722) within the 3'-UTR of the *COL5A1* gene, which encodes α1(V), with chronic AT in both a South African and Australian group. In particular, the CC genotype of this SNP was associated with a significantly decreased risk of developing AT. Recent investigation of this 3'-UTR within the *COL5A1* gene found further polymorphisms to be associated with AT, which were suggested to alter the predicted secondary structure of the 3'-UTR and regulate mRNA stability.[6,164] In particular, the alleles associated with AT appear to increase the stability of *COL5A1* 3'-UTR mRNA.[164] The α2(V) chain is encoded by the *COL5A2* gene which has been mapped to chromosome 2q32. The α3(V) chain is encoded by the *COL5A3* gene which is expressed in the developing tendons and ligaments of murine embryos and has been mapped to chromosome 19p13.2.[123] Like *COL5A1*, mutations within the *COL5A2* gene have also been associated with various forms of EDS and abnormal fibrillogenesis.[32,193,238] The *COL5A3* gene has been suggested as a suitable candidate locus for investigation in cases

University of Cape Town

of classical EDS in which these *COL5A1* and *COL5A2* gene variants have been excluded.[123] This emphasizes the importance of type V collagen fibrils in the ECM of connective tissue.

Type III collagen is a major fibrillar collagen and is important in the healing process of tendons.[173] It is laid down early in the repair process and rapidly forms disulphide cross-links which stabilise the newly synthesized matrix of the repair site.[52] These cross-links are easily degraded which allows for rapid growth and remodelling of tissue.[52] However, type III collagen fibres are thinner than type I collagen fibres and their presence therefore causes a decrease in the tensile strength of the tissue.[78] This is particularly relevant in AT where degeneration of tendon tissue leads to a chronic healing response and results in a slow accumulation of type III collagen.[78] Type III collagen exists as a homotrimer of three α1(III) chains,[268] encoded by the *COL3A1* gene which lays in tail-to-tail orientation 22kbp apart from *COL5A2*.[285] A missense polymorphism (rs1800255) in exon 30 of this gene has previously been associated with other connective tissue disorders such as pelvic organ prolapse and floppy mitral valve/mitral valve prolapse.[51,54,134,156]

The primary aim of this study was, therefore, to investigate the association of polymorphisms within the *COL3A1*, *COL5A2* and *COL5A3* genes with AT in South African and Australian groups. A second aim was to investigate interactions between these polymorphisms and the *BstUI* RFLP within *COL5A1* in modulating risk of AT.

University of Cape Town

4.2. MATERIALS AND METHODS

4.2.1. PARTICIPANTS

The SA-TEN (n=94), SA-CON (n=131), AUS-TEN (n=85) and AUS-CON (n=209) groups described in chapter two were used in this study. Participant characteristics are presented in Table 2.2 and described in 2.3.1.

4.2.2. VARIANT SELECTION

Lists of SNPs within the *COL5A3, COL5A2* and *COL3A1* genes were generated using the NCBI SNP database available at http://www.ncbi.nlm.nih.gov/snp/ and were interrogated using the Ensembl Genome Browser (http://www.ensembl.org/index.html) and the Genome Variation Server (GVS)(http://gvs.gs.washington.edu/GVS137/).[283] SNPs were selected for analysis based on their location within conserved regions of the genes, reported minor allele frequency (>0.20), reported heterozygosity (>0.35), biological significance and previously reported associations with other multi-factorial conditions.

Three SNPs within the *COL5A3* gene were investigated (Figure 4.1). The synonymous 453C>A transversion in exon 4 (rs2303099) does not change the alanine residue at position 151, the synonymous 1425C>G transversion in exon 14 (rs1559186) also does not change the leucine residue at position 475 but is situated in a splicing regulatory region, and the missense 3125G>C transversion in exon 42 (rs2161468) results in a change from arginine to proline at position 1042 (Arg1042Pro) and is situated in a splicing regulatory region. The GVS lists a high conservation score for SNPs rs2161468 (0.6130) and rs2303099 (0.8150), and in addition lists all three SNPs as Tag SNPs.[283]

Three SNPs within the *COL3A1* gene were selected for investigation (Figure 4.2). The two SNPs in intron 1 (rs2056156; 80-1018T>C) and intron 2 (rs3106796; 282+85A>G) are situated within regulatory regions, and the 2092G>A missense polymorphism (rs1800255) in exon 30 results in a change from an alanine to a threonine residue at position 698 (A698T). It should be noted that this SNP is often referred to as the "exon 31 polymorphism",[54,134] however the Ensembl Genome Browser localises this SNP to exon 30. In addition, all three SNP's are identified as Tag SNPs by the GVS.[283]

University of Cape Town

Two SNPs in *COL5A2* were also investigated (rs13031549; 97+13418A>C)(rs4667264; 97+34516G>C)(Figure 4.3), both of which are in intron 1. The rs4667264 SNP has a high conservation score (0.3310) listed on the GVS.[283]

4.2.3. ALLELIC DISCRIMINATION

SA-CON, SA-TEN, AUS-CON and AUS-TEN participants were genotyped for all eight polymorphisms. Standard PCR based RFLP techniques were used to amplify and genotype rs1559186. PCR reactions were carried out in a solution containing 200ng DNA template, 1X reaction buffer, 1.5mM Mg^{2+}, 2.5mM each of dATP, dCTP, dTTP and dGTP, 20pmol each of the forward and reverse primers and 0.5 units of either *Taq* DNA polymerase (*New England Biolabs® Inc.,* Ipswich, Massachusetts, USA) or SuperTherm *Taq* DNA polymerase (*Hoffmann-La-Roche*, US) using an XP Thermal Cycler Block (*Bioer Technology Co.*, Middlesex, UK). Reaction buffer was specific to the *Taq* used. Thermo-Pol® reaction buffer (*New England Biolabs® Inc.,* Ipswich, Massachusetts, USA) consisted of 10mM KCl, 10mM $(NH_4)SO_4$, 20mM Tris-HCl, 2mM $MgSO_4$ and 0.1% Triton X-100 at pH8.8. SuperTherm reaction buffer (*Hoffmann-La-Roche*, US) was provided with the SuperTherm *Taq* DNA polymerase and used as per manufacturer's instructions. PCR reactions were initiated with a five minute denaturing cycle at 94˚C, followed by 35 cycles consisting of 30 seconds denaturation at 94˚C, 30 seconds annealing at 58˚C and 40 seconds extension at 72˚C with a final five minute extension step at 72˚C. The oligonucleotide primer sequences, reaction conditions and amplicon sizes are listed in Appendix B. PCR amplicons were digested overnight for 15 hours with the *PvuII* restriction endonuclease as per manufacturer's guidelines. PCR amplicons and restriction fragments were mixed with *SYBER® Gold* nucleic acid gel stain (*Invitrogen Molecular Probes*™, Oregon, USA) and, together with a 100bp DNA size standard (*Promega Corporation*, Madison, Wisconsin, USA), separated by electrophoresis on 2% agarose (100V for 35 minutes) and 6% PAGE (120V for 2 hours) gels respectively. All gels were photographed under UV light using a photodocumentation system (*Uvitec Limited*, Cambridge, UK) [Appendix B].

University of Cape Town

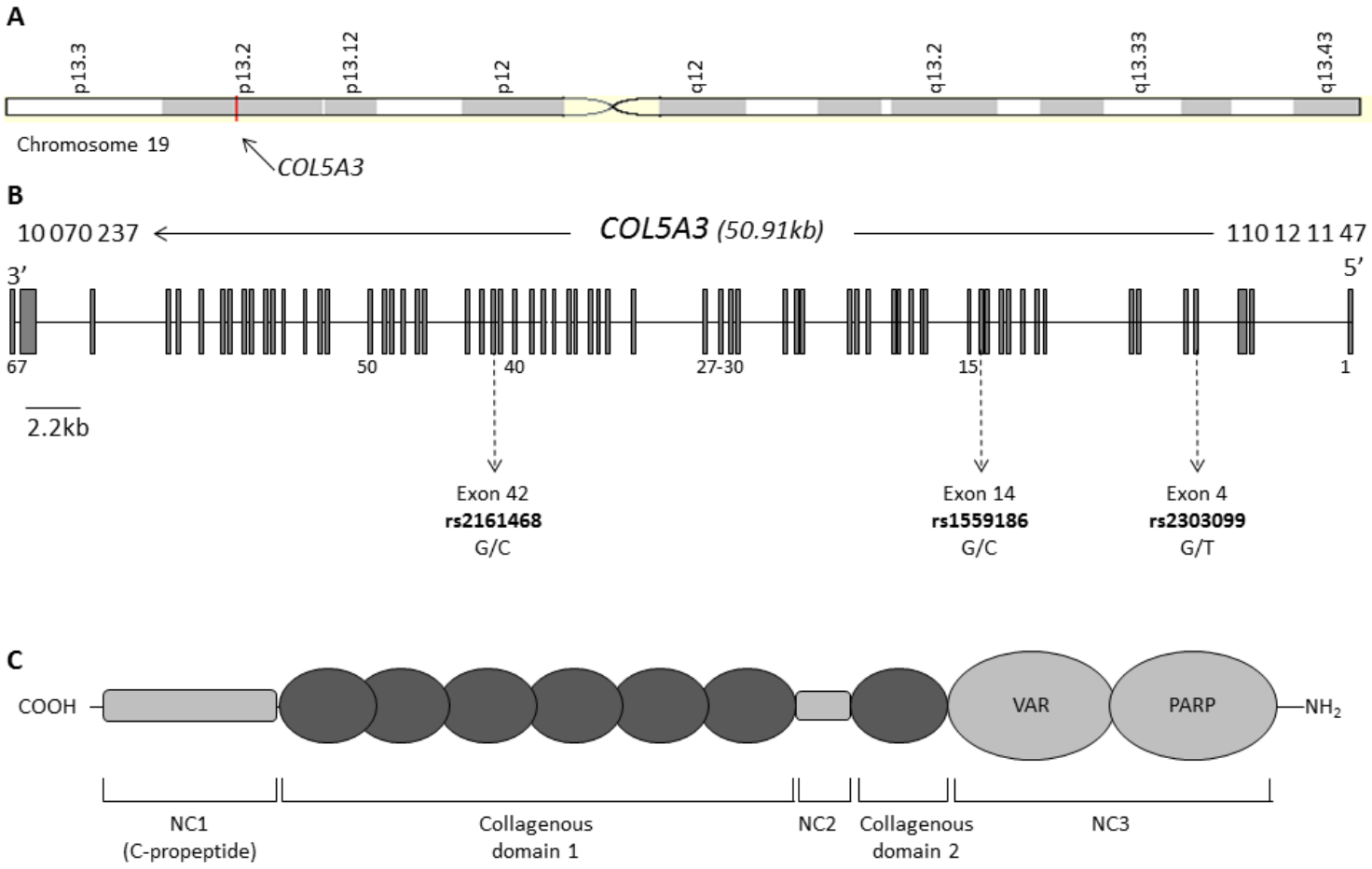

Figure 4.1: The alpha -3 chain of type V collagen

(A) Shaded bands indicate cytogenetic banding patterns for chromosome 19. Chromosomal banding position of COL5A3 gene is indicated with a n arrow (B) Genomic location and intron/exon structure of COL5A3. Solid bars represent numbered exons separated by introns (solid lines). Variants in bold are included in the present study. ***(C)*** *The α3 chain of type V collagen protein domain architecture. NC: Non-collagenous; VAR: Variable domain; PARP: Proline/arginine rich protein [Compiled from www.ensembl.org and Imamura et al. (2000)[123]]*

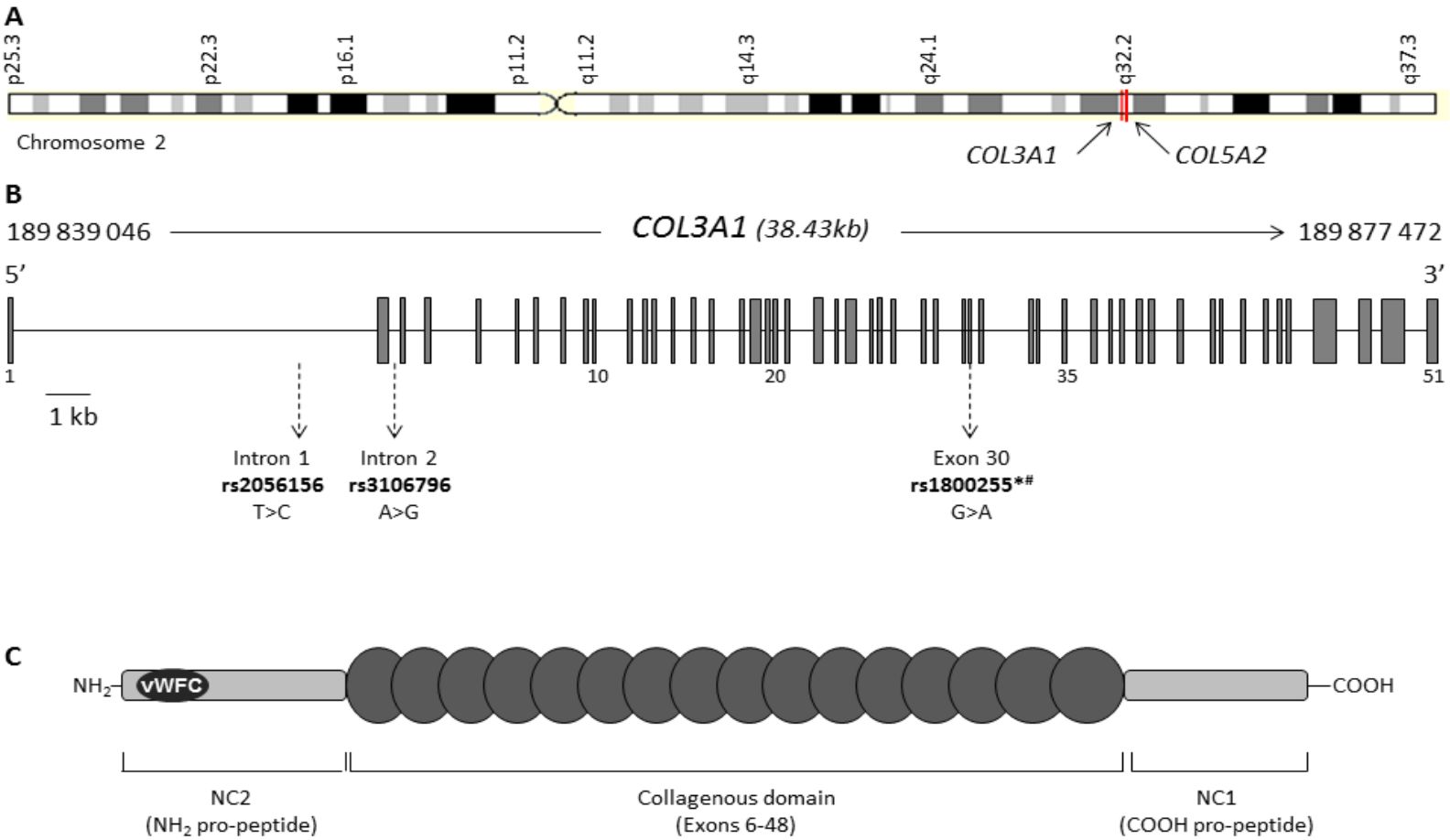

Figure 4.2: The alpha-1 chain of type III collagen

(A) *Shaded bands indicate cytogenetic banding patterns for chromosome 9. Chromosomal banding position of COL3A1 and COL5A2 genes are indicated with arrows* ***(B)*** *Genomic location and intron/exon structure of COL3A1. Solid bars represent numbered exons separated by intron (solid lines). Variants in bold are included in the present study.*Associated with pelvic organ prolapse.*[51,134,156] *#Associated with floppy mitral valve/mitral valve prolapse.*[54] ***(C)*** *The α1 chain of type III collagen protein domain architecture NC: Non-collagenous; vWFC: von Willebrand factor type C [Compiled from www.ensembl.org and Valkkila et al. (2001)*[119]*]*

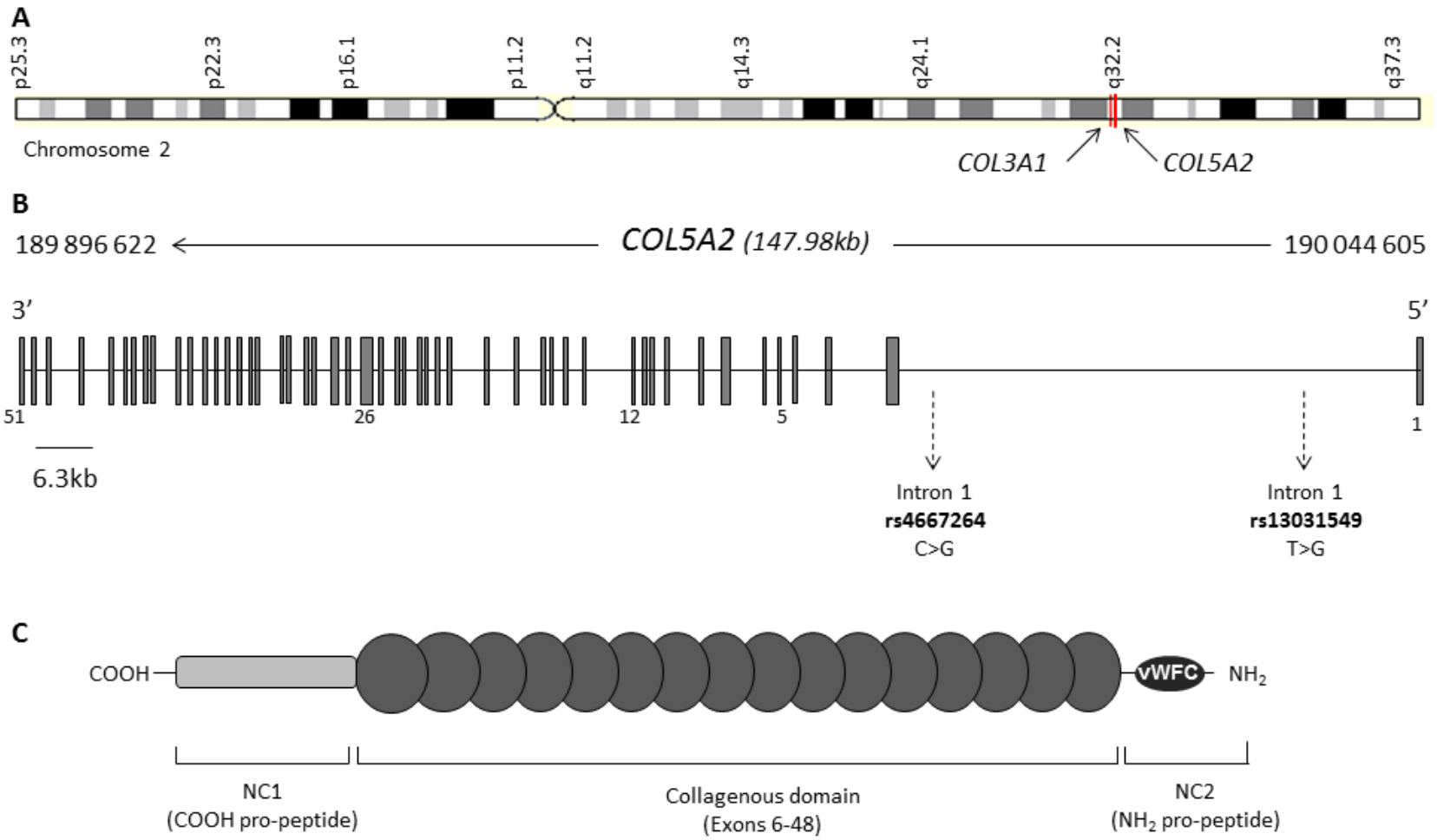

Figure 4.3: The alpha-2 chain of type V collagen

(A) *Shaded bands indicate cytogenetic banding patterns for chromosome 9. Chromosomal banding position of COL3A1 and COL5A2 genes are indicated with arrows* ***(B)*** *Genomic location and intron/exon structure of COL5A2. Solid bars represent numbered exons separated by intron (solid lines). Variants in bold are included in the present study.* ***(C)*** *The α2 chain of type V collagen protein domain architectecture. NC: Non-collagenous; vWFC: von Willebrand factor type C [Compiled from www.ensembl.org and Valkkila et al. (2001)[119]]*

All remaining polymorphisms (rs2303099, rs2161468, rs13031549, rs4667264, rs2056156, rs3106796, and rs1800255) were genotyped using inventoried TaqMan® SNP Genotyping Assays (*Applied Biosystems™*) that were amplified and distinguished using the StepOnePlus™ Real-Time PCR System (*Applied Biosystems™*). Briefly, real-time PCR reactions were performed in a final volume of 8µl and included 20ng DNA, TaqMan® Assay Mix and TaqMan® Genotyping Master Mix (2X). TaqMan® Assay Mix consisted of sequence specific forward and reverse primers to amplify the polymorphic sequence, and two MGB probes labelled with VIC® and FAM™ dyes to detect allele 1 and allele 2 respectively. Real-time PCR cycling started with an initial hold step at 95°C for 10 minutes followed by 40 cycles of 92°C for 15 seconds, and 60°C for 1 minute. Fluorescence signal for each MGB probe was used to discriminate alleles and genotypes using the StepOnePlus™ Real-Time PCR System software [Appendix B]. For quality control purposes, a number of positive controls of known genotype and DNA-free controls were randomly included on every 96-well PCR plate. In addition, a subset of samples were genotyped twice to ensure genotyping was consistent. The number and proportion of successfully genotyped participants for each of the polymorphisms is presented in Table 4.1. Below normal genotype call rates were observed for the AUS samples at the rs1559186 (88%), rs2303099 (78%) and rs13031549 (86%) loci.

University of Cape Town

4.2.4. STATISTICAL ANALYSIS

The programming environment R[230] and R packages were used for all analyses. The R package, Genetics,[296] was used to estimate genotype and allele frequencies and HWE probabilities. Logistic regression was used to compare genotype, allele and allele-combination frequencies between the TEN and CON groups, as well as SA and AUS groups. Results with a P-value of less than 0.05 were accepted as significant. Haplotype frequencies were inferred for both the SA and AUS groups using the R package, haplo.stats.,[249,266] and were based on the genotype data of the polymorphisms within each gene. All analyses were corrected for potential confounding by including age, sex, and group (SA or AUS) in the models as fixed effects. No adjustments were made for multiple testing because it has been suggested that these corrections, such as Bonferroni, markedly overcorrect for an inflated false-positive rate and unnecessarily reduce power in genetic association studies where

background LD exists between SNPs.[207,220] Genotype effects on age, height, weight, BMI and sex were determined by a one-way analysis of variance (ANOVA) using STATISTICA version 10 (*StatSoft Inc.*, Tulsa, OK, USA). Statistical significance was accepted when $P<0.05$.

Table 4.1: Number and proportion of recruited participants successfully genotyped for the eight polymorphisms investigated in this study

Total number of participants		Number of participants genotyped															
		rs2161468		rs1559186		rs2303099		rs2056156		rs3106796		rs1800255		rs4467264		rs13031549	
AUS-CON	**209**	204	98%	181	87%	153	73%	190	91%	193	92%	194	93%	204	98%	176	84%
AUS-TEN	**85**	80	94%	78	92%	76	89%	77	91%	80	94%	79	93%	81	95%	78	92%
TOTAL	**294**	**284**	**97%**	**259**	**88%**	**229**	**78%**	**267**	**91%**	**273**	**93%**	**273**	**93%**	**285**	**97%**	**254**	**86%**
SA-CON	**131**	131	100%	130	99%	131	100%	130	99%	130	99%	129	98%	130	99%	128	98%
SA-TEN	**94**	93	99%	93	99%	93	99%	92	98%	91	97%	90	96%	93	99%	91	97%
TOTAL	**225**	**224**	**99%**	**223**	**99%**	**224**	**99%**	**222**	**99%**	**221**	**98%**	**219**	**97%**	**223**	**99%**	**219**	**97%**

Values are n and proportions of the total number of participants recruited

University of Cape Town

4.3. RESULTS

4.3.1. GENOTYPE AND ALLELE FREQUENCIES

Genotype and minor allele frequency distributions for each of the polymorphisms, together with the HWE P-values, are shown in Tables 4.2, 4.3 and 4.4. The frequency distributions of the SNPs tested in this study were found to differ significantly between the SA and AUS groups at the rs2303099 (P<0.001) and rs3106796 loci (P=0.034). The participants from the two groups were therefore summarised separately. All genotype distributions were in HWE except at the rs1559186 locus that deviated significantly in the AUS-TEN group (P=0.034). After adjusting for the potential confounder's age and sex, there were no significant differences in either the genotype or allele frequency distributions between the CON and TEN groups. There were no significant genotype effects on age, weight, height, BMI or sex (Table 4.5).

University of Cape Town

Table 4.2: Genotype and minor allele frequency distributions of the sequence variants investigated within the COL5A3 gene in control (CON) and Achilles tendinopathy (TEN) groups of South African (SA) and Australian (AUS) populations

	CON		TEN		P-Values	
	AUS	SA	AUS	SA	Country	Group
COL5A3:						
rs2303099						
N	153	131	76	93		
T/T	0.25	0.32	0.24	0.35	**<0.001**	0.890
T/G	0.43	0.54	0.49	0.54		
G/G	0.32	0.14	0.28	0.11		
G	0.54	0.41	0.52	0.38	**<0.001**	0.885
HWE	0.105	0.207	0.821	0.192		
rs1559186						
N	181	130	78	93		
C/C	0.50	0.45	0.44	0.35	0.386	0.050
C/G	0.42	0.42	0.53	0.54		
G/G	0.08	0.13	0.04	0.11		
C	0.29	0.34	0.30	0.38	0.185	0.594
HWE	1.000	0.558	**0.034**	0.192		
rs2161468						
N	204	131	80	93		
C/C	0.38	0.42	0.36	0.34	0.886	0.473
C/G	0.46	0.41	0.46	0.52		
G/G	0.16	0.17	0.18	0.14		
G	0.39	0.37	0.41	0.40	0.691	0.916
HWE	0.660	0.191	0.816	0.522		

P-values are for the difference between countries and between diagnostic groups respectively, adjusted for each other, age, sex and whether or not a person was investigated in his/her country of birth. The genotype P-value is calculated using a 2-degrees of freedom test, with genotypes as categories and the allelic P-value is calculated using an additive allelic model. HWE gives exact P-values from tests of Hardy–Weinberg equilibrium. N is number of samples genotyped. Bold P-values are <0.05

University of Cape Town

Table 4.3: Genotype and minor allele frequency distributions of the sequence variants investigated within the *COL3A1* gene in control (CON) and Achilles tendinopathy (TEN) groups of South African (SA) and Australian (AUS) populations

	CON		TEN		P-Values	
	AUS	**SA**	**AUS**	**SA**	**Country**	**Group**
***COL3A1*: rs2056156**						
N	190	130	77	92		
T/T	0.31	0.25	0.26	0.25	0.347	0.928
T/C	0.49	0.53	0.56	0.49		
C/C	0.20	0.22	0.18	0.26		
C	0.44	0.49	0.46	0.51	0.147	0.925
HWE	0.884	0.598	0.361	0.837		
rs3106796						
N	193	130	80	91		
A/A	0.33	0.24	0.30	0.25	0.100	0.900
A/G	0.48	0.52	0.51	0.48		
G/G	0.20	0.24	0.19	0.26		
G	0.44	0.50	0.44	0.51	**0.034**	0.825
HWE	0.663	0.726	0.823	0.834		
rs1800255						
N	194	129	79	90		
G/G	0.56	0.55	0.63	0.57	0.688	0.726
A/G	0.26	0.35	0.30	0.36		
A/A	0.08	0.10	0.06	0.08		
A	0.26	0.28	0.22	0.26	0.698	0.463
HWE	0.453	0.183	0.334	0.579		

P-values are for the difference between countries and between diagnostic groups respectively, adjusted for each other, age, sex and whether or not a person was investigated in his/her country of birth. The genotype P-value is calculated using a 2-degrees of freedom test, with genotypes as categories and the allelic P-value is calculated using an additive allelic model. HWE gives exact P-values from tests of Hardy–Weinberg equilibrium. N is number of samples genotyped. Bold P-values are <0.05

University of Cape Town

Table 4.4: Genotype and minor allele frequency distributions of the sequence variants investigated within the *COL5A2* gene in control (CON) and Achilles tendinopathy (TEN) groups of South African (SA) and Australian (AUS) populations

	CON		TEN		P-Values	
	AUS	SA	AUS	SA	Country	Group
COL5A2:						
rs4667264						
N	204	130	81	93		
A/A	0.50	0.44	0.47	0.45	0.270	0.915
A/G	0.38	0.48	0.44	0.44		
G/G	0.11	0.08	0.09	0.11		
G	0.30	0.32	0.31	0.33	0.629	0.820
HWE	0.185	0.422	0.800	1.000		
rs13031549						
N	176	128	78	91		
G/G	0.47	0.34	0.42	0.40	0.077	0.969
G/T	0.40	0.56	0.47	0.47		
T/T	0.13	0.09	0.10	0.13		
T	0.33	0.38	0.34	0.37	0.231	0.805
HWE	0.236	0.037	0.801	1.000		

P-values are for the difference between countries and between diagnostic groups respectively, adjusted for each other, age, sex and whether or not a person was investigated in his/her country of birth. The genotype P-value is calculated using a 2-degrees of freedom test, with genotypes as categories and the allelic P-value is calculated using an additive allelic model. HWE gives exact P-values from tests of Hardy–Weinberg equilibrium. N is number of samples genotyped. Bold P-values are <0.05

Table 4.5: Genotype effects of investigated collagen gene SNP's on physiological characteristics of participants

POLYMORPHISM	P-VALUES					
	AGE	AGE INJURED	HEIGHT	WEIGHT	BMI	SEX
rs2161468	0.202	0.329	0.627	0.618	0.764	0.915
rs1559186	0.123	0.565	0.699	0.428	0.219	0.746
rs2303099	0.190	0.124	0.521	0.395	0.417	0.105
rs2056156	0.278	0.763	0.156	0.110	0.324	0.405
rs3106796	0.635	0.988	0.330	0.430	0.250	0.640
rs1800255	0.083	0.084	0.351	0.817	0.864	0.293
rs4667264	0.561	0.914	0.949	0.665	0.726	0.678
rs13031549	0.928	0.851	0.436	0.518	0.622	0.496

Age: Age at recruitment; Age Injured: Age at recruitment (CON) & age at diagnosis (TEN); P-values for age, age injured, height, weight and BMI are determined by one-way ANOVA; P-value for sex is determined by Pearsons χ^2 analysis

University of Cape Town

4.3.2. HAPLOTYPE ANALYSIS

There were no inferred haplotypes significantly associated with AT within any of the *COL5A3*, *COL3A1* and *COL5A2* genes, nor within the combined *COL3A1* and *COL5A2* genes (Table 4.6). The most common haplotypes within the *COL5A3* gene (rs2161468-rs1559186-rs2303099) were CCT and CCG in both the CON and TEN groups, however the distribution of all the haplotypes differed significantly between countries (P=0.002). Similarly, there was a significant difference in the distribution of haplotypes between SA and AUS within the *COL3A1* gene (P=0.025), with the most commonly occurring haplotype being CGG (rs2056156-rs3106796-rs1800255). The predominant haplotype within the *COL5A2* gene was GG (rs4667264-rs13031549) and there were no differences in the distribution of haplotypes between either country (P=0.626) or group (P=0.682). In the combined *COL3A1* and *COL5A2* haplotypes, the CGGGG haplotype occurred at the highest frequency. Furthermore, there was a significant difference in the distribution of these haplotypes between country groups (P=0.040).

4.3.3. INTERACTIONS WITH *COL5A1* RS12722

The probability of gene-gene interactions between any of the eight SNP's investigated and the *COL5A1* rs12722 polymorphism in modulating AT risk is shown in Table 4.7. There were no significant interactions, nor trends towards significance, between any of these SNPs and rs12722 (P<0.300 for all SNPs).

University of Cape Town

Table 4.6: Inferred haplotype frequencies for SNPs within the *COL5A3*, *COL3A1* and *COL5A2* genes, as well as the combined *COL3A1* and *COL5A2* haplotype

	CON		TEN		P-Values	
	AUS	SA	AUS	SA	Country	Group
COL5A3 (rs2161468-rs1559186-rs2303099):					**0.002**	0.895
CCT	0.24	0.30	0.28	0.32		
CCG	0.29	0.23	0.26	0.17		
GGG	0.15	0.10	0.09	0.12		
GGT	0.07	0.15	0.14	0.11		
GCT	0.12	0.09	0.06	0.12		
GCG	0.05	0.03	0.11	0.05		
CGT	0.03	0.05	0.00	0.07		
CGG	0.04	0.04	0.07	0.04		
COL3A1 (rs2056156-rs3106796-rs1800255):					**0.025**	0.997
CGG	0.42	0.48	0.43	0.49		
TAG	0.30	0.23	0.35	0.24		
TAA	0.25	0.27	0.20	0.25		
TGG	0.01	0.01	0.01	0.01		
CAG	0.01		0.00	0.01		
CGA	0.00	0.00	0.01	0.01		
TGA	0.01		0.00			
CAA	0.00		0.01			
COL5A2 (rs4667264-rs13031549):					0.626	0.682
GG	0.66	0.63	0.66	0.63		
CT	0.30	0.32	0.32	0.33		
GT	0.04	0.05	0.03	0.04		
CG	0.00	0.00	0.00			
COL3A1 & COL5A2:					**0.040**	0.828
CGGGG	0.29	0.32	0.28	0.34		
TAGGG	0.24	0.14	0.26	0.14		
CGGCT	0.12	0.14	0.14	0.15		
TAAGG	0.11	0.16	0.10	0.14		
TAACT	0.12	0.09	0.10	0.10		
TAGCT	0.05	0.08	0.06	0.07		
TAGGT	0.01	0.01	0.02	0.03		
CGGGT	0.01	0.02	0.00	0.00		
TAAGT	0.01	0.02	0.00	0.01		
CGAGG	0.00	0.00	0.01	0.01		
TGGCT	0.01	0.01	0.01	0.01		
CAGGG	0.01	0.00	0.00	0.01		
CAACT	0.00	0.00	0.01	0.00		
TGAGG	0.01	0.00	0.00	0.00		

Values are inferred haplotype frequencies and P-value for a test of equality of those frequencies. Combined COL3A1 and COL5A2 haplotypes that were not observed in any of the groups were excluded from the table. P-values are adjusted for country or group, age, sex and "born here"

University of Cape Town

Table 4.7: Interactions between *COL5A1* BstUI RFLP (rs12722) and the eight polymorphisms investigated

GENE	POLYMORPHISM	INTERACTION WITH *COL5A1* rs12722 (P-VALUE)
COL5A3	rs2161468	0.499
	rs1559186	0.384
	rs2303099	0.712
COL3A1	rs2056156	0.706
	rs3106796	0.698
	rs1800255	0.699
COL5A2	rs4667264	0.498
	rs13031549	0.525

University of Cape Town

4.4. DISCUSSION

The primary finding of this case-control genetic association study is that there were no independent associations between any of the eight SNP's investigated within the *COL5A3*, *COL3A1* and *COL5A2* genes and AT in a South African or Australian group. The hypothesis driven selection of the candidate genes was based on their biological function within the ECM of tendons. Furthermore, the selection of the eight SNPs investigated in this study was also hypothesis driven and based on the heterozygosity, location, conservation, and biological significance of these SNPs. Despite this, only rs1559186 within the *COL5A3* gene tended towards an association with AT (P=0.050). As discussed in chapter two, this study was only sufficiently powered to detect large effects on risk of developing AT (OR>2.0). It is therefore possible that the rs1559186 SNP may have a small effect on risk of developing AT that is unlikely to be detected in the current study. This is corroborated by the observation that the distribution of genotype frequencies in rs1559186 deviates from HWE in the AUS-TEN group (P=0.034), as this may be expected if one of these alleles or genotypes confers a harmful or protective effect. As in chapter two, the current study does not exclude the possibility that other polymorphisms – either rare with large effects, or common with

smaller effects - within these three genes are associated with AT.[34,175] However, it should be noted that many rare polymorphisms within the *COL3A1* and *COL5A2* genes have been shown to cause the severe connective tissue disorder, Ehlers-Danlos syndrome.[193,238,268] It is therefore possible that biologically significant variants within these highly conserved genes are more likely to result in severe clinical conditions, rather than common diseases, because of the limited redundancy within the collagen protein network. It is perhaps more likely that common polymorphisms with a smaller effect would be associated with the common condition of AT.[175]

The α2(V) chain forms heterotrimers with two α1(V) chains, while three α3(V) chains form homotrimers in type V collagen fibrils. In addition, both type V collagen and type III collagen ($\alpha1(III)_3$) are involved in determining the diameter and tensile strength of type I collagen fibrils. For these reasons, it was decided to further explore gene-gene interactions between any of the eight selected collagen gene SNPs and the rs12722 SNP in *COL5A1* in modulating the risk of AT. There were, however, no significant interactions, nor trends towards significance, between rs12722 and any of the eight SNPs investigated, including rs1559186. We can therefore conclude that any interactions between these genes at these SNP loci do not significantly contribute to the risk of AT. These results would further suggest that even if rs1559186 does have a small effect on the risk of developing AT, it is independent of the effect that rs12722 exerts. The *COL5A1* rs12722 polymorphism and the polymorphisms in the *IL-1RN, IL-6* and *IL-1β* genes have collectively been implicated in modulating risk of AT via a proposed cell-signalling pathway.[257] It is therefore reasonable to hypothesise that other collagen genes, such as those investigated in this study, may also interact with these interleukin genes in this pathway. This hypothesis will be investigated in the next chapter.

As with other studies in this thesis, there were several limitations to this study. These include the heavier weight of the TEN participants when compared to the CON participants, the lack of physical activity and smoking data for the Australian participants, and a small sample size which only results in sufficient power to detect large effects on risk of AT (OR>2.0). In addition, the low genotype call rate for rs1559186 (88%), rs2303099 (78%) and rs13031549 (86%) in the AUS samples is a limitation to this study. The reasons for this have

University of Cape Town

been discussed in chapter two. The strength of this study however, lays in the hypothesis driven selection of candidate genes and SNPs for investigation.

4.5. CONCLUSION

The eight polymorphisms investigated within the *COL5A3*, *COL3A1* and *COL5A2* genes were not independently associated with AT in a South African or Australian group.

University of Cape Town

CHAPTER 5: GENE-GENE INTERACTIONS IN THE EXTRACELLULAR MATRIX AND RISK MODELS FOR ACHILLES TENDINOPATHY

5.1. INTRODUCTION

A recent study by September et al. (2011)[257] used a pathway-based approach to investigate the association of functional sequence variants within the *IL-1β* (rs1143627; rs16944), *IL-6* (rs1800795) and *IL-1RN* (rs2234663) genes with AT. Although, no independent associations were found between the polymorphisms investigated and AT, these cell-signalling pathway polymorphisms and the *COL5A1* rs12722 polymorphism were collectively implicated in modulating risk of AT. Another recent study found an association of functional polymorphisms within the gene encoding caspase-8, a component of the apoptosis signalling cascade, with increased risk of AT.[205] In particular, rs1045486 and rs3834129 within the *CASP8* gene were found to be associated with AT both independently and as a haplotype. Considered collectively, the evidence above suggests an interaction between these components of the cell-signalling pathways involved in regulating the ECM, and the biology of the collagen fibril in tendon. It is therefore reasonable to hypothesize that polymorphisms within genes encoding other components of the ECM may also interact with these cell-signalling pathways in modulating the risk of AT.

Chapter 3 of this thesis presented the association of the potentially functional GCA haplotype within the *COL27A1* and *TNC* genes with AT in SA and AUS groups. Chapter 4 of this thesis reported the absence of any association between polymorphisms within the *COL5A3*, *COL3A1* and *COL5A2* genes with AT, however a tendency towards an independent association with AT was reported for rs1559186 within the *COL5A3* gene. Given the biological role of the proteins encoded by these genes in the ECM, it was decided to investigate potential gene-gene interactions between these ECM genes (*COL27A1*, *TNC*, *COL5A3*, *COL3A1* and *COL5A2*) and previously genotyped polymorphisms within genes encoding components of the cell-signalling pathways (*IL-6*, *IL-1β* and *CASP8*).[199,205,255,257] The primary aim of this study was, therefore, to test the gene-gene interactions between polymorphisms within genes encoding components of the collagen fibril and components of

University of Cape Town

cell-signalling pathways within the ECM, and to investigate the relative contribution of these polymorphisms to overall genetic risk in a polygenic model.

The number of distinct genes encoding various ECM proteins with either similar or distinct biological functions that have been implicated in modulating risk of AT is growing.[59,258] In addition to the various interactions and haplotypes which have been shown to influence risk of AT, a number of polymorphisms within several genes have been independently associated with AT in both the SA and AUS populations (Figure 5.1). As this number of confirmed, independent associations between polymorphisms and AT grows, and the biology underlying these associations is elucidated, it may become increasingly pertinent to incorporate genetic testing into the clinical management of athletes or individuals who present with other risk factors. Therefore, the second aim of this study was to develop and evaluate a clinically relevant genetic risk assessment (GRA) model for AT using (i) polymorphisms that have previously been independently associated with AT in two populations (Figure 5.1), and (ii) polymorphisms with a definitive AT risk genotype that were implicated in this thesis in a haplotype or gene-gene interactions.

5.2. METHODS

5.2.1. PARTICIPANTS

The SA-TEN (n=94), SA-CON (n=131), AUS-TEN (n=85) and AUS-CON (n=209) groups described in chapter two were used in this study. Participant characteristics are presented in Table 2.2 and described in 2.3.1.

5.2.2. GENOTYPING

Genotype data for the polymorphisms investigated in chapter three and four of this thesis were used in this study. These included rs4143245 (*COL27A1*), rs1249744 (*COL27A1*), rs753085 (*COL27A1*), rs946053 (*COL27A1*), rs13321 (*TNC*), rs2104772 (*TNC*), rs1330363 (*TNC*), rs2161468 (*COL5A3*), rs1559186 (*COL5A3*), rs2303099 (*COL5A3*), rs2056156 (*COL3A1*), rs3106796 (*COL3A1*), rs1800255 (*COL3A1*), rs4667264 (*COL5A2*) and rs13031549 (*COL5A2*),. Previously reported genotype data for polymorphisms rs1800795 (*IL-6*), rs1143627 (*IL-1β*), rs16944 (*IL-1β*), rs1045485 (*CASP8*), rs3834129 (*CASP8*), rs12722

University of Cape Town

(*COL5A1*), rs71746744 (*COL5A1*), rs16399 (*COL5A1*), rs1134170 (*COL5A1*), rs4919510 (*MIR608*), rs143383 (*GDF5*) and rs4789932 (*TIMP2*) was obtained for the SA-CON, SA-TEN, AUS-CON and AUS-TEN participants.[6,76,199,205,224,255,257]

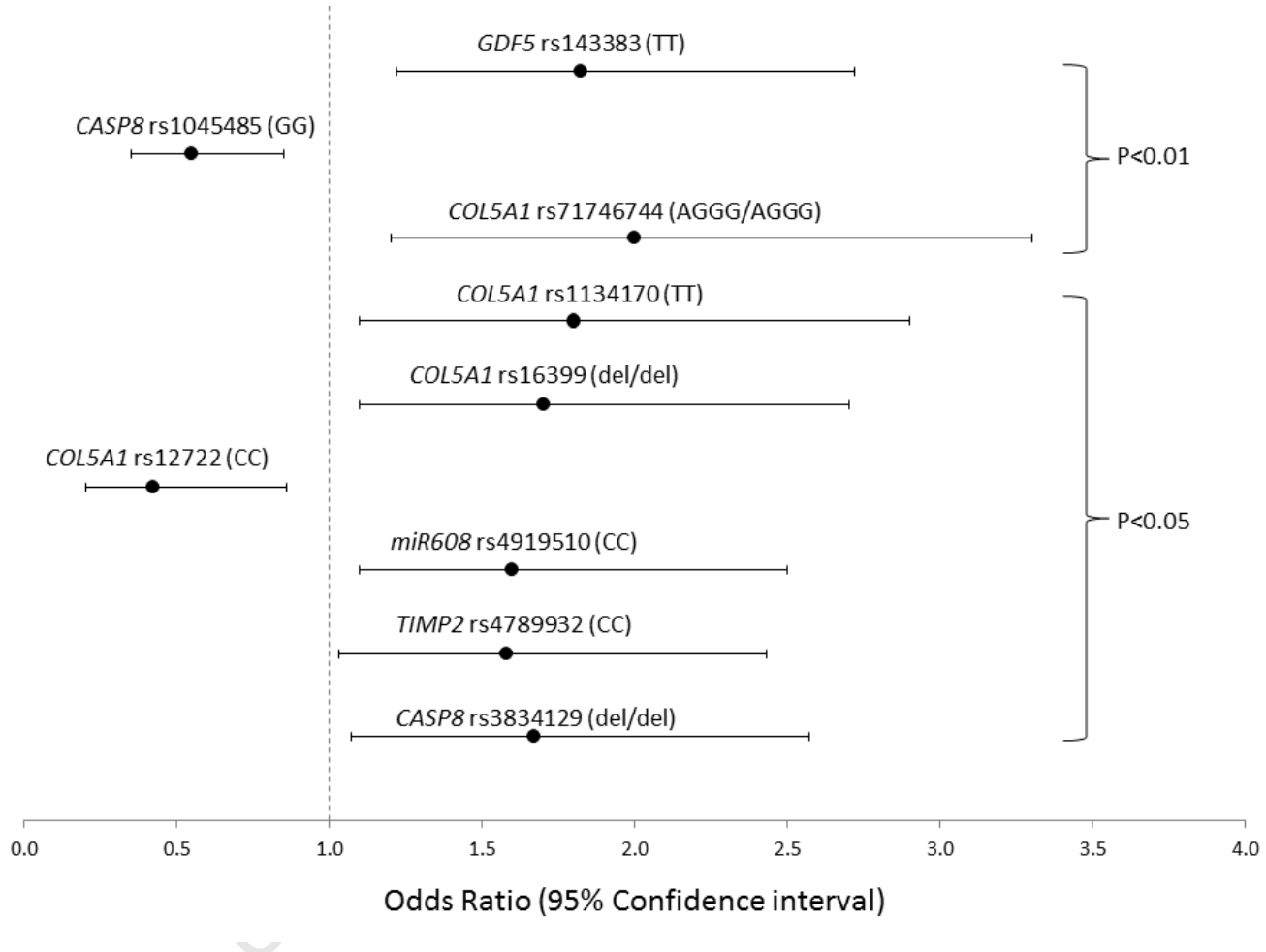

Figure 5.1: Polymorphisms previously associated with risk of Achilles tendinopathy in two populations

GDF5 rs143383: OR 1.82 (95% CI: 1.23-2.74) TT vs CC + CT; CASP8 rs1045485: OR 0.55 (95% CI: 0.35-0.84) GG vs CC + CG; COL5A1 rs71746744: OR 2.0 (95% CI: 1.2-3.3) AGGG/AGGG vs del/del + AGGG/del; COL5A1 rs1134170: OR 1.8 (95% CI: 1.1-2.9) TT vs AA + AT; COL5A1 rs16399: OR 1.7 (95% CI: 1.1-2.7) del/del vs ATCT/ATCT + ATCT/del; COL5A1 rs12722: OR 0.42 (95% CI: 0.20-0.86) CC vs TT + CT; miR608 rs4919510: OR 1.6 (95% CI: 1.1-2.5) CC vs GG + CG; TIMP2 rs4789932: OR 1.58 (95% CI: 1.03-2.43) CC vs TT + CT; CASP8 rs3834129: OR 1.67 (95% CI: 1.08-2.60) del/del vs AGTAAG/AGTAAG + AGTAAG/del

5.2.3. GENE-GENE INTERACTIONS AND LOGISTIC REGRESSION RISK MODELS FOR AT

The programming environment R[230] and R packages were used for all analyses. Stepwise logistic regression was used to compare allele-combination frequencies between the TEN and CON groups and to derive risk models for AT. Results with a P-value of less than 0.05 were accepted as significant. All analyses were corrected for potential confounding by including age, sex, and country group (SA or AUS) in the models as fixed effects. Two risk models (model A and B) were derived using stepwise logistic regression and receiver operating characteristic (ROC) curves were constructed, using the R package Epi,[45] to analyse the strength of each model.[83] Risk model A was derived by stepwise logistic regression of 20 variables: age, sex, country group, born "here" and genotype date for the 16 polymorphisms investigated within the *COL27A1*, *TNC*, *COL5A1*, *COL5A3*, *COL3A1* and *COL5A3* genes. These were discarded one by one until the model with the optimum specificity and sensitivity remained. Risk model B was derived by stepwise logistic regression of the same 20 variables above, as well as genotype date for the five polymorphisms investigated within the *IL-6*, *IL-1β* and *CASP8* genes. These were discarded one by one until the model with the optimum specificity and sensitivity remained. The area under each ROC curve (AUC) was used to quantify the overall ability of the model to discriminate between the TEN and CON groups based on genotype risk. No adjustments were made for multiple testing because it has been suggested that these corrections, such as Bonferroni, markedly overcorrect for an inflated false-positive rate and unnecessarily reduce power in genetic association studies where background LD exists between SNPs.[207,220]

5.2.4. GENETIC RISK ASSESSMENT MODELS

A stepwise approach was used to develop a series of potentially useful genetic risk assessment (GRA) models. Polymorphisms that were independently associated with AT in both populations were included in the GRA models (Figure 5.1). In addition, those polymorphisms with a clear at-risk or protective genotype that were implicated in either a haplotype or gene-gene interaction associated with AT were also included in the GRA models. Risk (and protective) genotype scores were calculated for each participant by allocating two points to the at-risk (or protective) genotype (Table 5.1 and 5.2), and zero

University of Cape Town

points for each of the other two genotypes. As the *COL5A1* polymorphisms rs71746744, rs16399 and rs1134170 are in strong LD[6] and inclusion of all three polymorphisms in the GRA models would therefore result in bias, only rs16399 was included based on the higher number of participants genotyped for this polymorphism. In each GRA model, risk (or protective) genotype scores for each polymorphism included in that model were summed for each participant. Differences in the distribution of risk (or protective) genotype scores between CON and TEN participants were calculated using Pearsons Chi-square analysis (STATISTICA version 10, *StatSoft Inc.*, Tulsa, OK, USA). Two-tailed Chi-square tests were used to determine specific differences in risk (or protective) genotype score between CON and TEN participants (GraphPad Prism 5.00 for Windows, *GraphPad Software*, San Diego, California, USA).

University of Cape Town

Table 5.1: List of polymorphisms and their risk genotypes included in genetic risk assessment (GRA) models 1-4

GENE	POLYMORPHISM	ASSOCIATION WITH AT	RISK GENOTYPE	ORIGINAL STUDY
COL5A1	rs71746744	Independent	AGGG/AGGG	Abrahams et al.[6]
COL5A1	rs16399	Independent	del/del	Abrahams et al.[6]
COL5A1	rs1134170	Independent	TT	Abrahams et al.[6]
miR608	rs4919510	Independent	CC	Abrahams et al.[6]
GDF5	rs143383	Independent	TT	Posthumus et al.[224]
CASP8	rs3834129	Independent & Haplotype	del/del	Nell et al.[205]
TNC	rs2104772	Haplotype	AA	Saunders et al.[247] Chapter 3
COL27A1	rs946053	Haplotype	GG	Saunders et al.[247] Chapter 3

Table 5.2: List of polymorphisms and their protective genotypes included in genetic risk assessment (GRA) models 5-7

GENE	POLYMORPHISM	ASSOCIATION WITH AT	PROTECTIVE GENOTYPE	ORIGINAL STUDY
COL5A1	rs12722	Independent	CC	Mokone et al.[199] September et al.[255]
CASP8	rs1045485	Independent	CC	Nell et al.[205]
TIMP2	rs4789932	Independent	CC	El Khoury et al.[76]
COL5A3	rs1559186	Gene-gene interaction	GG	Chapter 4
COL27A1	rs1249744	Gene-gene interaction	AA	Saunders et al.[247] Chapter 3

5.3. RESULTS

5.3.1. PARTICIPANT CHARACTERISTICS

Physiological characteristics for the TEN and CON groups are presented in chapter 2.3.1.

5.3.2. GENE-GENE INTERACTIONS

Frequencies were inferred for all allele-allele combinations between polymorphisms investigated in this thesis and previously genotyped polymorphisms within the cell-signalling pathways [Appendix C]. The distribution of these inferred allele-allele combinations was compared between the TEN and CON groups and the significance levels for differences in these combinations are shown in Table 5.3. Several significant associations with AT risk were noted. In particular, rs1249744 within the *COL27A1* gene significantly interacted with rs1800795 within the *IL-6* gene (P=0.004), rs16944 (P=0.030) and rs1143627 (P=0.046) within the *IL-1β* gene, and rs1045485 (P=0.026) and rs3834129 (P=0.034) within the *CASP8* gene to contribute to risk of AT. The *COL27A1* rs946053 polymorphism significantly interacted with *IL-6* rs1800795 (P=0.017) and *IL-1β* rs16944 (P=0.043). All three *TNC* polymorphisms significantly interacted with rs1800795 within the *IL-6* gene. The *TNC*

University of Cape Town

rs2104772 polymorphism further interacted with *IL-1β* rs16944 (P=0.034), *CASP8* rs1045485 (P=0.033) and *CASP8* rs3834129 (P=0.049), and *TNC* rs1330363 further interacted with *CASP8* rs1045485 (P=0.039) and *CASP8* rs3834129 (P=0.025). The *COL5A3* rs2303099 polymorphism significantly interacted with *IL-6* rs1800795 in the modulation of risk of AT (P=0.040).

The frequencies of specific allele combinations in the TEN and CON groups for all combinations are presented in appendix C. Briefly, the allele combinations for the significant interactions were generally consistent with the independent and haplotype associations observed for these polymorphisms. Allele combinations which included the *COL27A1* rs1249744 A-allele, *COL27A1* rs946053 T-allele, *TNC* rs2104772 T-allele, *TNC* rs1330363 A-allele, *IL-6* rs1800795 C-allele, rs16944 T-allele, rs1045485 C-allele and rs3834129 CTTACT-insertion allele occurred at a higher frequency in the CON group. Allele combinations which included the *COL27A1* rs1249744 G-allele, *COL27A1* rs946053 G-allele, *TNC* rs2104772 A-allele, *TNC* rs1330363 G-allele, *IL-6* rs1800795 G-allele, *IL-1β* rs16944 C-allele, *CASP8* rs1045485 G-allele and *CASP8* rs3834129 deletion allele occurred at a higher frequency in the TEN group. The only exception was the interaction between *TNC* rs13321 and *IL-6* rs180079, in which the distribution of allele combinations was the reverse of what would be expected from the haplotype association with AT presented in chapter three. The *TNC* rs13321 C-allele combined with the *IL-6* rs180079 C-allele was more frequently observed in the CON group (P=0.020), while the rs13321 G-allele in combination with the rs180079 G-allele was more frequently observed in the TEN group (P=0.018).

University of Cape Town

Table 5.3: Gene-gene interactions between genes encoding components of the ECM and cell-signalling pathways

	IL-6 rs1800795	*IL-1β* rs16944	*IL-1β* rs1143627	*CASP8* rs1045485	*CASP8* rs3834129
COL27A1 rs4143245	0.064	0.298	0.626	0.266	0.410
COL27A1 rs1249744	**0.004**	**0.030**	**0.046**	**0.026**	**0.034**
COL27A1 rs753085	0.080	0.270	0.526	0.361	0.380
COL27A1 rs946053	**0.017**	**0.043**	0.175	0.090	0.117
TNC rs13321	**0.039**	0.278	0.485	0.097	0.216
TNC rs2104772	**0.004**	**0.034**	0.090	**0.033**	**0.049**
TNC rs1330363	**0.012**	0.056	0.131	**0.039**	**0.025**
COL3A1 rs2056156	0.067	0.364	0.812	0.386	0.425
COL3A1 rs3106796	0.064	0.290	0.769	0.368	0.461
COL3A1 rs1800255	0.072	0.422	0.802	0.353	0.556
COL5A2 rs4667264	0.081	0.304	0.765	0.379	0.469
COL5A2 rs13031549	0.081	0.286	0.761	0.386	0.314
COL5A3 rs2161468	0.086	0.434	0.776	0.347	0.408
COL5A3 rs1559186	0.063	0.243	0.636	0.154	0.543
COL5A3 rs2303099	**0.040**	0.342	0.786	0.393	0.533

Values are P-values for interactions between the polymorphism listed on the left and the polymorphism listed on top in the modulation of risk of AT. Bold P-values are <0.05

University of Cape Town

5.3.3. LOGISTIC REGRESSION RISK MODELS FOR AT

As described in the methods, risk model A (Figure 5.2) for AT was derived by including 20 variables: age, sex, country group, born "here" and genotype data for 16 polymorphisms within the *COL27A1*, *TNC*, *COL5A1*, *COL5A3*, *COL3A1* and *COL5A2* genes. These were discarded one by one until the model with the optimum specificity and sensitivity remained. This best fit risk model included the variables country (SA), age (years), sex (male), *COL27A1* rs1249744 (AG, GG), *TNC* rs2104772 (AT, TT), *COL5A1* rs12722 (TC, TT) and *COL5A3* rs1559186 (CG, GG) (Figure 5.2)(Table 5.4). A ROC curve plot of the sensitivity and

specificity of the risk model when including/excluding different variables was generated. The optimal risk model had a sensitivity of 82.4%, where sensitivity is the proportion of participants who have the target condition (TEN) and are assessed by the model to be at high risk of AT. Specificity is the proportion of participants who do not have the target condition (CON) and are assessed to not be at risk of AT by the model,[83] and was calculated to be 58.0% for this model. The positive and negative predictive values were calculated to be 52.0% and 85.6% respectively. However these two statistics are of little value in case-control studies when the "diseased" and "not diseased" participants have been pre-selected. The AUC for the optimal risk model was 0.775.

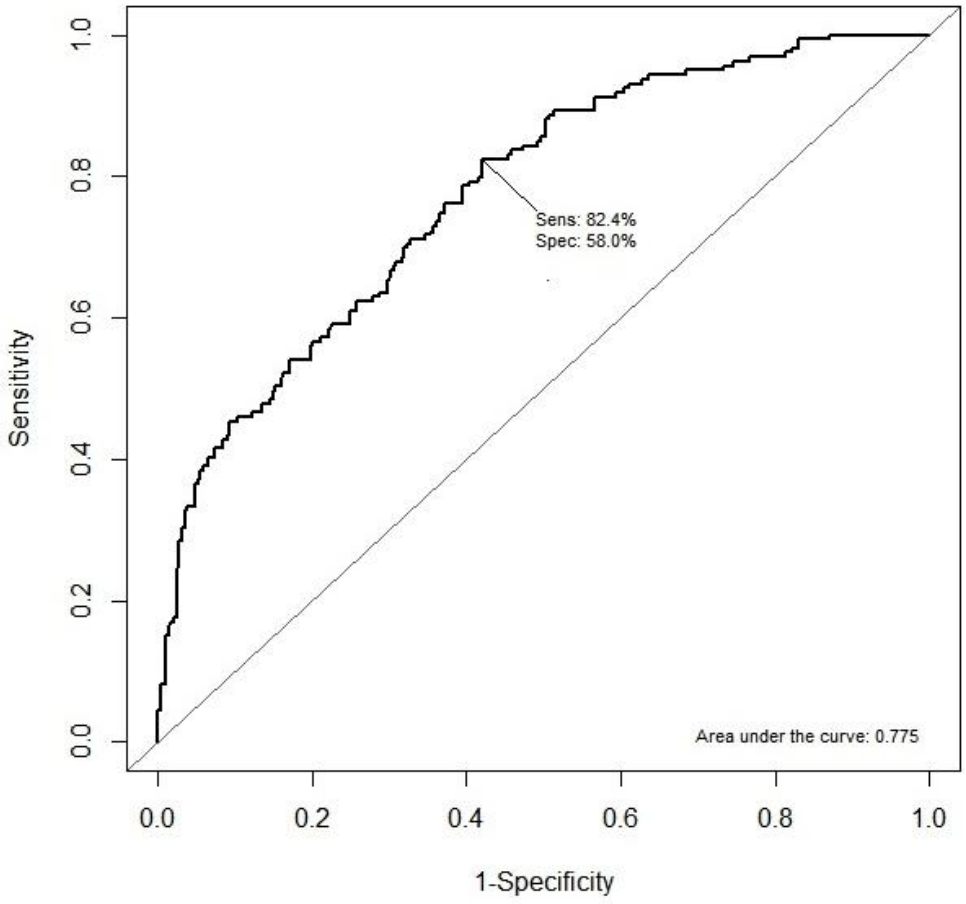

Figure 5.2: Receiver operating characteristic (ROC) curve of risk model A for Achilles tendinopathy showing sensitivity (Sens) and specificity (Spec) at the optimal cutpoint
ROC curve determines the true positive (sensitivity) versus true negative (specificity) rate for various cut-offs in determining risk of Achilles tendinopathy. The straight line indicates where sensitivity = 1-specificity and AUC = 0.5. The optimal cut-off which yields the maximum sensitivity plus specificity is indicated on the graph with an arrow, and the optimal model is summarised in Table 5.4

Table 5.4: List of variables included in risk model A for Achilles tendinopathy

Factor	Level	Coefficient	SE	P-value
Country of birth	AUS	0		
Country of birth	SA	0.43	0.23	0.065
Age	Years	0.07	0.01	**<0.001**
Sex	Female	0		
Sex	Male	0.74	0.24	**0.002**
COL27A1 rs1249744	AA	0		
COL27A1 rs1249744	AG	0.54	0.24	**0.024**
COL27A1 rs1249744	GG	0.22	0.42	0.600
TNC rs2104772	AA	0		
TNC rs2104772	AT	-0.39	0.25	0.122
TNC rs2104772	TT	-0.63	0.34	0.064
COL5A1 rs12722	CC	0		
COL5A1 rs12722	CT	1.04	0.31	0.001
COL5A1 rs12722	TT	0.43	0.34	0.197
COL5A3 rs1559186	CC	0		
COL5A3 rs1559186	CG	0.28	0.24	0.242
COL5A3 rs1559186	GG	-0.57	0.47	0.221

Coefficients are used to predict risk of AT, which is used to calculate points on the ROC curve. P-values are adjusted for other variables and assess the effect of each variable compared to its reference level indicated with coefficient 0. SE: Standard error of the coefficient.

University of Cape Town

Risk model B (Figure 5.3) for AT was derived by including the same 20 variables included in risk model A, as well as genotype data from polymorphisms within the *IL-6*, *IL-1β* and *CASP8* genes. The best fit risk model included the variables age (years), sex (male), *COL27A1* rs946053 (GT, TT), *COL5A1* rs12722 (TC, CC), *COL5A3* rs1559186 (CG, GG), *IL-6* rs1800795 (GC, CC), *CASP8* rs1045485 (GC, CC) and *CASP8* rs3834129 (del/CTTACT, del/del) (Table 5.5). Figure 5.3 shows the ROC curve analysis for the optimal risk model with a specificity of 86.5% and a sensitivity of 58.1%. The AUC for the optimal risk model was 0.801.

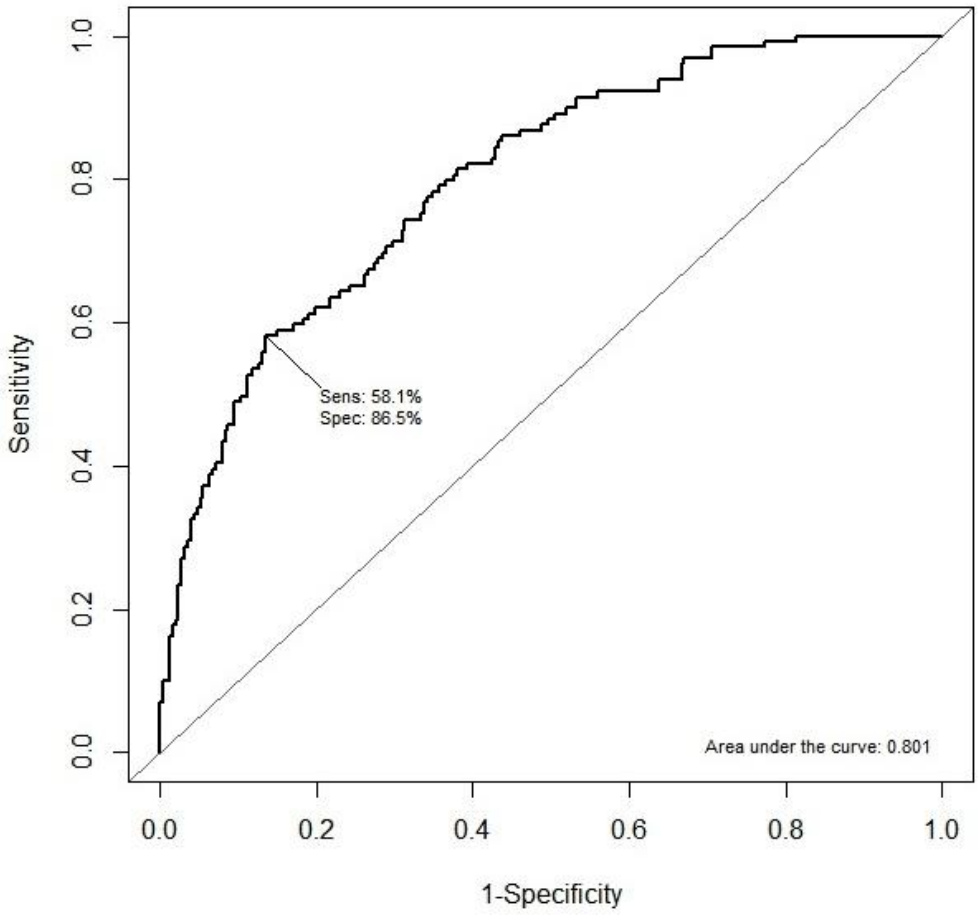

Figure 5.3: Receiver operating characteristic (ROC) curve of risk model B for Achilles tendinopathy showing sensitivity (Sens) and specificity (Spec) at the optimal cutpoint
ROC curve determines the true positive (sensitivity) versus true negative (specificity) rate for various cut-offs in determining risk of Achilles tendinopathy. The straight line indicates where sensitivity = 1-specificity and AUC = 0.5. The optimal cut-off which yields the maximum sensitivity plus specificity is indicated on the graph with an arrow, and the optimal model is summarised in Table 5.5

Table 5.5: List of variables included in risk model B for Achilles tendinopathy

Factor	Level	Coefficient	SE	P-value
Age	Years	0.073	0.01	**0.000**
Sex	Female	0		
Sex	Male	0.724	0.27	**0.007**
COL27A1 rs946053	GG	0		
COL27A1 rs946053	GT	-0.507	0.30	0.094
COL27A1 rs946053	TT	-1.098	0.38	**0.004**
COL5A1 rs12722	TT	0		
COL5A1 rs12722	TC	0.42	0.29	0.139
COL5A1 rs12722	CC	-0.49	0.38	0.194
COL5A3 rs1559186	CC	0		
COL5A3 rs1559186	CG	0.263	0.27	0.321
COL5A3 rs1559186	GG	-0.905	0.56	0.105
IL6 rs1800795	GG	0		
IL6 rs1800795	GC	-0.54	0.29	0.060
IL6 rs1800795	CC	-0.87	0.38	**0.023**
CASP8 rs1045485	GG	0		
CASP8 rs1045485	GC	-0.65	0.29	**0.024**
CASP8 rs1045485	CC	-0.75	1.31	0.566
CASP8 rs3834129	CTTACT/CTTACT	0		
CASP8 rs3834129	del/CTTACT	-0.265	0.32	0.402
CASP8 rs3834129	del/del	0.708	0.36	**0.048**

Coefficients are used to predict risk of AT, which is used to calculate points on the ROC curve. P-values are adjusted for other variables and assess the effect of each variable compared to its reference level indicated with coefficient 0. SE: Standard error of the coefficient

University of Cape Town

5.3.4. GENETIC RISK ASSESSMENT (GRA) MODELS

The previous analyses in this study show (i) that there are gene-gene interactions between ECM genes and cell-signalling pathway genes, and (ii) that some polymorphisms potentially contribute to the burden of risk of AT even if they are not independently associated with AT. However, the best fit risk models derived by logistic regression are complicated and, in the current form, difficult to interpret in a clinical setting. The underlying principle of a polygenic risk profile containing several implicated loci contributing to overall genetic risk of AT was therefore used to develop a more clinically relevant tool. A genetic risk assessment model using polymorphisms previously independently associated with risk of AT was developed first (Figure 5.4). Risk genotype scores for GRA model 1 were calculated using genotype data from rs16399 (*COL5A1*), rs4919510 (*miR608*), rs143383 (*GDF5*) and rs3834129 (*CASP8*) and therefore had a maximum score of eight (Table 5.1).[6,205,224] Participants with a risk genotype score of two or less were significantly more frequently observed in the CON group (OR 0.28, 95% CI: 0.15-0.51; P<0.001), whereas participants with a risk genotype score of six or more were significantly more frequently observed in the TEN group (OR 4.66, 95% CI: 2.51-8.68; P<0.001)(Figure 5.4).

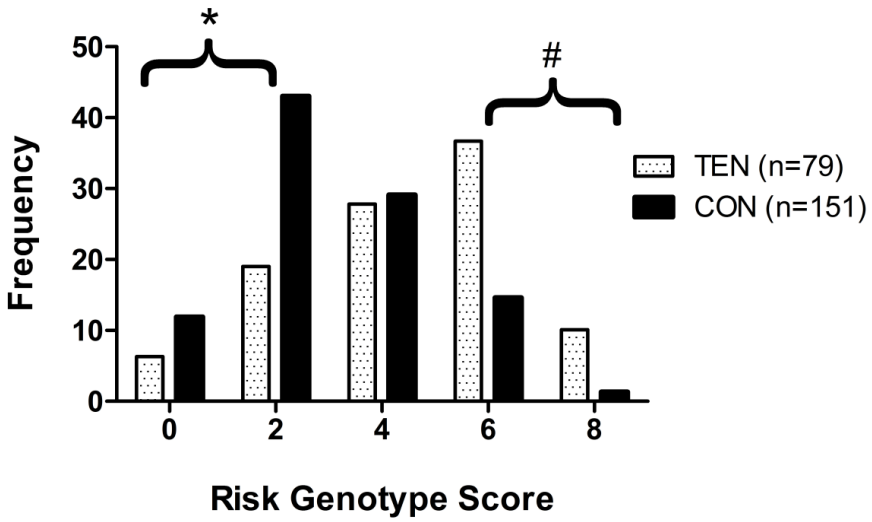

Figure 5.4: GENETIC RISK ASSESSMENT MODEL 1 - Distribution of risk genotype scores in participants with Achilles tendinopathy (TEN) and controls (CON)

Risk genotype score calculated using rs16399 (del/del), rs4919510 (CC), rs143383 (TT) and rs3834129 (del/del) with a maximum score of 8; n=combined SA and AUS

** 0-2 vs 4-8: χ^2 18.44, P<0.001, OR 0.28 (95% CI: 0.15-0.51)*

 # 6-8 vs 0-4: χ^2 25.48, P<0.001, OR 4.66 (95% CI: 2.51-8.68)

University of Cape Town

To assess whether the addition of two further genetic variables, with clear at-risk genotypes, would influence the efficacy of this model, a second GRA model was developed. Risk genotype scores for GRA model 2 were calculated using genotype data from rs946053 (*COL27A1*) and rs2104772 (*TNC*) only, and therefore had a maximum score of four (Table 5.1). Participants with a risk genotype score of four were significantly more frequently observed in the TEN group (OR 2.42, 95% CI: 1.35-4.32; P=0.002) (Figure 5.5).

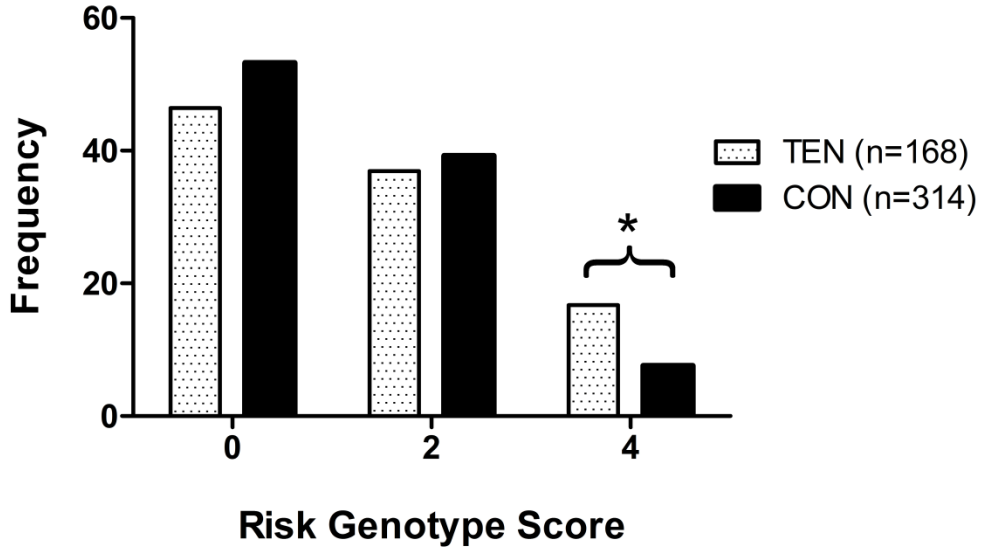

Figure 5.5: GENETIC RISK ASSESSMENT MODEL 2 - Distribution of risk genotype scores in participants with Achilles tendinopathy (TEN) and controls (CON)
Risk genotype score calculated using rs946053 (GG) and rs2104772 (AA) with a maximum score of 4; n=combined SA and AUS
** 0-2 vs 4: χ^2 9.26, P=0.002, OR 2.42 (95% CI: 1.35-4.32)*

University

Given the differences in the distribution of risk genotype scores observed in GRA model 2, two further GRA models were derived which sequentially tested the effect of adding the two new variables to GRA model 1. Risk genotype scores for GRA model 3 were calculated using genotype data from rs16399 (*COL5A1*), rs4919510 (*miR608*), rs143383 (*GDF5*), rs3834129 (*CASP8*) and rs2104772 (*TNC*), and therefore had a maximum score of ten (Table 5.1). Participants with a risk genotype score of eight or more were significantly more frequently observed in the TEN group (OR 6.88, 95% CI: 2.76-17.13; P<0.001) (Figure 5.6). Risk genotype scores for GRA model 4 were calculated using genotype data from rs16399

(*COL5A1*), rs4919510 (*miR608*), rs143383 (*GDF5*), rs3834129 (*CASP8*), rs2104772 (*TNC*) and rs946053 (*COL27A1*), and therefore had a maximum score of 12 (Table 5.1). Again, participants with a risk genotype score of eight or more were significantly more frequently observed in the TEN group (OR 5.89, 95% CI: 2.39-14.52; P<0.001) (Figure 5.7). A graphical representation comparing the ORs of GRA models 1-4 is shown in Figure 5.8. Taking into account the OR and the 95% CI for each model, GRA model 4 was chosen as the best model for distinguishing between TEN and CON participants. The proportion of TEN and CON participants that would be classified as either at-risk or not at-risk by GRA model 4 is presented in Figure 5.9. Thirty-seven percent (37%) of TEN participants are classified as at-risk, and 90% of CON participants are classified as not at-risk by GRA model 4.

Town

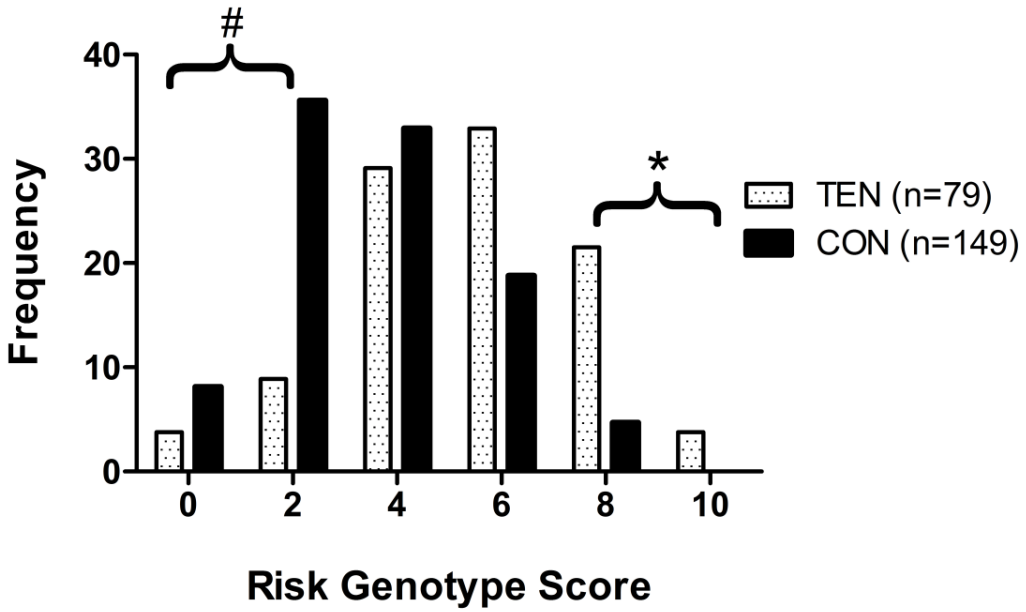

Figure 5.6: GENETIC RISK ASSESSMENT MODEL 3 - Distribution of risk genotype scores in participants with Achilles tendinopathy (TEN) and controls (CON)

Risk genotype score calculated using rs16399 (del/del), rs4919510 (CC), rs143383 (TT), rs3834129 (del/del) and rs2104772 (AA) with a maximum score of 10; n=combined SA and AUS

** 0-6 vs 8-10: χ^2 21.02, P<0.001, OR 6.88 (95% CI: 2.76-17.13)*

 # 0-2 vs 4-10: χ^2 22.43, P<0.001, OR 5.34 (95% CI: 2.55-11.17)

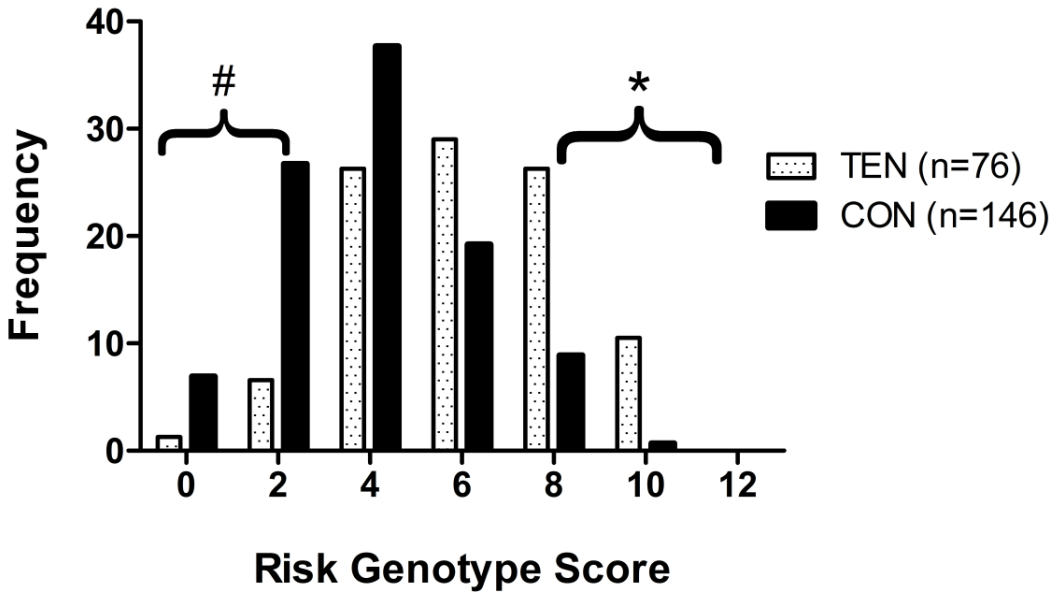

Figure 5.7: GENETIC RISK ASSESSMENT MODEL 4 - Distribution of risk genotype scores in participants with Achilles tendinopathy (TEN) and controls (CON)

Risk genotype score calculated using rs16399 (del/del), rs4919510 (CC), rs143383 (TT), rs3834129 (del/del), rs2104772 (AA) and rs946053 (GG) with a maximum score of 12; n=combined SA and AUS
** 0-6 vs 8-12: χ^2 24.20, P<0.001, OR 5.89 (95% CI: 2.39-14.52)*
 # 0-2 vs 4-12: χ^2 17.67, P<0.001, OR 5.34 (95% CI: 2.55-11.17)

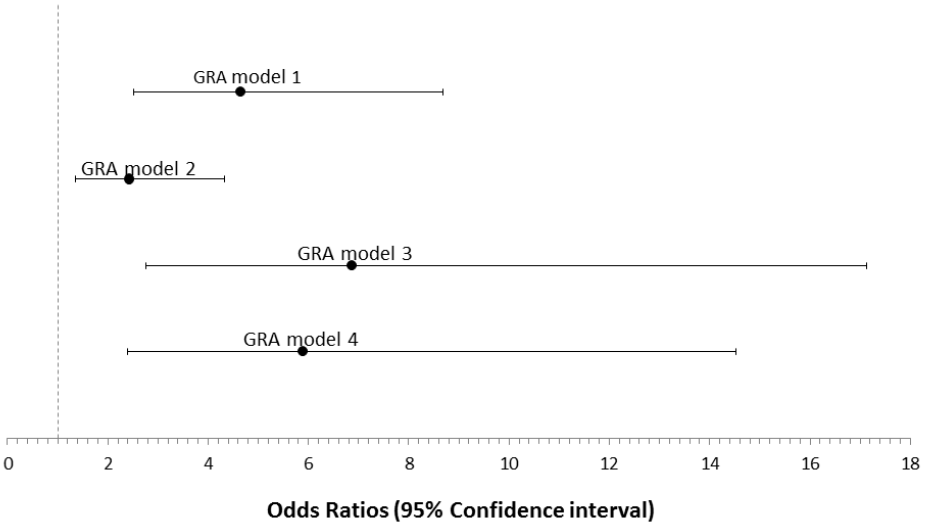

Figure 5.8: Odds ratios and 95% confidence intervals of genetic risk assessment (GRA) models 1-4

GRA model 1: rs16399, rs4919510, rs143383, rs3834129; GRA model 2: rs2104772, rs946053; GRA model 3: rs16399, rs4919510, rs143383, rs3834129, rs2104772; GRA model 4: rs16399, rs4919510, rs143383, rs3834129, rs2104772, rs946053

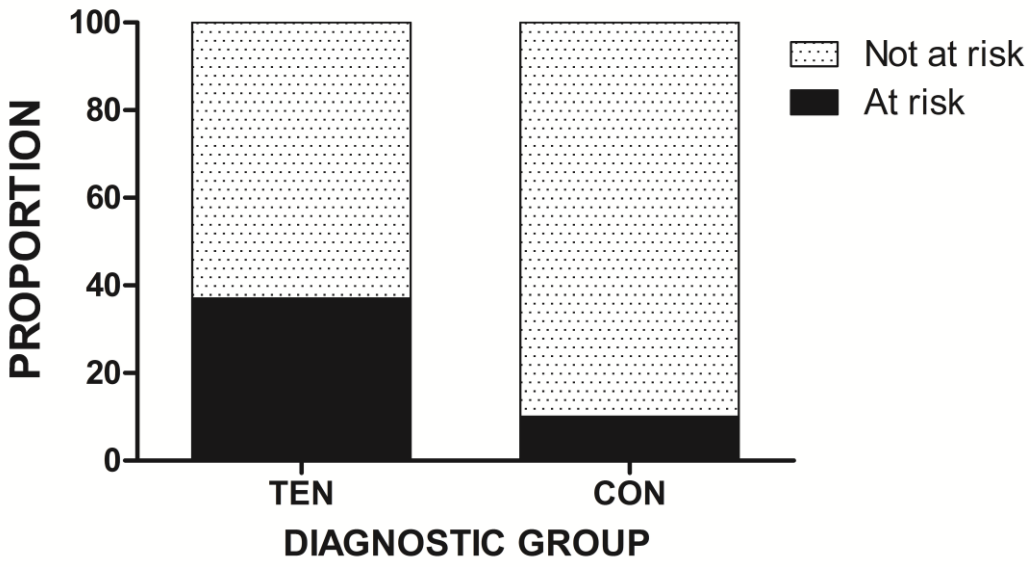

Figure 5.9: The proportion of TEN and CON participants classified as at risk or not at risk of Achilles tendinopathy by genetic risk assessment model 4

At risk: GRA model 4 risk genotype score ≥8; Not at risk: GRA model 4 risk genotype score ≤6; GRA model 4: rs16399, rs4919510, rs143383, rs3834129, rs2104772, rs946053. TEN: n=76, CON: n=146

A GRA model using polymorphisms previously independently associated with protection against AT was also sequentially developed. Protective genotype scores for GRA model 5 were calculated using genotype data from rs12722 (*COL5A1*), rs4789932 (*TIMP2*) and rs1045485 (*CASP8*) and therefore had a maximum score of six (Table 5.2).[76,199,205,255] Participants with a protective genotype score of two or more were significantly more frequently observed in the CON group compared to the TEN group (OR 0.44, 95% CI: 0.28-0.69; P<0.001)(Figure 5.10). The addition of rs1559186 (*COL5A3*) to this model resulted in a larger effect with participants who had a protective genotype score of two or larger once more observed more frequently in the CON group (OR 0.39, 95% CI: 0.23-0.65; P<0.001)(Figure 5.11). A similar result was observed when *COL27A1* rs1249744 was added to the model (OR 0.38, 95% CI: 0.23-0.64; P<0.001), however this addition did not increase the effect size (Figure 5.12). The ORs and 95% confidence intervals for each protective GRA model are shown in Figure 5.13. Taking into account the OR and 95% confidence interval for each model, GRA model 6 was chosen as the best protective model. The proportion of TEN

and CON participants that would be classified as either protected or not protected by GRA model 6 is presented in Figure 5.14. The large majority (83%) of CON participants were classified as protected, however 66% of TEN participants were also classified as protected by GRA model 6.

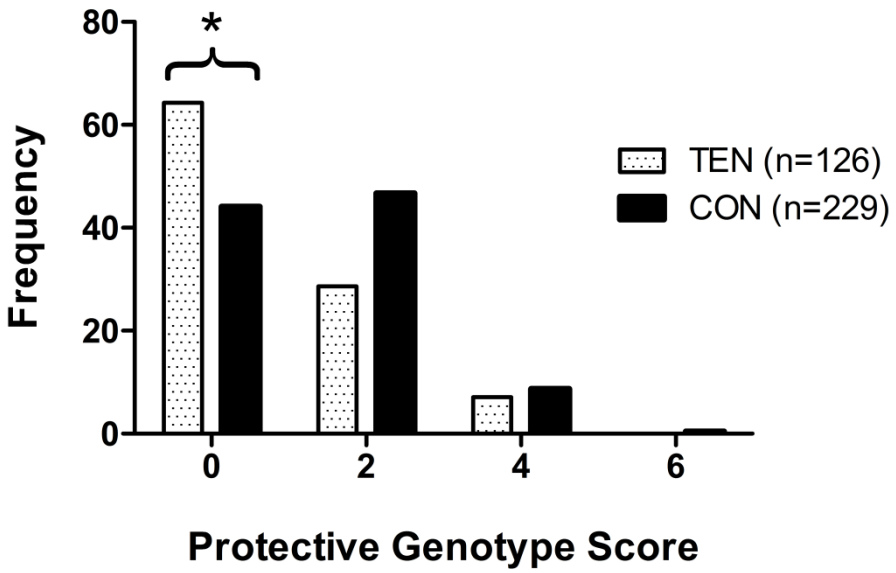

Figure 5.10: GENETIC RISK ASSESSMENT MODEL 5 - Distribution of protective genotype scores in participants with Achilles tendinopathy (TEN) and controls (CON)
Protective genotype score calculated using rs1045485 (CC), rs12722 (CC) and rs4789932 (CC) with a maximum score of 6; n=combined SA and AUS
** 0 vs 2-6: χ^2 13.25, P<0.001, OR 0.44 (95% CI: 0.28-0.69)*

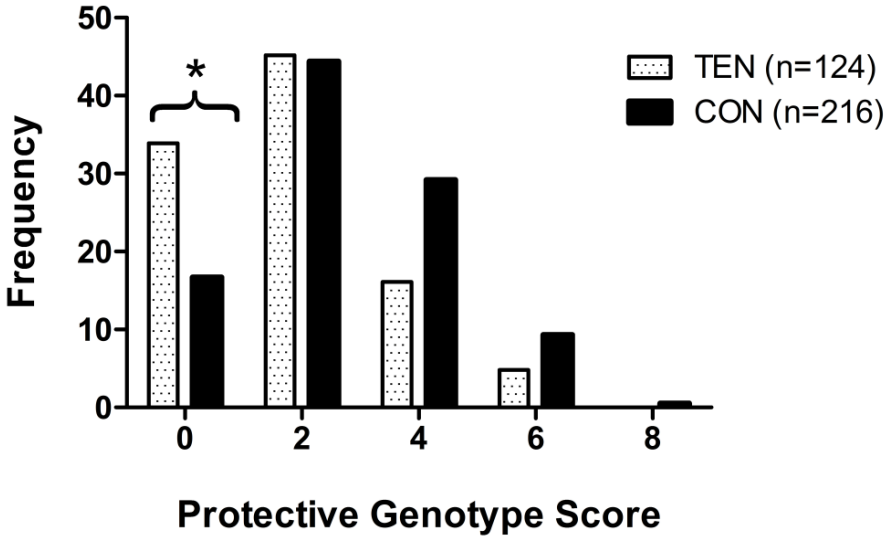

Figure 5.11: GENETIC RISK ASSESSMENT MODEL 6 - Distribution of protective genotype scores in participants with Achilles tendinopathy (TEN) and controls (CON)
Protective genotype score calculated using rs1045485 (CC), rs12722 (CC), rs4789932 (CC) and rs1559186 (GG) with a maximum score of 8; n=combined SA and AUS
** 0 vs 2-8: χ^2 13.19, P<0.001, OR 0.39 (95% CI: 0.23-0.65)*

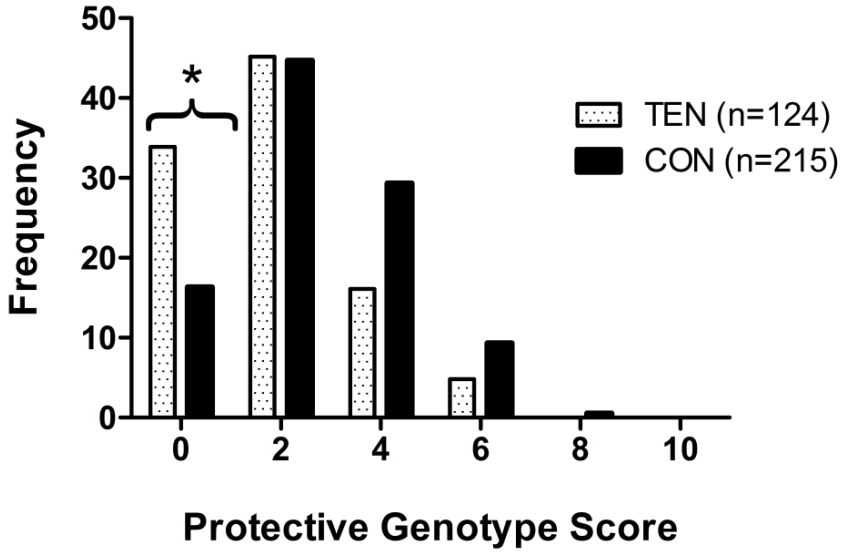

Figure 5.12: GENETIC RISK ASSESSMENT MODEL 7 - Distribution of protective genotype scores in participants with Achilles tendinopathy (TEN) and controls (CON)
Protective genotype score calculated using rs1045485 (CC), rs12722 (CC), rs4789932 (CC), rs1559186 (GG) and rs1249744 (AA) with a maximum score of 10; n=combined SA and AUS
** 0 vs 2-10: χ^2 13.86, P<0.001, OR 0.38 (95% CI: 0.23-0.64)*

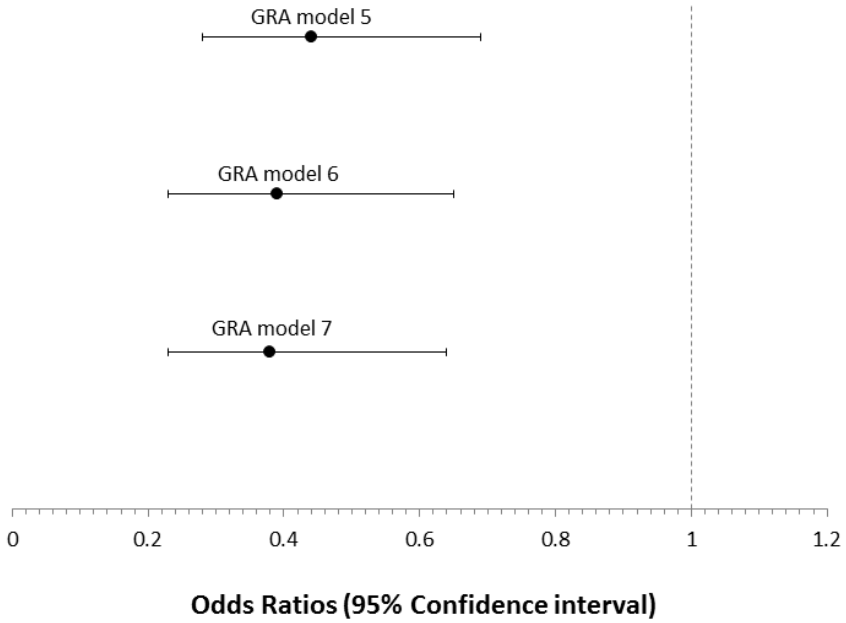

Figure 5.13: Odds ratios and 95% confidence intervals for protective genetic risk assessment (GRA) models 5-7

GRA model 5: rs1045485, rs12722, rs4789932; GRA model 6: rs1045485, rs12722, rs4789932, rs1559186; GRA model 7: rs1045485, rs12722, rs4789932, rs1559186, rs1249744

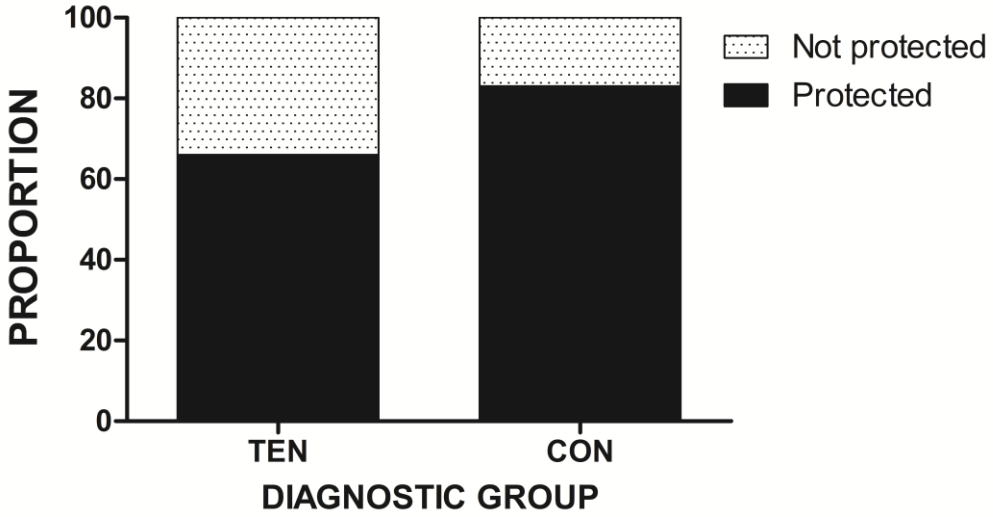

Figure 5.14: The proportion of TEN and CON participants classified as protected or not protected against AT by genetic risk assessment model 6

Protected: GRA model 6 protective genotype score ≥2; Not protected: GRA model 6 protective genotype score =0; GRA model 6: rs1045485, rs12722, rs4789932, rs1559186. TEN: n=124, CON: n=216

5.4. DISCUSSION

The primary finding of this study is that there are significant interactions between genes encoding components of the ECM and genes encoding components of cell-signalling pathways in modulating risk of AT. Variations in these genes may act in parallel through different pathways, or within the same pathway to modulate risk of AT. In addition, the analysis of risk models derived from both logistic regression and risk genotype scores shows that polymorphisms with a small effect (OR<2.0) contribute to overall genetic risk of AT in a cumulative manner.

Stepwise logistic regression analyses implicated interactions between genes encoding components of the ECM (*COL27A1*, *TNC*, *COL5A1* and *COL5A3*) and genes encoding components of the cell-signalling pathways (*IL-6*, *IL-1β* and *CASP8*) in a cumulative risk model for AT. September et al. (2011)[257] have previously proposed pathways in which the IL-6 and IL-1β proteins influence the expression of type V collagen and tenocyte apoptosis (Figure 5.15). Briefly, IL-1β up-regulates the expression of both itself and IL-6 as well as other inflammatory mediators which up-regulate the expression of MMP-1, MMP-3 and MMP-13.[257,281] These MMPs target type V collagen for degradation.[33] In addition, both IL-1β and IL-6 specifically affect the expression of the *COL5A1* gene via transforming growth factor β (TGF-β),[143,291] and TGF-β has also been shown to regulate the expression of *COL27A1*, *COL3A1* and *COL5A2*.[30,277,290] Interestingly, TGF-β has been shown to induce expression of a miRNA gene (*MIR455*) which, in humans, is located in intron 8 of the *COL27A1* gene, and miR455 in turn regulates TGF-β signalling.[277] In addition to its effects on TGF-β expression IL-6, together with caspase-8, has also been shown to play a role in tenocyte apoptosis.[196] Increased expression of both *IL-6* and *CASP8* is associated with tenocyte apoptosis and has been observed in tendinopathy.[196,205,257] The evidence above emphasises the intricate relationship between cell-signalling pathways and their effects on the ECM collagen network.

University of Cape Town

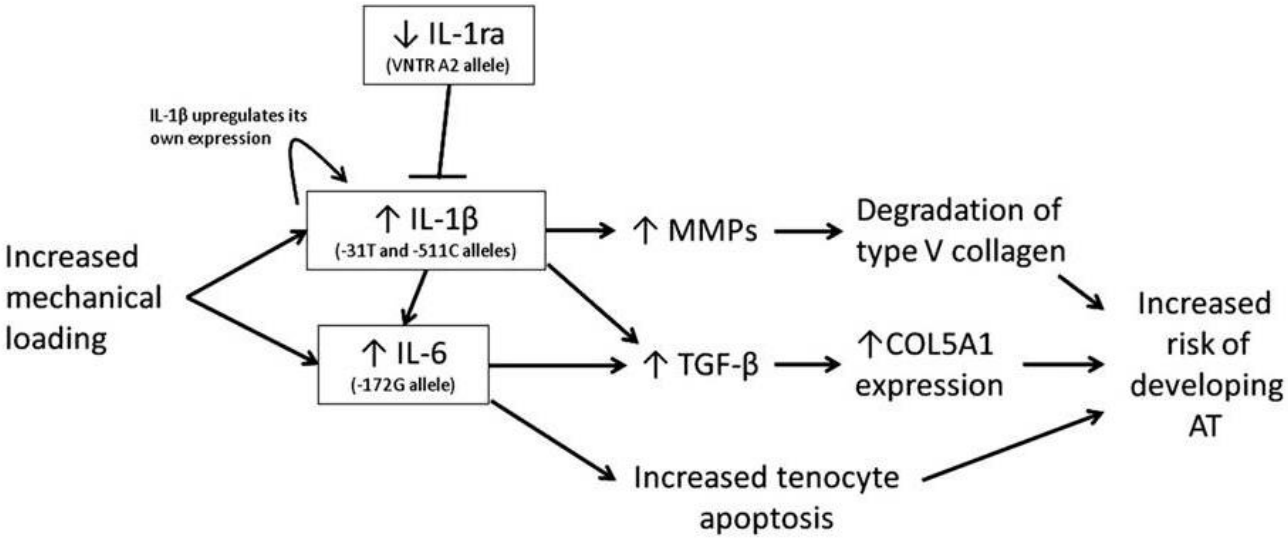

Figure 5.15: Proposed cell-signalling pathways which may lead to increased risk of AT
Proposed pathways through which increased mechanical loading induces expression of IL-1β, IL-1ra and IL-6, which in turn may act to influence the expression of type V collagen thereby potentially modulating the risk of AT.[33,55,143,166,279,291] *IL-1β increases MMP expression by increasing expression of inflammatory mediators such as COX-2 and PGE2.*[279,281] *IL-6 induces tenocyte apoptosis causing the production of reactive oxygen species and subsequent activation of caspases 8 and 3.*[196] *Some of these interactions have been observed in chondrocytes and are assumed to be similar in tenocytes. Up and down arrows refer to increased or decreased expression, respectively. Boxed molecules are those investigated in this study. The functional alleles are given in brackets and have previously been shown to affect the expression of the proteins as indicated by the arrows. IL-6, interleukin-6; IL-1β, interleukin-1β; IL-1ra, interleukin-1 receptor antagonist; MMP, matrix metalloproteinase; TGF- β, transforming growth factor β*

[Figure and caption reprinted from September et al. (2011)[257] *with permission from BMJ Publishing Group Ltd.]*

University of Cape Town

In addition, the gene-gene interactions observed between *TNC* and the *IL-6*, *IL-1β* and *CASP8* genes lend further support to recent evidence that implicates TN-C in the innate and adaptive immune responses.[90,194] In recent studies of inflammatory joint disease, TN-C was shown to be an endogenous activator of toll-like receptor (TLR) 4 which, together with TLR 2, is the predominant TLR in recognizing damage-associated molecular patterns (DAMPs).[90] DAMPs are generated as endogenous molecules are released from the ECM and cells in response to stress or injury and are recognised by pattern recognition receptors such as TLRs.[90] The recognition of DAMPs by TLRs is important during sterile immune responses such as seen in inflammatory conditions like rheumatoid arthritis, Alzheimers disease and

tumorigenesis.[90] TN-C in particular has been shown to have a key role in promoting the innate and adaptive immune responses in inflammatory joint disease.[194] It is required for chronic, destructive joint inflammation and stimulates the synthesis of tumor necrosis factor alpha (TNF-α), IL-6 and interleukin-8 in macrophages, and IL-6 in synovial fibroblasts.[194] In addition, the expression of *TNC*, together with *COL5A1*, *COL5A3* and *COL27A1*, is upregulated in tendinopathy[133] which, when considered with the gene-gene interactions noted in this study, further implicates these cellular pathways in the ECM response to load, adaptation and repair.

Considering the interactions observed in this study between the genes encoding components of the collagen fibril and genes involved in the cell-signalling pathways, as well as the biological relationship between these pathways, preliminary models for predicting genetic risk of AT incorporating these polymorphisms were investigated. When interpreting the ROC curve for the two risk models derived from stepwise logistic regression, a number of parameters are important. Sensitivity is the proportion of participants who have AT (TEN) who are predicted to be at risk of AT by the risk model, whereas specificity is the proportion of participants in the CON group who are predicted to not be at risk of AT.[83] The AUC quantifies the ability of the risk model to effectively discriminate between those participants with AT (TEN) and those participants without AT (CON), where an AUC of 1.00 represents the ideal model and an AUC of 0.50 represents a model that discriminates between TEN and CON no better than chance. Risk model A is, therefore, moderately able to discriminate between CON and TEN participants (AUC=0.775). Although the specificity of this model is low, the sensitivity is relatively high. In particular, 82.4% of TEN participants were predicted to be at risk of AT by this model. Conversely, risk model B has a higher specificity than sensitivity and is better able to discriminate CON and TEN participants (AUC=0.801). It is important to note that in this study, the optimal risk model was assumed where both sensitivity and specificity were optimised. It is however, possible to derive a risk model in which either sensitivity or specificity is optimised depending on the desired outcomes of the test. For example, using the ROC curve plotted in Figure 5.3, it is possible to derive a risk model in which sensitivity is optimised to approximately 98% with a corresponding specificity of 30%. This type of risk model would be able to predict risk of TEN in 98% of TEN participants, but would falsely identify more CON participants as having had a risk of TEN.

University of Cape Town

Contrary to this, deriving a risk model in which specificity is optimised to approximately 95% would decrease the sensitivity to approximately 35%. This more conservative model would only have assessed 35% of TEN participants to be at risk of AT, but would have assessed more CON participants to not be at risk of AT.

The strength of risk models derived from a statistical stepwise logistic regression lays in the relative weighting of risk attached to each factor or variable. Polymorphisms which have a larger biological effect on risk of AT will be more heavily weighted in their contribution to the risk model. The results obtained from this analysis are perhaps more applicable to a population or group, and are effective in highlighting the biologically significant factors that should be considered when evaluating disease risk for a group of individuals. However, in the current form, these risk models are not easy to interpret in a clinical setting. In this era of personalised medicine, a clinical tool that is pertinent to individuals may be of value. It was therefore decided to derive and test a risk model which may be more clinically appropriate and allow for easy interpretation at an individual level. Although the GRA models presented in this study may be more clinically applicable, they do not currently take into account the relative contribution of each polymorphism as all risk genotypes are weighted equally. An initial GRA model based on previously published genetic associations with AT observed in two populations was derived (GRA model 1). Participants with a risk genotype score of six or higher in this model, had 4.7-fold higher odds of having AT compared to those with a score of four or less. The *COL27A1* and *TNC* SNPs, rs946053 and rs2104772, were investigated in this thesis and found to be associated with AT in a haplotype and significantly contributed to risk of AT in the logistic regression models of risk. As they had a clearly distinguishable at-risk genotype (rs946053 GG; rs2104772 AA), the effect of their inclusion on the efficacy of GRA model 1 was tested. A GRA model was first derived using only these two SNPs (GRA model 2) and found that participants with a risk genotype score of four were twice as likely to have AT when compared to participants with a score of two or less. These two polymorphisms were therefore consecutively added to the initial risk model. When either *TNC* rs2104772 alone (GRA model 3) or both polymorphisms (GRA model 4) were included in the initial risk model, participants with a risk genotype score of eight or higher had approximately seven or six fold higher odds of AT respectively. This

University of Cape Town

suggests that the addition of the rs2104772 and rs946053 polymorphisms to the initial model may strengthen its ability to assess risk of AT in these participants.

Similarly, a GRA model derived using polymorphisms that have previously been associated with protection against AT was investigated (GRA model 5). Participants with a protective genotype score of two or more were half as likely to have AT than participants with a score of zero. The addition of the *COL5A3* rs1559186 marginally improved the model (GRA model 6), but the addition of *COL27A1* rs1249744 did not improve the model further (GRA model 7). Analysis of the proportion of CON and TEN participants predicted to be protected or not protected from AT by GRA model 6 showed that 83% of CON participants were predicted to be protected against AT. However, 66% of TEN participants were also predicted to be protected against AT. This model is, therefore, of limited value to the clinician as a risk assessment tool.

The identification of these preliminary risk models may support the inclusion of genetic risk testing in clinical care. However, it should be clearly noted that only 37% of TEN participants were predicted to be at risk of AT using GRA model 4. Conversely, 63% of TEN participants were predicted not to be at risk of AT using this model. This result emphasises that genetic susceptibility is only one contributing factor in the determination of overall susceptibility to AT. Using the current example, approximately 63% of the risk of AT can be attributed to polymorphisms other than those already identified ("missing" genes) and/or other risk factors. The more polymorphisms that are incorporated into risk models, the more likely they are to effectively assess risk.[205] In order to develop cost effective screening strategies, it is however important to determine the minimum number of these polymorphisms that can be used while still maintaining the efficacy of the model. As more polymorphisms which contribute to risk of AT are identified and the accuracy of genetic risk assessment models improves, genetic screening tests may become important tools in the clinical management of athletes and at-risk individuals. However, as genetic tests are neither diagnostic nor predictive, they should not be conducted in isolation. They would be of more value when considered in conjunction with other clinical information in athletes and individuals who present with other risk factors for AT including previous Achilles tendon injury, overweight, fluoroquinolone treatment and unfavourable biomechanics.

University of Cape Town

As only participants who were successfully genotyped for all polymorphisms included in each model could be included, small sample size is recognised as a limitation to this study. In addition, the applicability of these results to other racial groups and geographically defined populations may be limited.

5.5. CONCLUSION

The primary finding of this study is that there are significant genetic interactions between genes encoding components of the ECM and genes encoding components of the cell-signalling pathways in the modulation of AT risk. The biological mechanisms which may underlie these interactions are diverse and point to an intricate homeostatic relationship between cell-signalling pathways and the structural integrity of the ECM collagen network. This study also provides proof of concept that as more polymorphisms which influence risk of AT, and the biological mechanisms underpinning these associations, are elucidated, genetic screening tests for individuals presenting with other risk factors for AT may be implementable in a clinical setting.

University of Cape Town

CHAPTER 6: SUMMARY & PERSPECTIVES

Although there are clear clinical, histological and imaging diagnostic criteria for AT, the exact aetiology of this condition is still unclear. Several models have been proposed to explain the initiation and progression of this condition and a review of these models (Chapter 1) suggests that the underlying mechanism of pathology in AT is likely to be a combination of several mechanisms which are influenced by both intrinsic and extrinsic risk factors. It is therefore clear that Achilles tendinopathy is a multi-factorial condition. The most notable risk factors are a history of previous Achilles tendon injury, evidence of neovascularisation on examination with ultrasound, treatment with fluoroquinolone antibiotics, high running loads and heritable elements (reviewed in Chapter 1). In particular, there is strong evidence to support genetic variation as an independent risk factor for AT. Family history of AT has been identified as a significant risk factor for AT,[157] and several genes encoding structural and associated elements of the ECM have previously been associated with risk of AT.[6,76,164,197,199,205,224,231,255,257] Considering the number of these genes and the distinct pathways that they are components of, it is clear that the heritable component of AT is polygenic in nature and it is therefore unlikely that there is one causative gene variant for AT. It is more likely that many polymorphisms with small to moderate effects collectively contribute to an individual's genetic predisposition to develop AT. The identification of the full array of these variants is vital in elucidating the underlying mechanisms contributing to AT, as well as for the development of injury prevention practices.[82]

The primary aim of this thesis was therefore to identify additional genetic elements which predispose individuals to risk of AT and to propose the biological mechanisms underlying this genetic risk. In chapters 2 to 4 of this thesis, several genes encoding components of the ECM were identified as good candidate genes for association with AT based on their chromosomal location and biological function. A case-control, genetic association study approach was adopted to investigate the association of specific polymorphisms within these genes and AT. The selection of polymorphisms investigated was hypothesis driven and based on chromosomal location, biological function, heterozygosity, minor allele frequency and previous associations with other multifactorial phenotypes.

University of Cape Town

The *COL5A1* gene, encoding the α1(V) chain, is the most investigated gene in AT. A number of polymorphisms within the 3'-UTR of this gene have been independently associated with AT in two separate populations.[6,199,255] These polymorphisms have been suggested to alter the predicted secondary structure of the 3'-UTR and, subsequently, mRNA stability.[6] In particular, increased mRNA stability has been associated with Achilles tendinopathy when compared to the stability in asymptomatic controls.[164] This evidence has led to the hypothesis that the increased stability of the *COL5A1* mRNA results in an increased relative content of type V collagen in the collagen fibril, and leads to a collagen fibre that contains smaller, more densely packed collagen fibrils (Figure 6.1).[57,236] The subsequent changes in the mechanical properties of tendon tissue may lead to an increased susceptibility to injury.[57,236] It is therefore clear that the role of type V collagen in fibrillogenesis is strongly involved in the pathogenesis of tendinopathy. Type V collagen predominantly exists as a heterotrimer of two α1(V) and one α2(V) chains, but also exists as a homotrimer of three α3(V) chains. It is therefore reasonable to hypothesise that the *COL5A2* and *COL5A3* genes, which encode the α2(V) and α3(V) chains respectively, may play a role in the pathogenesis of tendinopathy. As with *COL5A1*, polymorphisms within the *COL5A2* gene have been associated with EDS and abnormal fibrillogenesis, and the *COL5A3* gene has also been suggested as a suitable candidate locus for investigation in EDS.[32,123,198,238] Despite this, the polymorphisms investigated in the *COL5A3* (rs2303099; rs1559186; rs2161468), *COL5A2* (rs13031549; rs4667264) and neighbouring *COL3A1* (rs2056156; rs3106796; rs1800255) genes were found not to be significantly associated with risk of AT in the SA and AUS participants. There was, however, a trend towards association with AT for the *COL5A3* rs1559186 polymorphism. In summary, the hypothesis for the selection of these genes as candidate genes in this thesis remains strong. It may therefore be of value to interrogate these genetic loci further with whole gene sequencing, using next generation sequencing technologies together with bioinformatics analyses to uncover potential functional variants involved in the pathogenesis of AT.

University of Cape Town

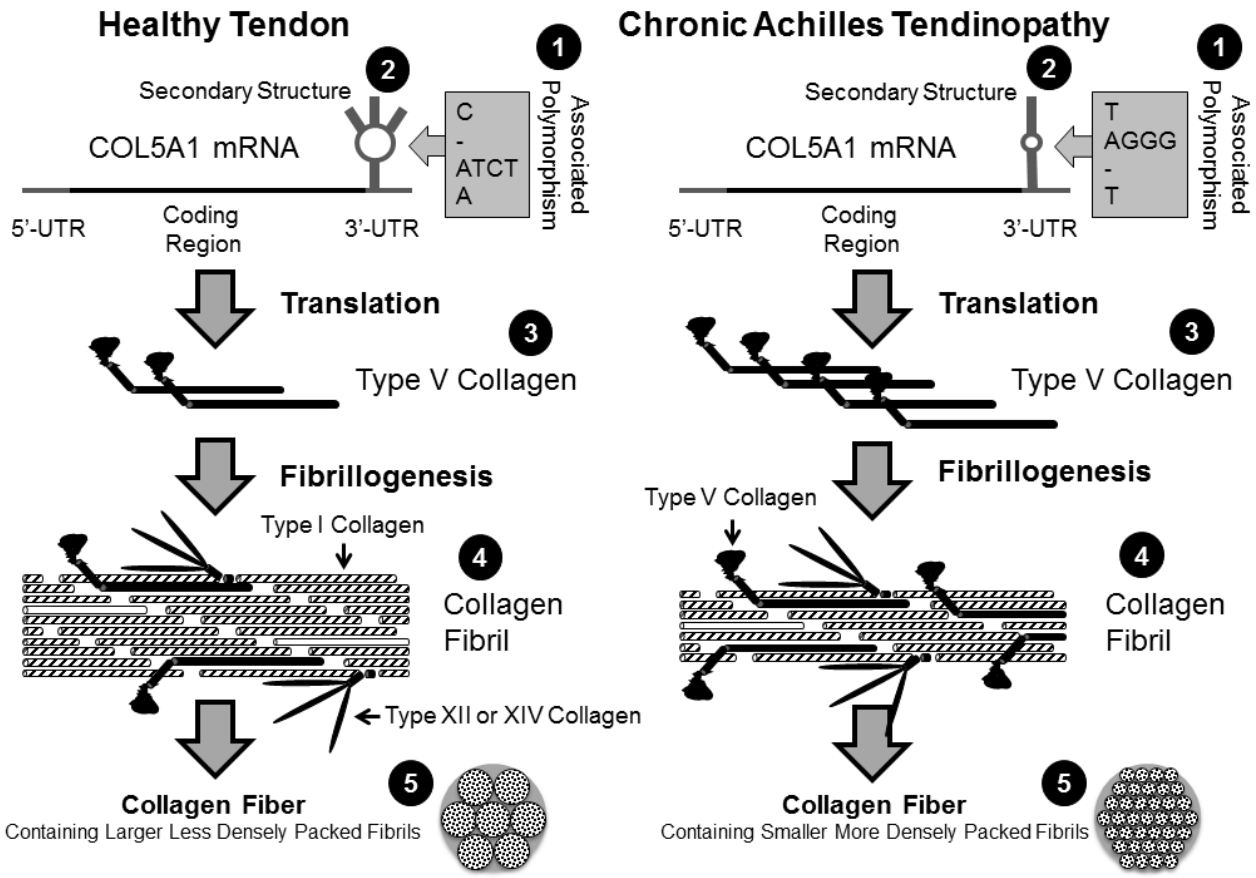

Figure 6.1: Ribbans and Collins'[236] hypothesis of the mechanism through which the *COL5A1* polymorphisms affect risk of AT

1) Four polymorphisms within the mRNA are independently associated with tendinopathy; 2) These polymorphisms appear to alter the predicted secondary structure of the 3'-UTR; 3) The secondary structure regulates mRNA stability and by implication type V collagen production; 4) Type V collagen regulates fibrillogenesis and thus the mechanical properties of tendons; 5) The healthy tendon contains larger less densely packed fibrils while the injured tendon contains smaller more densely packed fibrils.[57] [Figure and caption reprinted from RIbbans et al. (2013)[236]]

Similarly, the selection of the polymorphisms within the *COMP* and *THBS2* genes for investigation was also hypothesis driven. COMP (TSP-5) is a mechanosensitive, structural glycoprotein that plays a role in matrix assembly and tissue repair, and rs730079 has previously been associated with osteoarthritis.[14,108,223,243,272,284] TSP-2 is implicated in the healing response, MMP homeostasis, cell-matrix interactions as well as the regulation of collagen fibrillogenesis and fibril diameter, and both rs6422747 and rs9283850 have been associated with other multifactorial musculoskeletal conditions.[7,36,37,112,163] However, the polymorphisms investigated in the *COMP* (rs28494505; rs730079) and *THBS2* (rs9505888;

rs6422747; rs9283850) genes were also found not to associate with risk of AT in the AUS and SA participants investigated. Notwithstanding, whole gene sequencing of *COMP* and *THBS2*, as well as the *COL5A3*, *COL5A2* and *COL3A1* genes, should be considered as viable methods to conclusively test whether polymorphisms within these genes are associated with AT.

An additional novel finding of this thesis which addressed the primary aim was the association of the GCA haplotype, constructed from the *COL27A1* rs946053 (T>G), *TNC* rs13321 (G>C) and *TNC* rs2104772 (T>A) SNPs, with AT in the Australian and South African groups. This finding implicates the 759kbp region containing the 3'-end of the *COL27A1* gene and the 5'-end of the *TNC* gene in influencing risk of AT. Functional analysis of this region suggests that genetic variation within this haplotype may have functional effects on mRNA transcription, as well as the structure and properties of the TN-C and α1(XXVII) proteins. This potentially results in modified protein-protein and protein-environment interactions and, subsequently, compromised wound healing, tissue remodelling and adaptation (Chapter 3). Future research prospects include genome sequencing of this region in order to identify more functional variants and/or signature molecular motifs which may predispose to AT.

As briefly discussed in chapter 1, the ultimate objective of injury research at any level is the development and implementation of clinically relevant and practical injury prevention strategies. A six stage framework proposed by Finch (2006)[82] describes an approach for translating research into injury prevention practice (TRIPP). Briefly, stage 1 requires high-quality injury surveillance in order to establish the extent of the problem. Stage 2 requires a multidisciplinary research approach to establish the aetiology and mechanisms of injury as well as the factors associated with increased risk and severity of injury. The findings which fulfil the primary objective of this thesis contribute towards the increasing body of knowledge at this stage of the TRIPP framework. Findings from these genetic association studies also start to unveil the underlying molecular mechanisms in tendinopathy as seen in the *COL5A1* example.[6,164,199,255]

University of Cape Town

Stage 3 of the TRIPP framework is the identification and development of appropriate injury prevention strategies using a multidisciplinary approach.[82] Whilst AT is recognised as a multifactorial condition of which genetic variation is only one component contributing to overall risk, tools which gauge the extent of genetic risk of AT for individuals will be of value to sports clinicians. The secondary aim of this thesis was therefore to develop and test preliminary models assessing genetic risk of developing AT in order to start addressing stage 3 of the TRIPP framework. This secondary aim was addressed in chapter 5 of this thesis.

September et al. (2011)[257] reported the interaction of sequence variants within the *COL5A1*, *IL-6*, *IL-1β* and *IL-1RN* genes in the modulation of risk of AT using a pathway-based approach. Another study reported the independent and haplotype association of sequence variants within the *CASP8* gene, encoding a component of the apoptosis pathway, with AT.[205] It was therefore hypothesized that polymorphisms within genes encoding other components of the ECM may also interact with these cell-signalling pathways in the modulation of risk of AT. In chapter 5, two approaches were used to develop models which assessed genetic risk of AT. The first, stepwise logistic regression with ROC curve analysis, optimised both sensitivity and specificity of the model by sequentially adding factors to the models. Two separate, biologically plausible, risk models were developed which identified the polymorphisms contributing, and their relative contribution, to overall genetic risk of AT. The best fit risk model with the largest area under the curve included the variables age (years), sex (male), *COL27A1* rs946053 (GT,TT), *COL5A1* rs12722 (TC,CC), *COL5A3* rs1559186 (CG,GG), *IL-6* rs1800795 (GC,CC), *CASP8* rs1045485 (GC,CC) and *CASP8* rs3834129 (del/CTTACT,del/del), and had sensitivity and specificity indicators of 58.1% and 86.5% respectively. These risk models are valuable in that they indicate the relative contribution of each factor to the risk of AT assessed by that model, and highlight the biologically significant factors which should be considered when assessing risk. The risk models developed in chapter 5 emphasise the interactions between genes encoding components of the collagen fibril and genes involved in the cell-signalling pathways in modulating risk of AT. In addition, they emphasise the fact that the ability of risk models to effectively assess risk is increased when more biomarkers are included in the model.

University of Cape Town

The risk models described above are of value at a population level but are not easy to interpret in a clinical setting. Therefore, the underlying principles were used to develop a series of genetic risk assessment (GRA) models which may be more relevant in the provision of personal medical care. The GRA models included polymorphisms that (i) have previously been associated with AT in two independent populations and, (ii) have a definitive risk genotype that was implicated in a haplotype or gene-gene interactions associated with AT. A genetic risk assessment model which included risk genotype scores for polymorphisms rs16399 (*COL5A1*), rs4919510 (*MIR608*), rs143383 (*GDF5*), rs3834129 (*CASP8*), rs2104772 (*TNC*) and rs946053 (*COL27A1*) was developed and accurately assessed 90% of CON participants as not at-risk of AT but only accurately assessed 37% of TEN participants as at-risk of AT (OR 5.89, 95% CI:2.39-14.52; P<0.001). Although these preliminary GRA models are not fully implementable in a clinical setting, they provide proof of concept that genetic screening tests for AT are of value as a clinical tool.

Stages 4, 5 and 6 of the TRIPP framework centre around the evaluation of the injury prevention strategies under ideal conditions, understanding factors which influence the implementation of these measures and evaluating the effectiveness of the strategy in real-world conditions.[82] It is, therefore, worthwhile discussing some of the ethical and practical considerations which may influence the implementation and effectiveness of genetic screening tests.

It is important for clinicians and patients to understand the limitations of genetic screening tests, as well as issues of validity. Before polymorphisms that have been associated with risk of AT are included in genetic screening tests, the associations should have been repeated in at least one other independent population.[1,47] In addition, the association should be shown to be biologically plausible through investigation of the mechanisms underlying the association. For example, the initial association of the rs12722 SNP in *COL5A1* with AT in a South African population has been replicated in an independent Australian population and, together with other polymorphisms, has been mapped to a functional region of the 3'-UTR and suggested to play a role in the stability of *COL5A1* mRNA.[6,164,199,255]

University of Cape Town

It is particularly important to emphasise that predisposition does not equal predetermination. Polygenic profiling tests are not predictive and can only identify relative risk of disease/injury.[236,256] While some rare genetic mutations are causative and result in severe or lethal phenotypes without the involvement of environmental factors, milder non-lethal phenotypes are often a result of environmental exposure interacting with genetic risk factors (Figure 6.2).[236] Although there is an increasing number of sequence variants shown to be associated with AT, it is still not fully understood how these polymorphisms interact with environmental risk factors to either mask or modify risk of AT, nor the extent to which environmental factors modulate genetic risk through epigenetics.[59,236,256] This point is particularly pertinent when considering the ethical implications of genetic screening tests. It is imperative that these tests are interpreted by trained geneticists and that qualified genetic counsellors are involved in the process. As the identification of genetic risk factors increases and our understanding of personal genomics grows, so too does the potential for genetic discrimination. Genetic testing for sports performance and predisposition to concussion is already commercially available,[2,3] and has the potential to be considered in the selection and contracting of athletes to sports teams. In addition, there is potential for medical insurance companies to start considering genetic predisposition to disease/injury in a similar light to pre-existing conditions and therefore adjusting risk estimates and medical cover accordingly.[15]

Notwithstanding these limitations, genetic screening tests may ultimately allow for better prescription of exercise and preventative measures as well as treatment of AT. The feasibility of gene therapy has been shown in animal tendon studies and could be used to manipulate the healing environment of genetically "compromised" tendons.[110,312] Future challenges include identifying the most appropriate gene targets and refining the timing and mechanism of delivery to the affected tissue.[110] As is the case with many advances in medical treatment, the potential for abuse of these therapies is present. The World Anti-doping agency (WADA) prohibits "the transfer of polymers of nucleic acids or nucleic acid analogues" and "the use of normal or genetically modified cells" as methods of gene doping (http://list.wada-ama.org/list/m3-gene-doping/; accessed March 2013). As gene therapies

University of Cape Town

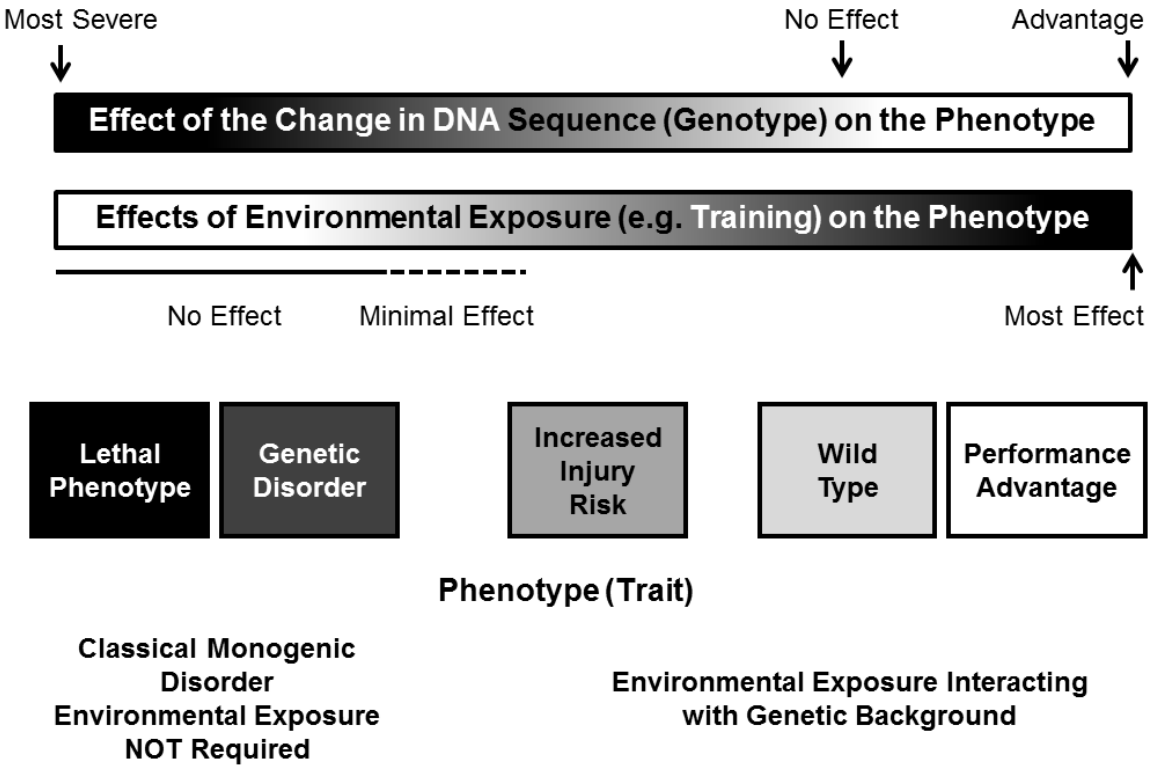

Figure 6.2: The gene-environment continuum described by Ribbans and Collins (2013)[236]
The heat bars indicate the effect of genotype and environmental exposure on the phenotype. Black shading represents the largest effect while the clear shading represents minimal or no effect. The black and dark grey boxes represent the most severe phenotypes and are caused by rare genetic mutations with minimal or no environmental exposure. Common orthopaedic injuries are caused by environmental exposure interacting with the individual's genetic background (middle box). [Figure and caption reprinted from RIbbans et al. (2013)[236]]

for the treatment of overuse injuries evolve, these therapies may be both legitimately used and abused by athletes partaking in high-load repetitive sports such as marathon running, rowing, canoeing and swimming to both improve recovery and increase resistance to injury during training. Currently, these prohibited methods are difficult to detect without baseline genetic profiling of athletes.

The studies included in this thesis have a number of limitations. Firstly, the sample size provides statistical power to detect associations of polymorphisms with AT that have an effect where OR>2. However, previous research in this area has shown that genetic variants in genes encoding structural components of the collagen fibril often have large effects on

risk of developing AT.[197,199,231] Future studies with larger cohorts are required to repeat these findings, as well as many of the previously published associations with AT, in independent populations. Secondly, the TEN participants were significantly heavier than the CON participants at recruitment. Although weight of TEN participants at time of injury could not be obtained, several TEN participants reported weight gain due to a decrease in physical activity after injury. In addition, the lack of physical activity and smoking data for the Australian participants is recognised as a limitation to this study. Lastly, due to the effects of population stratification and differences in genotype distribution, the applicability of these findings to other racial groups and geographically defined populations may be limited.

6.1. FUTURE DIRECTIONS

Future research in this field should exploit a multi-disciplinary approach to fully elucidate the role of genetic variation in Achilles tendinopathy. As mentioned previously in this chapter, it is likely that many polymorphisms with small to moderate effects collectively contribute to an individual's genetic predisposition to AT, and the identification of many more of these variants is vital in building a model to describe this genetic predisposition.[82] The studies presented in this thesis and those published to date have only been sufficiently powered to identify polymorphisms with a moderate effect due to the experimental sample size. As the field of tendinopathy genomics grows and collaborations lead to larger cohorts, it may become possible to explore a GWAS approach to define further genomic regions which are associated with AT.[186] Although GWAS are not always of value in detecting rare variants because of their exclusion in the currently available commercial arrays, they are likely to detect common variants with small effects, as well as copy number variations in newer GWAS chips.[186] Whilst GWAS in tendinopathy have value to add to the field, the value of candidate-gene association studies should not be underestimated as GWAS are based purely on statistical computation and do not take biological plausibility into account.[186] As larger study populations become available, it is also important to investigate the repeatability of reported associations with AT in independent populations, which will facilitate the identification of biologically significant pathways underlying tendon disease mechanisms.

University of Cape Town

In addition to undertaking a purely genetic approach, an integrated approach to investigating AT using techniques from molecular and cell biology, biochemistry, bioinformatics and genetics is necessary. A particularly interesting area of expansion for this field lies in the investigation of epigenetic modulation of gene expression profiles due to gene-environment interactions.[59,236,256] More specifically, alterations in the methylation profiles of CpG sites within genes encoding proteins of the extracellular matrix and cell-signalling pathways should be investigated.[59,242]

Currently, only the associated variants within the 3'-UTR of the *COL5A1* gene have been investigated at a functional level.[6,164] Future research should therefore focus on elucidating the molecular mechanisms underlying other reported genetic associations with AT. A suggested approach is the combination of next generation sequencing of genomic regions of interest to identify rare and/or functional variants which are likely to be associated with AT, with subsequent bioinformatic analysis to identify signature molecular elements within these regions. Furthermore, functional molecular research will help to elucidate and confirm the biological pathways and mechanisms which underpin the ECM response to load, adaptation and healing, as well as the role of genetic variation in this process.

Lastly, large prospective cohort studies are required to assess the ability of genetic screening tests to effectively and accurately assess risk of AT.

University of Cape Town

APPENDICES

(A) ETHICAL APPROVAL & RECRUITMENT FORMS

1. Approval letters from human research ethics committee
2. Recruitment information sheet
3. Informed consent forms
4. Participant questionnaire
5. Diagnostic criteria for Achilles tendinopathy (TEN) participants

(B) ALLELIC DISCRIMINATION METHODOLOGY

1. PCR conditions for *COMP* and *THBS2* polymorphisms investigated in chapter two
2. PCR conditions for *COL27A1* and *TNC* polymorphisms investigated in chapter three
3. PCR conditions for *COL5A3*, *COL3A1* and *COL5A2* polymorphisms investigated in chapter four

(C) SUPPLEMENTARY RESULTS

Table C.1: Allele combination frequency distributions in CON and TEN participants for significant interactions in the modulation of risk of AT

University of Cape Town

(A) ETHICAL APPROVAL & RECRUITMENT FORMS

1. APPROVAL LETTERS FROM HUMAN RESEARCH ETHICS COMMITTEE

UNIVERSITY OF CAPE TOWN

Research Ethics Committee
E53 Room 44.1, Old Main Building
Groote Schuur Hospital, Observatory, 7925
Queries : Xolile Fulo
Tel : (021) 406-6492 Fax: 406-6411
E-mail : Xfulo@cumie.uct.ac.za

15 September 2004

REC REF: 289/2004

Dr M Collins
Human Biology
Exercise Science & Sports Medicine

Dear Dr Collins

THE GENETIC BASIS OF ACHILLES TENDON PATHOLOGY IN AN AUSTRALIAN POPULATION

Thank you for your letter to the Research Ethics Committee dated 13 September 2004.

It is a pleasure to inform you that the Research Ethics Committee has approved the amended English Patient Information Leaflet and Informed Consent Sheet dated 13 September 2004.

Please quote the REC. REF in all your correspondence

Yours sincerely

PROF. T. ZABOW
CHAIRPERSON

UNIVERSITY OF CAPE TOWN

Research Ethics Committee
E53 Room 44.1, Old Main Building
Groote Schuur Hospital, Observatory, 7925
Queries : Xolile Fulo
Tel : (021) 406-6492 Fax: 406-6411
E-mail : Xfulo@curie.uct.ac.za

21 February 2005

REC REF: 086/2005

Dr M Collins
Human Biology

Dear Dr Collins

THE GENETIC BASIS OF TENDINOPATHY

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 10th February 2005

The contents are noted and added to our files.

Please quote the REC. REF in all your correspondence

Yours sincerely

PROF T. ZABOW
CHAIRPERSON

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: nosi.tywabi@uct.ac.za

13 August 2009

REC REF: 086/2005 & 289/2004

A/Prof M Collins
Department of Human Biology
Sports Science Institute

Dear A/Prof Collins

PROTOCOL TITLE: ADDENDUM TO OUR RESEARCH STUDIES ON THE "THE GENETIC BASIS OF TENDINOPATHY" AND "THE GENETIC BASIS OF TENDON PATHOLOGY IN AN AUSTRALIAN POPULATION".

Thank you for your letter to the Research Ethics Committee dated 05th August 2009.

It is a pleasure to inform you that the Ethics Committee has **noted and approved** the amendment to the above mentioned study.

Please note that this will be the last amendment for 289/2004 as it is 5 years old. Please can we get simple updated reports for both studies.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON, HSF HUMAN ETHICS

2. RECRUITMENT INFORMATION SHEET

Department of Human Biology
UCT/MRC Research Unit for Exercise Science & Sports Medicine
Faculty of Health Sciences, University of Cape Town
Private Bag, Rondebosch 7700, South Africa
Tel: + 27 21 650 1661
Fax: + 27 21 686 7530

THE GENETIC BASIS OF EXERCISE-INDUCED CHRONIC TENDON PATHOLOGY

Although there is a high incidence of tendon overuse 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 tendon injuries. In an attempt to determine whether there is a genetic basis for tendon pathology, we are interested in studying whether certain genes are associated with chronic tendinopathies. 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 and the Division of Human Genetics within the Department of Clinical Laboratory Sciences at 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 at least 30 minutes, you will be asked to donate 5 ml (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. At a later stage, some participants will be asked to visit a doctor (radiologist) for a tendon scan at no cost to themselves.

All the information received 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 tendon pathology. You are also free to request that your DNA sample be destroyed before the completion of the study.

If you are part of the tendon 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 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. Malcolm Collins, PhD
(021) 650 4574
mcollins@sports.uct.ac.za

Prof. Martin Schwellnus, MBChB, MD
(021) 650 4576
mschwell@sports.uct.ac.za

Colleen Saunders, MSc student
(021) 650 1569
csanders@sports.uct.ac.za

The University of Cape Town is committed to policies of equal opportunity and affirmative action which are essential to its mission of promoting critical inquiry and scholarship.

Department of Human Biology
University of Cape Town
South Africa

16 July 2004

STUDY ON THE GENETIC BASIS OF TENDON PATHOLOGY

The Musculoskeletal Research Centre at La Trobe University in Melbourne, Australia, in collaboration with the UCT/MRC Research Unit for Exercise Science and Sports Medicine at the University of Cape Town (UCT) in Cape Town, South Africa, are currently studying the genetic basis of chronic tendon pathology in the Australian population.

Studies have suggested that some individuals have a genetic predisposition to tendon injury and that genes (those traits which you inherit from your parents), such as Type V collagen (COL5A1) and Tenascin C (TNC), which encode for important components of tendons are associated with tendon pathology.

The aim of this study is to determine whether the TNC, the COL5A1 or other similar genes are associated with chronic tendon pathology.

In order to participate in this study, you will be required to donate five mill litres of venous blood after giving written consent. The blood sample will be used for the extraction and analysis of genetic material (DNA). The extracted DNA will be sent to UCT in South Africa for analysis. The samples will be shipped to and analysed by UCT anonymously. The DNA will be genotyped (analysed) for variations (polymorphisms) within the *COL5A1*, *TNC* and other candidate genes believed to be associated with tendon pathology. You will be required to complete a number of questionnaires regarding personal particulars, sporting participation and medical history, as well as a stretching and warm up questionnaire. In addition, you will be required to visit a doctor (radiologist) at a later stage for a tendon(s) scan free of charge. All the information collected during the study will be treated with the strictest confidentiality and will only be used for scientific research purposes. Your name and personal particulars will not be released under any circumstances and all the data obtained will be analysed anonymously.

If you would like to participate in the study and/or obtain any additional information, please contact **Dr Jill Cook**, on phone: **9479 5789**, or e-mail: **J.Cook@latrobe.edu.au.**

3. INFORMED CONSENT FORMS

Department of Human Biology
UCT/MRC Research Unit for Exercise Science & Sports Medicine
Faculty of Health Sciences, University of Cape Town
Private Bag, Rondebosch 7700, South Africa
Tel: + 27 21 650 4561
Fax: + 27 21 686 7530

GENETIC BASIS OF EXERCISE-INDUCED CHRONIC TENDON PATHOLOGY

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 and the Division of Human Genetics within the Department of Clinical Laboratory Sciences at the University of Cape Town's study on the genetic basis of exercise induced chronic tendon pathology. I have agreed to donate five millilitres of venous blood 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 not be released under any circumstances and that all data will be analysed anonymously.

If requested, I am also prepared to visit a doctor (radiologist) at a later stage for a tendon scan at no cost to myself (please delete this sentence if not applicable). If requested, I am also prepared to visit the SSISA for measurements to determine musculo-tendinous stiffness.

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 tendon pathology. I also understand that I will be free to request that my DNA sample be destroyed before the completion of the study.

FULL NAME OF SUBJECT:

SUBJECT'S SIGNATURE:

DATE:

INVESTIGATOR: _____

INVESTIGATOR'S SIGNATURE

The University of Cape Town is committed to policies of equal opportunity and affirmative action which are essential to its mission of promoting critical inquiry and scholarship

13TH JULY 2004

LA TROBE UNIVERSITY

GENETIC BASIS OF TENDON PATHOLOGY

INFORMED CONSENT

I, (the participant), have been fully informed about this study on the genetic basis of tendon pathology to be conducted by the Musculoskeletal Research Centre at La Trobe University in Melbourne, Australia, in conjunction with the UCT/MRC Research Unit for Exercise Science and Sports Medicine at the University of Cape Town in Cape Town, South Africa.

I have agreed to donate five millilitres of venous blood, which will be used for the extraction and analysis of genetic material (DNA), and will be taken by a registered medical physician or nurse. The DNA will only be used for scientific research purposes relating to the genetic basis of tendon pathology only. 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 all data will be analysed anonymously and my DNA sample will be destroyed on completion of the study.

I understand that the DNA extracted form the donated blood sample will be sent to UCT in South Africa for analysis. I understand that the DNA samples will be shipped to and analysed by UCT anonymously. I understand that the DNA will be genotyped (analysed) for variations (polymorphisms) within the COL5A1 and TNC genes, as well as other similar genes relating to the genetic basis of tendon pathology only.

In addition, I am also prepared to visit a doctor (radiologist) at a later stage for a tendon(s) scan at no cost to myself.

I understand that whilst there is no direct benefit to myself, if a genetic predisposition for tendinopathy 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 have read (or, where appropriate, have had read to me) and understood 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 anytime

and, further, to demand that data arising from my participation is not used in the research project provided that this right is exercised within four weeks of the completion of my participation in the project. 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 neither my name nor any other identifying information is used.

Any questions regarding this project may be directed to the Chief Investigator: **Dr Jill Cook**, of the **School of Physiotherapy** on telephone number **9479 5789**.

If you have any complaints or queries that the investigator has not been able to answer to your satisfaction, you may contact the Ethics Liaison Officer, Human Ethics Committee, La Trobe University, Victoria, 3086, (ph: 03 9479 1443, e-mail: humanethics@latrobe.edu.au).

Name of Participant: _____

Signature: _____ **Date:** _____

Name of Researcher: _____

Signature: _____ **Date:** _____

4. **PARTICIPANT QUESTIONNAIRE**

Department of Human Biology
UCT/MRC RESEARCH UNIT FOR EXERCISE SCIENCE & SPORTS MEDICINE
Faculty of Health Sciences, University of Cape Town
Private Bag, Rondebosch 7700, South Africa
Tel: + 27 21 650 4561
Fax: + 27 21 686 7530

GENETIC BASIS OF TENDON 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)			
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		Nationality	
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 _____		
Do you know your blood group?	Yes ☐ A ☐ B ☐ AB ☐ O ☐ No ☐ Rh Pos ☐ Rh Neg ☐		

The University of Cape Town is committed to policies of equal opportunity and affirmative action which are essential to its mission of promoting critical inquiry and scholarship

(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 Current or past participation	Sport 1 Current ☐ Past ☐	Sport 2 Current ☐ Past ☐	Sport 3 Current ☐ Past ☐
Year started participation			
Years involved in the sport Current hours of training per week (1-12 months)			
Current hours of training per week (13-24 months)			
Hours of training per week prior to first injury (1-12 months)			
Hours of training per week prior to first injury (13-24 months)			

Type of sport(s) you have participated in Current or past participation	Sport 4 Current ☐ Past ☐	Sport 5 Current ☐ Past ☐	Sport 6 Current ☐ Past ☐
Year started participation			
Years involved in the sport Current hours of training per week (1-12 months)			
Current hours of training per week (13-24 months)			
Hours of training per week prior to first injury (1-12 months)			
Hours of training per week prior to first injury (13-24 months)			

C. GENERAL MEDICAL DETAILS		If Yes, How long ago or how often
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
Do you suffer from any Connective Tissue and Rheumatological Diseases and Disorders?	Yes ☐ No ☐	If Yes, please select from the list below

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

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

Have any other members of your family suffered from any tendon pathology?	Yes ☐ No ☐	If Yes, please specify the family member (eg Mother, Son) and type of injury Acute Injury ☐ Chronic Pain and Swelling ☐ Other ☐
Do you suffer from elevated blood cholesterol?	Yes ☐ No ☐	Do any other members of your family suffer from elevated blood cholesterol? Yes ☐ No ☐
Have you been diagnosed with any of the following systemic diseases?	☐ No systemic disease ☐ Diabetes mellitus ☐ Adrenal disorders ☐ Thyroid disorders ☐ Amyloidosis ☐ Renal disease ☐ Other endocrine and metabolic disease (Specify)	

D. TENDON INJURY - MEDICAL DETAILS

Symptoms

How many times have you had tendon injuries?	Tendon Injured	Date of Injury	Acute or Chronic Injury	Sudden[1] or Gradual[2] Onset
[1]Sudden onset is within a few seconds or minutes	1			
	2			
[2]Gradual onset is over days or weeks	3			
	4			
	5			

Please complete a **separate form**, Part D only, for each Tendon Injury you have had

Injury Number (1,2,3,4,or 5)	☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ _____
Which tendon did you injure?	Rotator cuff tendon ☐ • Supraspinatus ☐ • Infraspinatus ☐ • teres minor ☐ Patellar tendon ☐ Wrist extensor tendons ☐ Achilles tendon ☐
Which side was injured?	☐ Left ☐ Right ☐ Both
Which region of your tendon was injured? Please indicate on a diagram. (Only if applicable)	☐ Upper 1/3 ☐ Middle 1/3 ☐ Lower 1/3
To what extent was your Tendon ruptured?	☐ Complete ☐ Partial ☐ None
How were you injured? (e.g. sport, walking)	
Grade of injury **at the time** of injury	☐ pain only after exercise ☐ pain during exercise, but did not cause you to alter training ☐ pain during exercise, which causes you to alter training ☐ pain which causes you to stop training ☐ no pain ☐ not sure ☐ Other (Specify)
Grade of injury **currently**	☐ pain only after exercise ☐ pain during exercise, but did not cause you to alter training. ☐ pain during exercise, which causes you to alter training ☐ pain which causes you to stop training ☐ no pain ☐ not sure ☐ Other (Specify _____)

Which of the following symptoms were present **before** the injury	☐ Pain (less than 1 week) ☐ Pain (1-4 weeks) ☐ Pain (> 4 weeks)	☐ Stiffness ☐ Swelling ☐ None
Which of the following symptoms were present **after** the injury If you have or had chronic tendon pain, what seems to alleviate the pain?	☐ Pain (less than 1 week) ☐ Pain (1-4 weeks) ☐ Pain (> 4 weeks)	☐ Stiffness ☐ Swelling ☐ None

Diagnosis

Which type of Tendon Disease were you diagnosed with e.g. Rupture, Tendinitis, etc.		
Diagnosed by (Please indicate the name and contact number of the clinician who diagnosed you)	☐ Doctor ☐ Physiotherapist ☐ Biokineticist ☐ Podiatrist ☐ Other	_____
If you had a tendon rupture. How was it treated?	☐ Surgically	☐ Non-surgically
If applicable, who was the surgeon?	Surgeon	Phone
If applicable, what diagnostic imaging was performed?	☐ Ultrasound ☐ MRI ☐ CT	Other _____
If applicable, who did the imaging?	Clinician	Phone

General Information		**If Yes, How long ago or how often**	
Have you ever used oral corticosteroids **prior** to your symptoms of tendon pathology (cortisone tablets)?	Yes ☐ No ☐	☐ 3 months ☐ 12 months	☐ 6 months ☐ 24 or more months
Have you ever been given an injection with corticosteroids **prior** to your symptoms of tendon pathology?	Yes ☐ No ☐	☐ 3 months ☐ 12 months	☐ 6 months ☐ 24 or more months
Have you ever been given an injection of corticosteroids in or around the injured tendon **prior** to your symptoms of tendon pathology?	Yes ☐ No ☐	☐ Once ☐ 3 times	☐ Twice ☐ >3 times
Have you ever used anabolic steroids **prior** to your symptoms of tendon pathology?	Yes ☐ No ☐	☐ 3 months ☐ 12 months	☐ 6 months ☐ 24 or more months
Have you ever used fluoroquinoline antibiotics **prior** to your symptoms of tendon pathology?	Yes ☐ No ☐	☐ 3 months ☐ 12 months	☐ 6 months ☐ 24 or more months

E. STRETCHING AND WARM UP (PRIOR TO YOUR INJURY)		
I usually stretch each week as follows: (Please tick ALL the appropriate boxes)		Never Occasionally ___ Before sport ___ After sport ___ Once daily ___ Twice daily More than twice daily
Which of these muscle groups do you stretch?	Lower Back	Always ☐ Occasionally ☐
	Buttock	___ Always ☐ Occasionally
	Hip Flexors	___ Always ☐ Occasionally
	Quads	Always ☐ Occasionally
	Hamstrings	Always ☐ Occasionally
	Calf Muscles	Always ☐ Occasionally
How many times do you stretch per week?		Never ___ < 5 min ___ 5 min 10 min 15 min 20 min 25 min > 30 min
Do you warm up before exercise?		Yes ☐ No ☐
If yes, for how many minutes and how?		
Do you cool down after exercise?		Yes ☐ No ☐
If yes, for how many minutes and how?		

5. DIAGNOSTIC CRITERIA FOR ACHILLES TENDINOPATHY (TEN) PARTICIPANTS

Department of Human Biology
UCT/MRC Research Unit for Exercise Science & Sports Medicine
Faculty of Health Sciences, University of Cape Town
Private Bag, Rondebosch 7700, South Africa
Tel: + 27 21 650 4561
Fax: + 27 21 686 7530

CLINICAL DIAGNOSIS OF ACHILLES TENDINOPATHY

SUBJECT NUMBER/CODE: _____

Clinical criteria [1,2]	Present
Gradual progressive pain over the posterior lower leg - Achilles tendon area (> 6 weeks)	
Early morning pain	
Early morning stiffness	
History of swelling over the Achilles tendon area	
Tenderness to palpation over the Achilles tendon	
Palpable nodular thickening over the affected Achilles	
Positive "shift" test (movement of the nodular area with planter- /dorsi-flexion)	

Other criteria	Present
Confirmation of the diagnosis by ultrasound *	
Confirmation of the diagnosis by MRI *	
Confirmation of the diagnosis by CT scan *	

*. One of these criteria must be present to confirm the diagnosis

Date: / /

Investigator: Prof M Schwellnus

Signature: _____

References:

1. Schapsis AA, Jones IE, Haas AL. Achilles tendon disorders in athletes. *Am.J Sports Med* 2002;30:287-305.
2. Kader D, Saxena A, Movin T, Maffulli N. Achilles tendinopathy: some aspects of basic science and clinical management. *Br.J Sports Med* 2002;36:239-49

The University of Cape Town is committed to policies of equal opportunity and affirmative action which are essential to the mission of promoting critical inquiry and scholarship

(B) ALLELIC DISCRIMINATION METHODOLOGY

1. PCR CONDITIONS FOR *COMP* & *THBS2* POLYMORPHISMS INVESTIGATED IN CHAPTER TWO

dbSNP:	**rs9283850**
CHROMOSOMAL POSITION:	6:169637456
BASE PAIR CHANGE:	R; A>G
MINOR ALLELE FREQUENCY:	44.8%
HETEROZYGOSITY:	0.346

FASTA SEQUENCE:

CGTCCTGCCGGACTAACACAAGAAGC**R**GAGAGAGATCAGGCTGTGCCGCCTA

Bolded base pair indicates SNP location

PCR CONDITIONS: TaqMan® SNP Genotyping Assays (Applied Biosystems™)
Assay ID C____180588_10

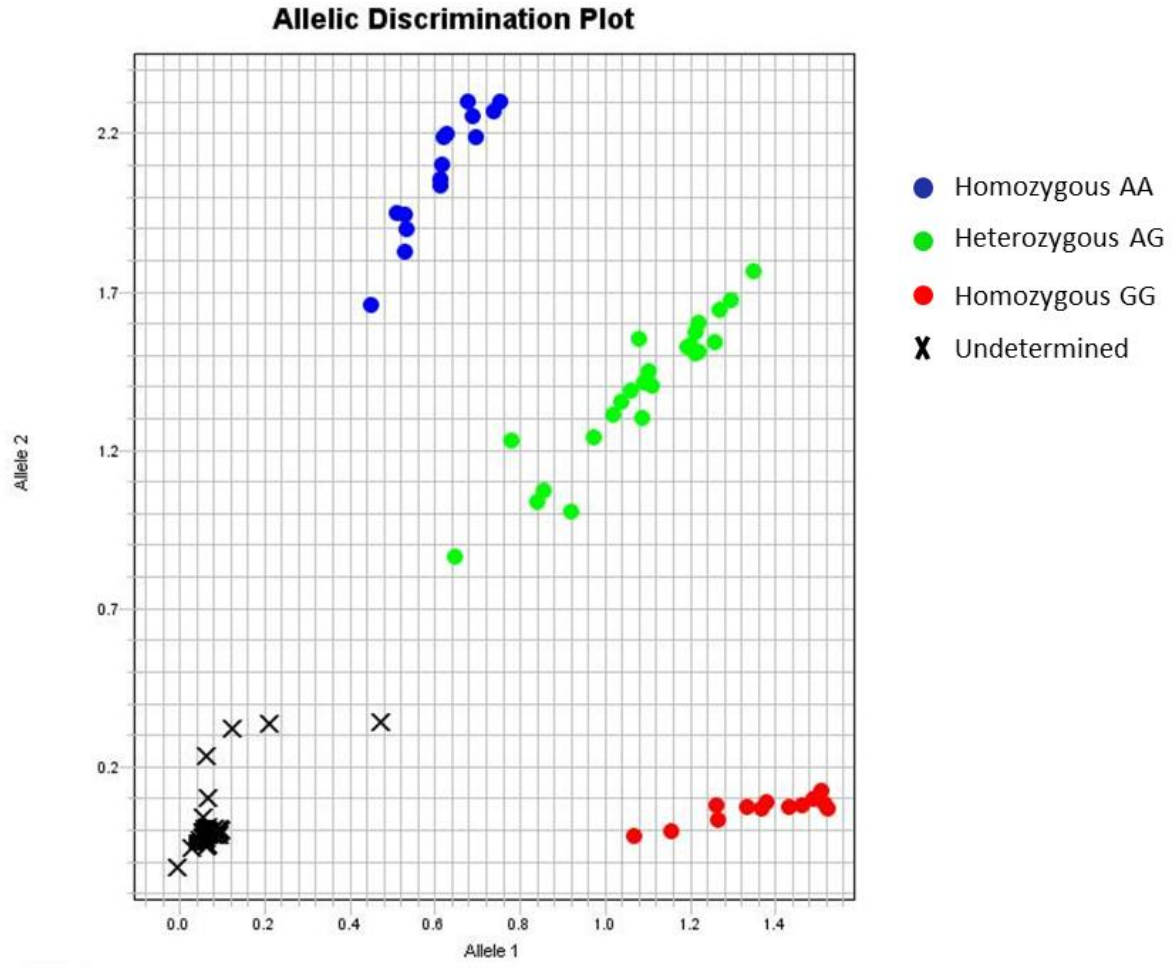

Figure B.1: A typical allelic discrimination plot using the TaqMan® SNP Genotyping Assay for rs9283850 on the StepOnePlus™ Real-Time PCR System

dbSNP:	**rs6422747**
CHROMOSOMAL POSITION:	6:169629783
BASE PAIR CHANGE:	R; A>G
MINOR ALLELE FREQUENCY:	48.1%
HETEROZYGOSITY:	0.445

FASTA SEQUENCE:

CACTGGCTCCTCAGCAGAAGGAGAGGAGGTCTTCTATTTTTTTGCAGTTAAAATAGAAAAAATTATTCAAAACATTTCCGGC
AGCACACCCTGGGCAGCCCCATGCACTGAGGTAAAATCCGCTTTTGCAAAGGACTGTCTTTATAGTGACCTTTGGGAACTGC
CTATGCACAAACGCTTGCAGGACAATGGCAGTTACGTTGACAAGGGCTGTTTGGTATAAAGTGGCCTCCTGTGGACCTCGCT
GCTCCTGGAACCCAGGGCAGGTGCGCGGCACAAGCTGCTCCTACCTTCTCATCGGTCACACCGTCATTGTCATCGTCATCATC
ACAGGCATCGCCAATCCCGTCCTTGTCAAAGTCTTCCTGCCCAGAATTTGGCAGATGGGGGCAGTTATCCTGCAATTRGAGA
AGGAAGAATACTTGGAAAAACATTTAGGAAAGCGCCTCTTTCAGGAATGCAGGTGAAATGAAACCAGCATGGCGCCGAGGA
AGGAAGCGTGCCCCATCAGTCCTCGATCTCAAGCG

Underlined sequences indicate oligonucleotide primer sequences; Bolded base pairs indicate SNP location; Highlighted sequences indicate nuclease cutting sites.

PCR CONDITIONS:	
Polymerase	Super Therm DNA Polymerase
Cycling	Standard PCR for 35 cycles
Annealing temperature	55˚C
Mg^{2+} concentration	2.0mM
Oligonucleotides	5'-CGCTTGAGATCGAGGACTG-3' 5'-CACTGGCTCCTCAGCAGAAG-3'
Amplicon size	527bp
RESTRICTION CONDITIONS:	
Nuclease	*MfeI*
Cutting site	C'AATT_G
Reaction Buffer	New England Biolabs® Inc. Buffer 4
Incubation temperature	37˚C
Fragment sizes	A allele - 527bp G allele - 405 and 122bp

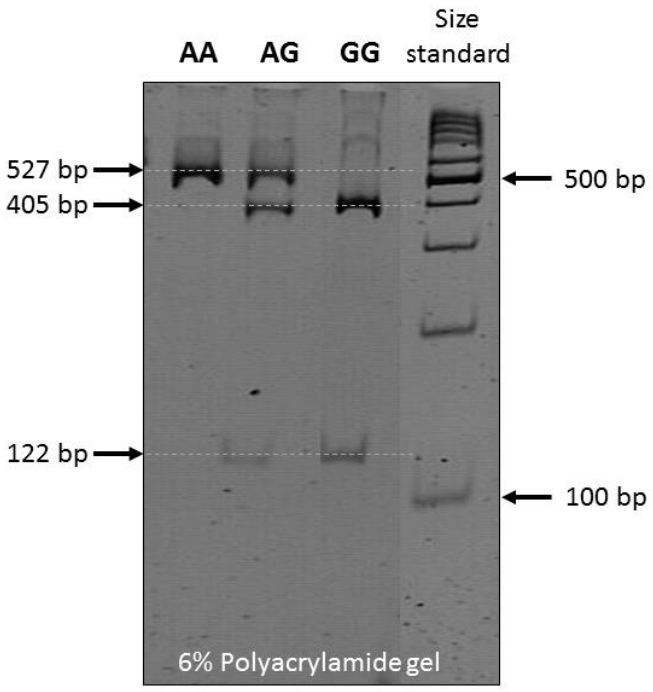

Figure B.2: A typical 6% PAGE image discriminating restriction fragments from individuals who are homozygous AA (lane 1), heterozygous AG (lane 2) and homozygous GG (lane 3) for the rs6422747 polymorphism

Fragment sizes are indicated with arrows on the sides of the image with a 100bp size standard in lane 4. Samples were electrophoresed at 120V for 2 hours.

University of Cape Town

dbSNP ID: **rs9505888**

CHROMOSOMAL POSITION: 6:169625874
BASE PAIR CHANGE: R; A>G
MINOR ALLELE FREQUENCY: 42.3%
HETEROZYGOSITY: 0.487

FASTA SEQUENCE:

AAGGAAGATGTGCTTCTTCGCTGGAAAAGCATGAACTACCAGGTAATTATGTCTCTTTTTTACACCATTTGCGCCATTTACAC
CATTTACATCATAGGATAATATATCCTATGAGCTGCATCAAGATAACTGTCCTCCCTTGTCTCTTTCAAACGCRAACCTGGAA
ATGAAGCCAATTCCTATGTTGGCATCAGGAAAATGGGGAGAGAATTCTGTGAAGACACAACCTCGCTCCAGCTGATGAGGTG
AAGGCGTCTATTCATTTGGCCTGTGGCCACTGAGATACAAGAATAGTGCTGAGATTTGAGTTGAATAAATGTTTTCAATTGAA
ATCATTGAAAACTCCATGTTATCCTGTCTTTGTGGAGCTCATGGAAAACT

Underlined sequences indicate oligonucleotide primer sequences; Bolded base pairs indicate SNP location; Highlighted sequences indicate nuclease cutting sites.

PCR CONDITIONS:	
Polymerase	Super Therm DNA Polymerase
Cycling	Standard PCR for 35 cycles
Annealing temperature	55˚C
Mg^{2+} concentration	2.0Mm
Oligonucleotides	5'-AAGGAAGATGTGCTTCTTCGCTGGAAAA-3' 5'-AGTTTTCCATGAGCTCCACAAAGACAGG-3'
Amplicon size	381bp
RESTRICTION CONDITIONS:	
Nuclease	*BstUI*
Cutting site	CG'CG
Reaction Buffer	Fermentas Buffer R
BSA concentration	0%
Incubation temperature	37˚C
Fragment sizes	A allele - 381bp G allele - 226 and 155bp

iversity of Cape Town

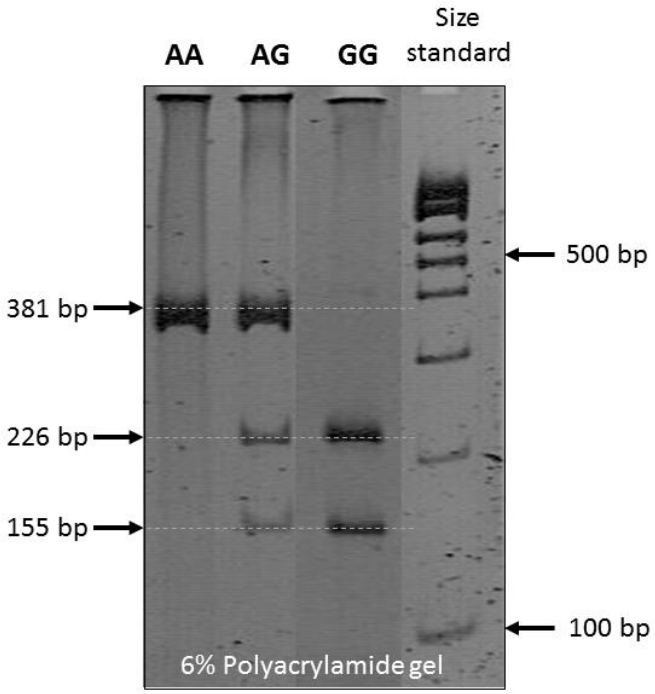

Figure B.3: A typical 6% PAGE image discriminating restriction fragments from individuals who are homozygous AA (lane 1), heterozygous AG (lane 2) and homozygous GG (lane 3) for the rs9505888 polymorphism

Fragment sizes are indicated with arrows on the sides of the image with a 100bp size standard in lane 4. Samples were electrophoresed at 120V for 2 hours.

dbSNP ID: **rs730079**

CHROMOSOMAL POSITION: 19:18903495
BASE PAIR CHANGE: S; C>G
MINOR ALLELE FREQUENCY: 27.5%
HETEROZYGOSITY: 0.417

FASTA SEQUENCE:

GAGTTACAAAACCCCTTAACCTCCTG**S**AATCAGTTAACACTCTCCACTCCCA

Bolded base pair indicates SNP location

PCR CONDITIONS: TaqMan® SNP Genotyping Assays (Applied Biosystems™)
Assay ID C___2257395_10

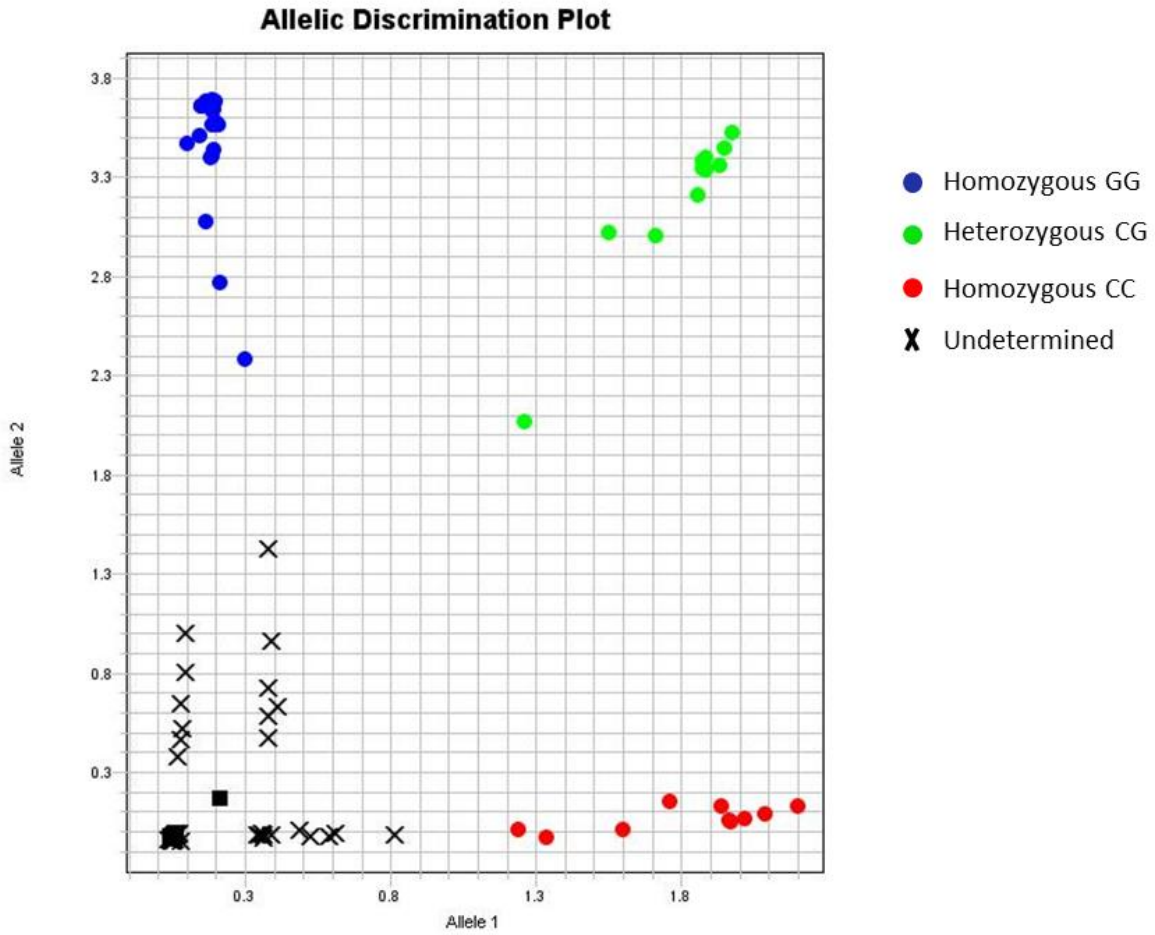

Figure B.4: A typical allelic discrimination plot using the TaqMan® SNP Genotyping Assay for rs730079 on the StepOnePlus™ Real-Time PCR System

dbSNP:	**rs28494505**
CHROMOSOMAL POSITION:	19:18893811
BASE PAIR CHANGE:	R; A>G
MINOR ALLELE FREQUENCY:	15.0%
HETEROZYGOSITY:	Undetermined

FASTA SEQUENCE:

CTACTGTCTCCATGCAGC//CTCTGAGCCCTTCTCACTTCCCCCTCAGGACGGCCACCCCTTGGGGCTGGGTGCAGAGCCCCCA
TCCAGCCGCGGTGAGGGTGGCTGTCATCCGGCGGGTCCTCACCCTGGTCCCTAGGCTTGCCGCAGCTGATGGGTCTCATAGTC
CTCTGGGATGGTGTCTGCAGGGAGAGGGCAGGCGGGTGAGGGCTGAGAAGGCC**R**GCAGGGCCTGTGTGCAGACTCCCCGCC
CACGGCCCGCTGGCCCTCGGCTCACCATTGCAGCGGTAACGCAGGTTGGCCCAGATGATGTTCTCCTGGGAGAAGCAGAAGA
CCCCCAGGCGGCCACCCCGCATGGTTGTGTCCAAGACCACG

Underlined sequences indicate oligonucleotide primer sequences; Bolded base pairs indicate SNP location; Highlighted sequences indicate nuclease cutting sites.

PCR CONDITIONS:	
Polymerase	Super Therm DNA Polymerase
Cycling	Standard PCR for 35 cycles
Annealing temperature	58˚C
Mg^{2+} concentration	1.5mM
Oligonucleotides	5'-CTACTGTCTCCATGCAGC-3' 5'-CGTGGTCTTGGACACAAC-3'
Amplicon size	415bp
RESTRICTION CONDITIONS:	
Nuclease	*MspI*
Cutting site	C'CG_G
Reaction Buffer	New England Biolabs® Inc. Buffer 4
BSA concentration	0%
Incubation temperature	37˚C
Fragment sizes	T allele – 258 and 157bp C allele – 157, 151 and 107bp

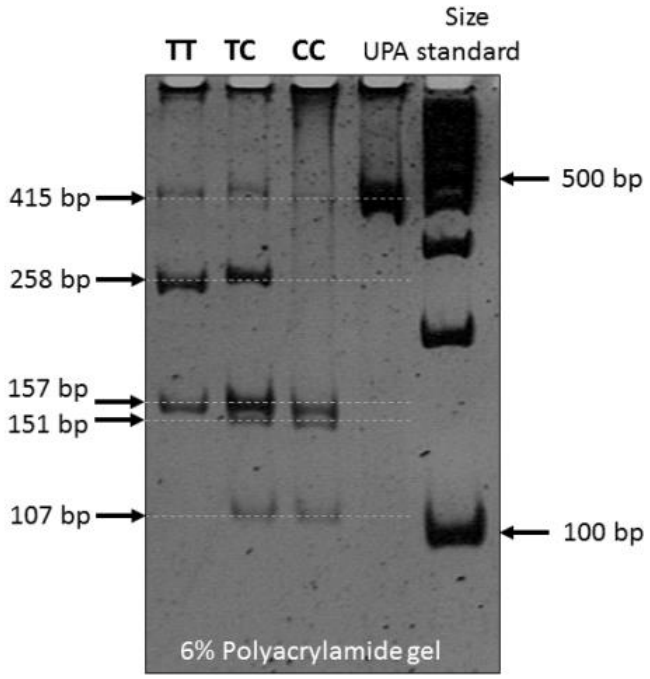

Figure B.5: A typical 6% PAGE image discriminating restriction fragments from individuals who are homozygous TT (lane 1), heterozygous TC (lane 2) and homozygous CC (lane 3) for the rs28494505 polymorphism, together with an uncut PCR amplicon (UPA) in lane 4 and a 100bp size standard (lane 5)
Fragment sizes are indicated with arrows on the sides of the image. Samples were electrophoresed at 120V for 2 hours.

University of Cape Town

2. PCR CONDITIONS FOR *COL27A1* & *TNC* POLYMORPHISMS INVESTIGATED IN CHAPTER THREE

dbSNP ID:	**rs4143245**
CHROMOSOMAL POSITION:	9:117033022
BASE PAIR CHANGE:	Y – T>C
MINOR ALLELE FREQUENCY:	43.8%
HETEROZYGOSITY:	0.473

FASTA SEQUENCE:

CAACCAGCCTAGAGATGATGTTCAGGACAAATCCATGACAAGCACCTTCTAAATTCGTGACCTCATGTAGCCCCCACCCCAC
GTCACCCTGCAGAGGGAGCTGTCGGTAGCTTGGGGGTCTGCTCCCCTTTTCTGGAGAGCTGAGGGTGGGCACCTCAACCTTCC
TGTTCATTTGCTAACTCAGCCCCTGGCTACCATGCACCCACTGTGTGCCAGGGCCCCCAGGGTCCCACTGCCTGGAGATCCAT
GGGTCAGGGTCCTGCCCCAGGGAGGGGGAAGAGAGGAAAAGCACACCGGGGCCTCCTCCCCAGTAACATCTCCTCCTCTTTC
TCCATCACCTTTTCCAGGGAGAGCCAGGCCTYGAGGGTGACAGTGGCCCCATGGGACCTGATGGGCTGAAGGTAAGTGCCCT
TTTAGGGCAGGGCCTGGGGACCCCGGCAGGGCATTGCTTTCCCATCCAAGACCTAAGTCTCCTCCCAGAAACACTGTGTTCCC
ATAGAGATCTGAGATGACAGCTCCAGCTCC

Underlined sequences indicate oligonucleotide primer sequences; Bolded base pairs indicate SNP location; Highlighted sequences indicate nuclease cutting sites.

PCR CONDITIONS:	
Polymerase	Super Therm DNA Polymerase
Cycling	Standard PCR for 35 cycles
Annealing temperature	57˚C
Mg^{2+} concentration	2.0mM
Oligonucleotides	5'-CAACCAGCCTCGAGATGATG-3' 5'-GGAGCTGGAGCTGTCATCTCA-3'
Amplicon size	525bp
RESTRICTION CONDTIONS	
Nuclease	*XhoI*
Cutting site	C'TCGA_G
Reaction Buffer	New England Biolabs® Inc. Buffer 2
BSA concentration	10%
Incubation temperature	37˚C
Fragment sizes	C allele - 351, 165 and 9bp T allele - 516 and 9bp

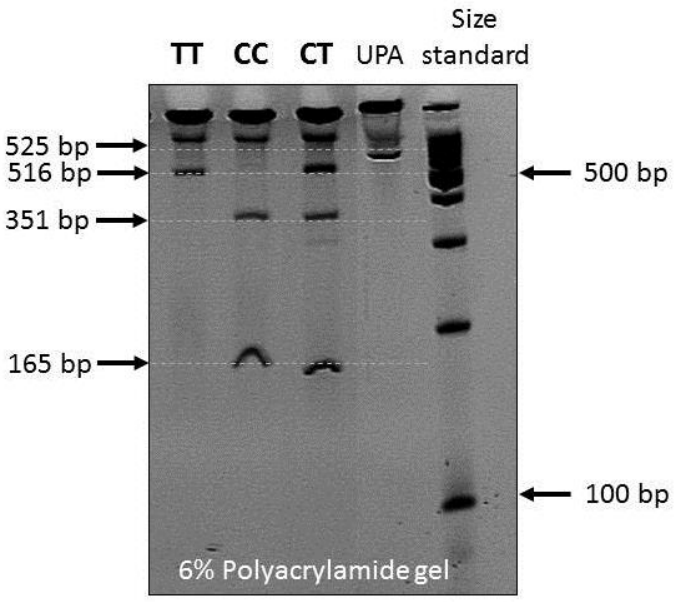

Figure B.6: A typical PAGE image discriminating restriction fragments from individuals who are homozygous TT (lane 1), homozygous CC (lane 2) and heterozygous CT (lane 3) for the rs4143245 polymorphism, together with an uncut PCR amplicon (UPA) in lane 4 and a 100bp size standard (lane 5)
Fragment sizes are indicated with arrows on the sides of the image. Samples were electrophoresed at 120V for 2 hours.

University of Cape Town

dbSNP:	**rs1249744**
CHROMOSOMAL POSITION:	9: 117043352
BASE PAIR CHANGE:	R; A>G
MINOR ALLELE FREQUENCY:	21.2%
HETEROZYGOSITY:	0.381

FASTA SEQUENCE:

GTTGATGCCGCATTAGTCTTTAGAGCAGACATTTCTCCCTTGTGCTGAATCAAAGCCAGCCGCCCCATAATTTTCATGCGCTG
GCCACCCTAGAGAACATGGCTGGCCCTGCCTCCAGAACAGCTGCAGAAATCTGCAGACAGGAACCAAGTCCTCTGGGTTGGT
CCAGATGTTTCCTTGTTCCTCTAGGACCTCCTCTTCCTGTCCCTCTTCTATCACAATTAGTCAGTGCCTCTTCTTAAAGTACAAT
ACCT**R**TTCCCAAATGGGACATCACCTCCTTCATTCTACATCTGGTGCCTCTGTTAATGCAACCAGAGGGATTACATCAAAGCA
GCTCCTTAGTGCTTACCAAGAGCATAGACATTCTTTCTAGATCCTTTGAGGCTTGCAGTCA

Underlined sequences indicate oligonucleotide primer sequences; Bolded base pairs indicate SNP location

PCR CONDITIONS:	
Polymerase	New England Biolabs® Inc. *Taq* DNA Polymerase
Cycling	T-ARMS protocol
Annealing temperature	56˚C
Mg^{2+} concentration	2.0mM
Oligonucleotides	5'-GTTGATGCCGCATTAGTCTTTAGAGCAG-3' 5'-AGTGCCTCTTCTTAAAGTACAATACATA-3' 5'-TGAAGGAGGTGATGTCCCATTTGGGCAC-3' 5'-TGACTGCAAGCCTCAAAGGATCTAGAAA-3'
Amplicon size	Control – 394bp A Allele – 167bp G Allele – 282bp

University of Cape Town

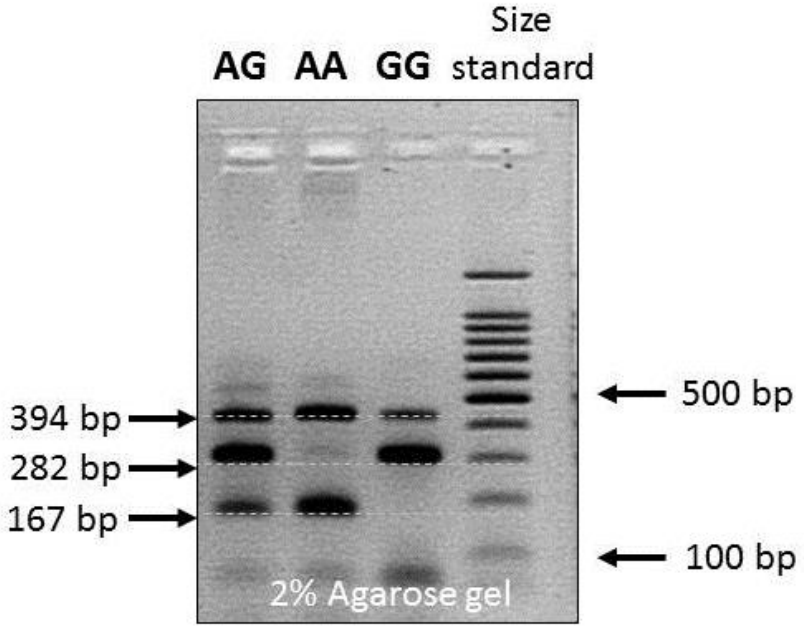

Figure B.7: A typical agarose gel image discriminating restriction fragments from individuals who are heterozygous AG (lane 1), homozygous AA (lane 2) and homozygous GG (lane 3) for the rs1249744 polymorphism, together with a 100bp size standard (lane 4)
Fragment sizes are indicated with arrows on the sides of the image. Samples were electrophoresed at 100V for 35 minutes.

dbSNP ID:	**rs753085**
CHROMOSOMAL POSITION:	9: 117045447
BASE PAIR CHANGE:	Y; C>T
MINOR ALLELE FREQUENCY:	25.0%
HETEROZYGOSITY:	0.425

FASTA SEQUENCE:

GGTCTATTGCATCTGTGCGTGTTGGTGGAAATACTGCCTAAGGGAGGGCTGCCGTGTAACTCCTCCCCACTCTGCATTCAGTG
ACGTCATGTCAGTGGCTTGAAACCAGACATGGAGGGAGCATTTACA**Y**CGCAGAAACTGGCAAATACCACAAACCAGCCTGG
TTGCTTAACCTTACACACAACCAGATCGGAGCTGCTGTGTCACCCTCTTCAGCCTCCTTCCCTGGCTTCCCCCTGGGGTCTTCA
TCTCCTGGAGGTATCAAAGCCCTCTGGGCCCCTTCCTTCAGAGCAGCCCAGCAATGCTGGGACACCCTCAGAGTCATCCAGG
GACCTGCAGCAGCATCACTCCATTCCTCCTGAATCTATTTCCAGGGGGAAATCTCCTTGTTTGATAAAAACACCCTAAGAATC
TGGTCCTTTCTCCCCCGATGCTAATAACAAGGCTTTATAATTTGGGTCCATAGAGGGCTTCAGG

Underlined sequences indicate oligonucleotide primer sequences; Bolded base pairs indicate SNP location

PCR CONDITIONS:	
Polymerase	Super Therm DNA Polymerase
Cycling	T-ARMS protocol
Annealing temperature	60.5˚C
Mg^{2+} concentration	2.0mM (1% Dimethyl Sulphoxide incl.)
Oligonucleotides	5'-GGTCTATTGCATCTGTGCGTGTTGGTGG-3' 5'-AAACCAGACATGGAGGGAGCATTTAAAC-3' 5'-GGTTTGTGGTATTTGCCAGTTTCTGAGA-3' 5'-CCTGAAGCCCTCTATGGACCCAAATTAT-3'
Amplicon size	Control – 477bp T Allele – 157bp C Allele – 375bp

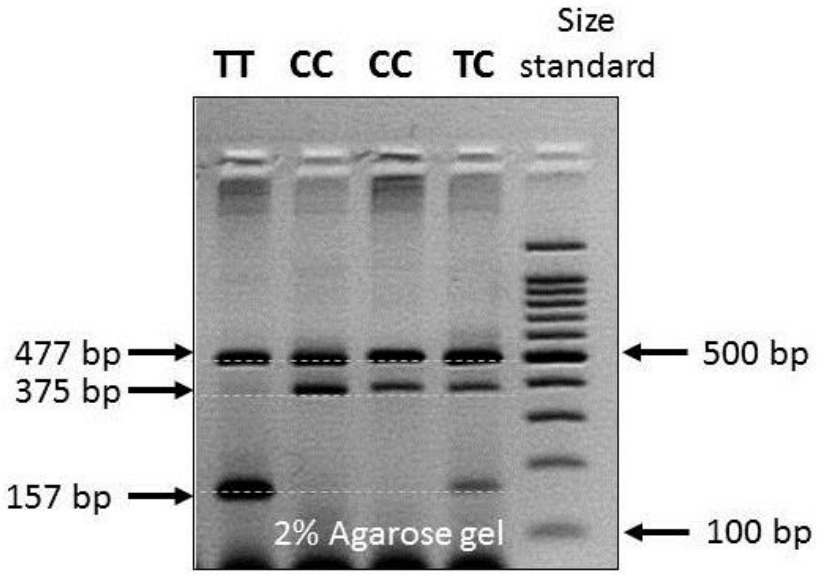

Figure B.8: A typical agarose gel image discriminating restriction fragments from individuals who are homozygous TT (lane 1), homozygous CC (lanes 2 & 3) and heterozygous TC (lane 4) for the rs753085 polymorphism, together with a 100bp size standard (lane 5)

Fragment sizes are indicated with arrows on the sides of the image. Samples were electrophoresed at 100V for 30 minutes.

dbSNP:	**rs946053**
CHROMOSOMAL POSITION:	9: 117049891
BASE PAIR CHANGE:	K; T>G
MINOR ALLELE FREQUENCY:	46.5%
HETEROZYGOSITY:	0.359

FASTA SEQUENCE:

CTCTGTCCAGCCTCTTCCGGCCCCCTTTCAATCCCTATGAACAAAGGGGCTCAGGGGTGCCTGGCTTGAGCCCTGTGCCCAGG
GGAGTGACACCAGTTTCTGGGGCTGGAGTCTTCTGAGCCCTGAGATTCTGCCATTTGCCGTTCCGCTTACACCTTCCTTGTAGT
TGAAAGGTGGCGTCTACACCTGTCCTGTGGGTGGACCAGGGCCTGGGGTGGCCATGGCATAAGGACCTCACCTCTAAACTCT
GTTCCGCCCCCCCCCACCCCAGTCTCTGGCCATGTCTATATGACACAGTCTCACTGCTGCTATTGTTACCGGCTACGGCAATTTT
AACGCCATCTGCTGAGCCCCTGCCACGTGKGCAGCAGCCTGCTGAGTGCTTCCTGTGCCTTTCCTCACCGAGTCCCCTGAACA
GCACCGGGAGGCGCTCCAGTCCCGCTCTGCACATAGGAACC

Underlined sequences indicate oligonucleotide primer sequences; Bolded base pairs indicate SNP location; Highlighted sequences indicate nuclease cutting sites.

PCR CONDITIONS:	
Polymerase	Super Therm DNA Polymerase
Cycling	Standard PCR for 35 cycles
Annealing temperature	52˚C
Mg^{2+} concentration	3.5mM
Oligonucleotides	5'-CTCTGTCCAGCCTCTTC-3' 5'-GGTTCCTATGTGCAGAG-3'
Amplicon size	457bp
RESTRICTION CONDITIONS	
Nuclease	*BglI*
Cutting site	GCC_ _ _ _ '_GGC
Reaction Buffer	New England Biolabs® Inc. Buffer 3
BSA concentration	0%
Incubation temperature	37˚C
Fragment sizes	T allele - 457bp G allele - 361 and 96bp

University of Cape Town

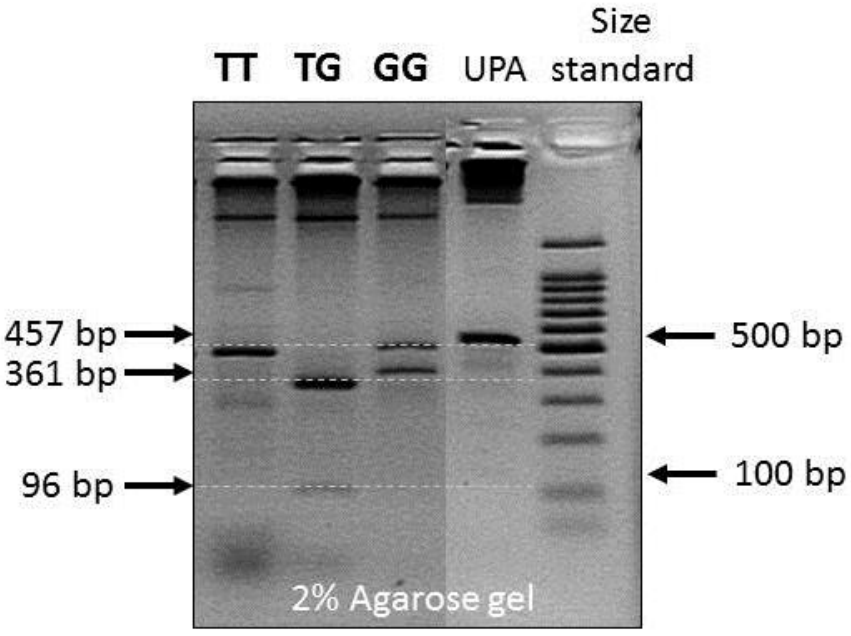

Figure B.9: A typical agarose image discriminating restriction fragments from individuals who are homozygous TT (lane 1), heterozygous TG (lane 2) and homozygous GG (lane 3) for the rs946053 polymorphism, together with an uncut PCR amplicon (UPA) in lane 4 and a 100bp size standard (lane 5)
Fragment sizes are indicated with arrows on the sides of the image. Samples were electrophoresed at 120V for 2 hours.

dbSNP:	**rs13321**
CHROMOSOMAL POSITION:	9: 117792583
BASE PAIR CHANGE:	S; C>G
MINOR ALLELE FREQUENCY:	30.0%
HETEROZYGOSITY:	0.459

FASTA SEQUENCE:

CAGCTCTAAGACCCATCTACTAGGCATAGTTCCAAGATTGCCCATTTCTATGAGTTTTTAATTGTTTGCATCACATTATTGCAC
CTGGAAATTCCCCAGGCATTCTAGTCCATAGAGTTAAGCTAAGAATTAAGCTCCAGCTCAGTCCTCCTGCACTGAGAGCATC
ACCATTTCTCCTTATTCCACAGTTGGACTCCTGTACCCCTTCCCCAAGGACTGCTCCCAAGCAATGCTGAATGGAGACACGAC
CTCTGGCCTCTACACCATTTATCTGAATGGTGATAAGGCTSAGGCGCTGGAAGTCTTCTGTGACATGACCTCTGATGGGGGTG
GATGGATTGTGAGTACCACCAGGGACTGCAGAGCCCTCAGGCCTAGGGGAAGAGGGAGAGGGCAAAGCAGGGGAGGAGAA
GAGGGTGATCTCCTCCTGATGCACTC

Underlined sequences indicate oligonucleotide primer sequences; Bolded base pairs indicate SNP location; Highlighted sequences indicate nuclease cutting sites.

PCR CONDITIONS:	
Polymerase	Super Therm DNA Polymerase
Cycling	Standard PCR for 35 cycles
Annealing temperature	55.0˚C
Mg^{2+} concentration	2.0mM
Oligonucleotides	5'-CAGCTCTAAGACCCATCTACTAGGC-3' 5'-GAGTGCATCAGGAGGAGATCACC-3'
Amplicon size (RFLP)	438bp
RESTRICTION CONDITIONS	
Nuclease	*BbvCI*
Cutting site	CC'TCA_GC
Reaction Buffer	New England Biolabs® Inc. Buffer 4
BSA concentration	0%
Incubation temperature	37˚C
Fragment sizes	C allele - 438bp G allele - 148 and 289bp

University of Cape Town

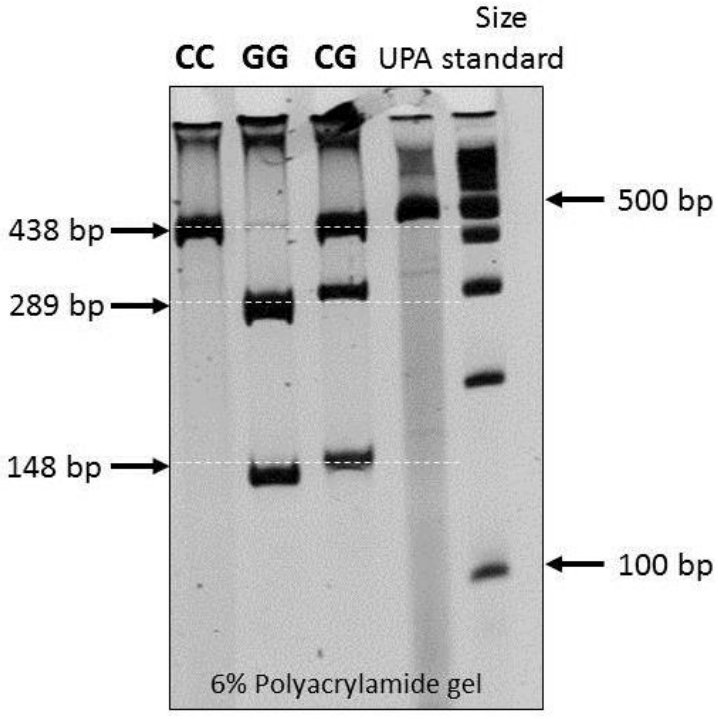

Figure B.10: A typical PAGE image discriminating restriction fragments from individuals who are homozygous CC (lane 1), homozygous GG (lane 2) and heterozygous CG (lane 3) for the rs13321 polymorphism, together with an uncut PCR amplicon (UPA) in lane 4 and a 100bp size standard (lane 5) *Fragment sizes are indicated with arrows on the sides of the image. Samples were electrophoresed at 120V for 2 hours.*

dbSNP:	**rs2104772**
CHROMOSOMAL POSITION:	9: 117808785
BASE PAIR CHANGE:	W; T>A
MINOR ALLELE FREQUENCY:	41.7%
HETEROZYGOSITY:	0.500

FASTA SEQUENCE:

GCCGAACCGGAAGTTGACAACCTTCTGGTTTCAGATGCCACCCCAGACGGTTTCCGTCTGTCCTGGACAGCTGATGAAGGGG
TCTTCGACAATTTTGTTCTCAAAATCAGAGATACCAAAAAGCAGTCTGAGCCACTGGAAATAACCCTACTTGCCCCCGAACG
TACCAGGGACWTAACAGGTCTCAGAGAGGCTACTGAATACGAAATTGAACTCTATGGAATAAGCAAAGGAAGGCGATCCCA
GACAGTCAGTGCTATAG

Underlined sequences indicate oligonucleotide primer sequences; Bolded base pairs indicate SNP location; Highlighted sequences indicate nuclease cutting sites.

PCR CONDITIONS:	
Polymerase	Super Therm DNA Polymerase
Cycling	Standard PCR for 35 cycles
Annealing temperature	52.0˚C
Mg^{2+} concentration	2.0mM
Oligonucleotides	5'-GCCGAACCGGAAGTTGACAACC-3' 5'-CTATAGCACTGACTGTCTGG-3'
Amplicon size	262bp
RESTRICTION CONDITIONS	
Nuclease	*MseI*
Cutting site	T'TAA
Reaction Buffer	New England Biolabs® Inc. Buffer 4
BSA concentration	10%
Incubation temperature	37˚C
Fragment sizes	A allele - 262bp T allele - 175 and 87bp

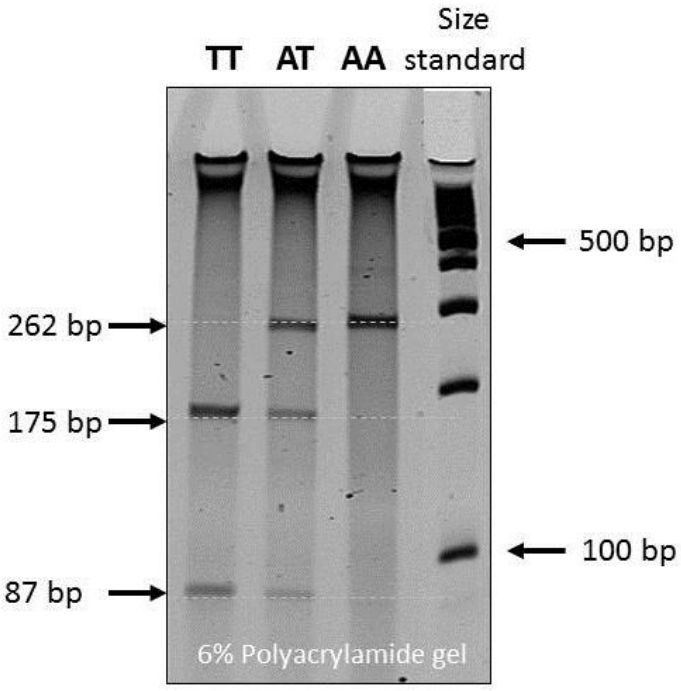

Figure B.11: A typical PAGE image discriminating restriction fragments from individuals who are homozygous TT (lane 1), heterozygous AT (lane 2) and homozygous AA (lane 3) for the rs2104772 polymorphism, together with a 100bp size standard (lane 4)
Fragment sizes are indicated with arrows on the sides of the image. Samples were electrophoresed at 120V for 2 hours.

University of Cape Town

dbSNP:	**rs1330363**
CHROMOSOMAL POSITION:	9: 117813990
BASE PAIR CHANGE:	R; A>G
MINOR ALLELE FREQUENCY:	39.8%
HETEROZYGOSITY:	0.479

FASTA SEQUENCE:

ACAGCTGCTCTTAGGAAGTGGAGAGAACGTTCTCAAGAGTTGTATAAAGCTTTTAGGGTTTTTTGTGCCCAGTTTTATTGGCT
GTTTTATCCCTAGTTTAGAGATTCTGCCTAAAAAGTGACTCCTAAGTTAGTAAAGGCCTAAGATCGAAATCCTGTTCATAAAT
CATGCTAGGTAACRCAGTTCATATATGTGCCAAAAGTATAACCATCTTTTTCTTTAAAAAATAATTATTCTTAAAACTGACAA
CACAGATGCAGCTAGATGCAGCTAATTTGAAATTCCTGTCTCCTTTTTTTCTTTTGATTAAAAAAAAAAATGCATTTACCAGTA
CTGTCAGAAAAAAAAAAAAAAGAAAAAAAAGAAAAAAGAAAAATCACCTAGAAAAGAGACGCTGAACCTTCTCCACTAAT
CCTCCTTA

Underlined sequences indicate oligonucleotide primer sequences; Bolded base pairs indicate SNP location; Highlighted sequences indicate nuclease cutting sites.

PCR CONDITIONS:	
Polymerase	Super Therm DNA Polymerase
Cycling	Standard PCR for 35 cycles
Annealing temperature	55.0˚C
Mg^{2+} concentration	2.0mM
Oligonucleotides	5'-ACAGCTGCTCTTAGGAAGTGGAGAGAAC-3' 5'-TAAGGAGGATTAGTGGAGAAGGTTCAGC-3'
Amplicon size (RFLP)	420bp
RESTRICTION CONDITITIONS	
Nuclease	*TaaI (HpyCH4III)*
Cutting site	AC_n'GT
Reaction Buffer	Fermentas Life Sciences Buffer Tango
BSA concentration	0%
Incubation temperature	65˚C
Fragment sizes	A allele - 182 and 238bp G allele - 420bp

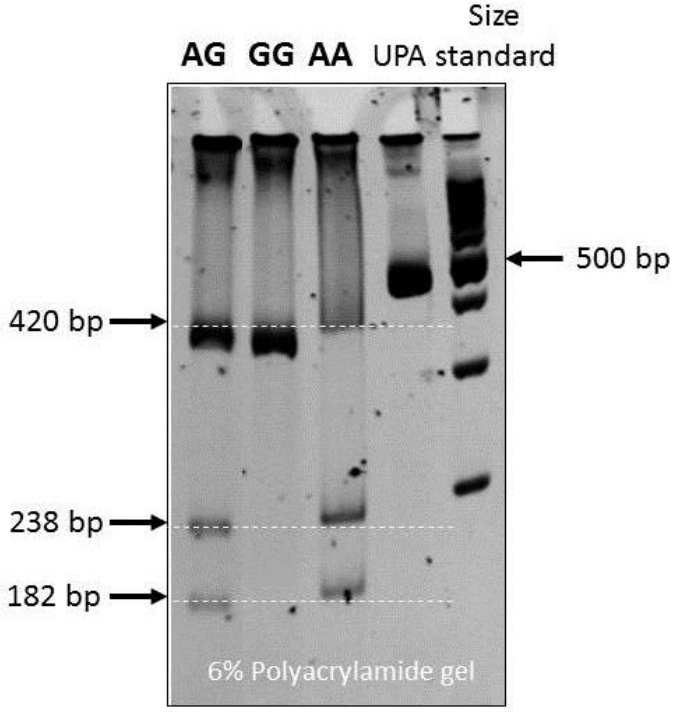

Figure B.12: A typical PAGE image discriminating restriction fragments from individuals who are heterozygous AG (lane 1), homozygous GG (lane 2) and homozygous AA (lane 3) for the rs1330363 polymorphism, together with an uncut PCR amplicon (UPA) in lane 4 and a 100bp size standard (lane 5)

Fragment sizes are indicated with arrows on the sides of the image. Samples were electrophoresed at 120V for 2 hours.

University of Cape Town

3. PCR CONDITIONS FOR *COL5A3, COL3A1* AND *COL5A2* POLYMORPHISMS INVESTIGATED IN CHAPTER FOUR

dbSNP:	**rs2161468**
CHROMOSOMAL POSITION:	19:10088271
BASE PAIR CHANGE:	S; C>G
MINOR ALLELE FREQUENCY:	36%
HETEROZYGOSITY:	0.49

FASTA SEQUENCE:

TGGCCCTGCAGGCAAGAAGGGGTCCC**S**GGTAAGTGACTTGCCCATCCCATTC

Bolded base pair indicates SNP location

PCR CONDITIONS: TaqMan® SNP Genotyping Assays (Applied Biosystems™) Assay ID C__2885020_10

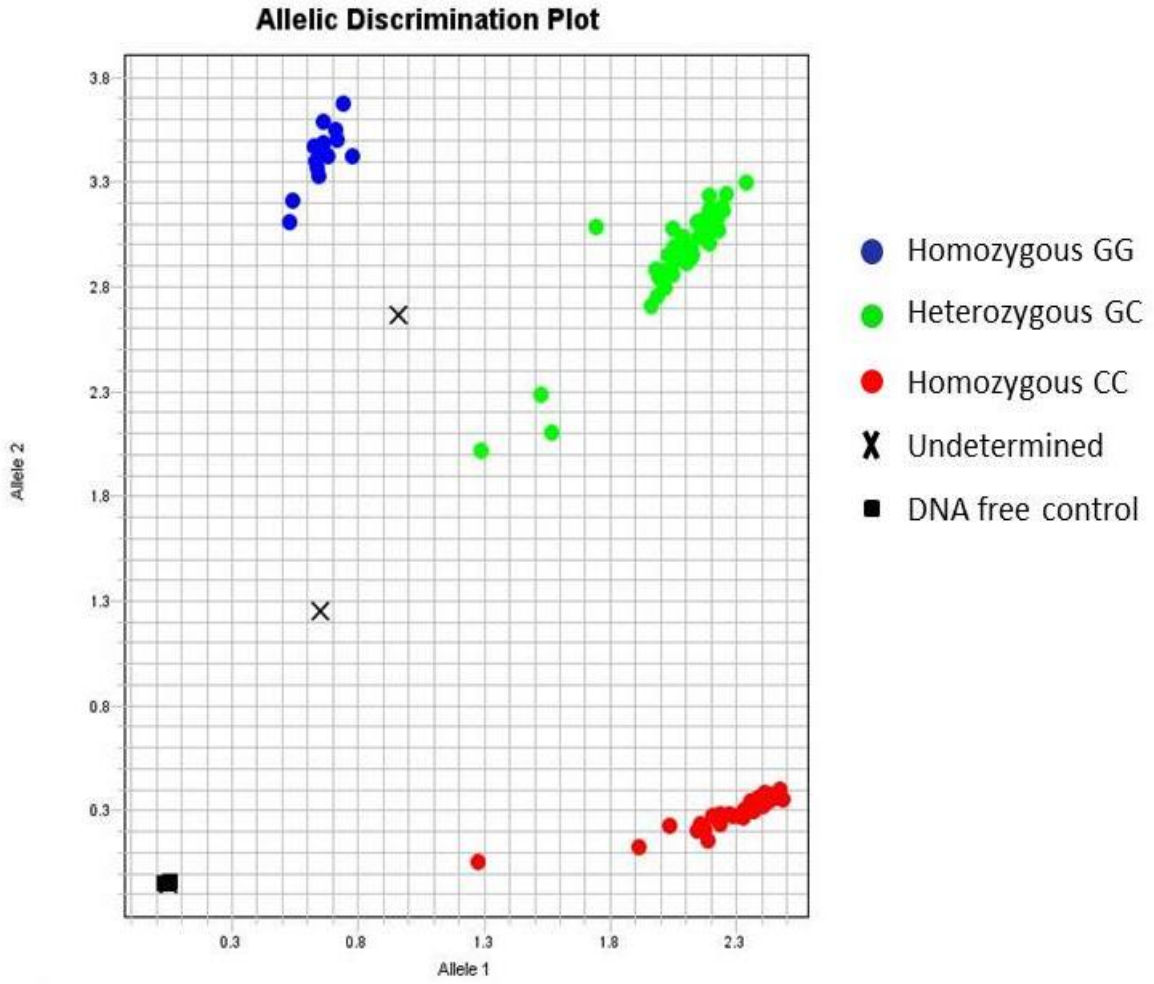

Figure B.13: A typical allelic discrimination plot using the TaqMan® SNP Genotyping Assay for rs2161468 on the StepOnePlus™ Real-Time PCR System

dbSNP:	**rs1559186**
CHROMOSOMAL POSITION:	19:10106936
BASE PAIR CHANGE:	S; C>G
MINOR ALLELE FREQUENCY:	40%
HETEROZYGOSITY:	0.43

FASTA SEQUENCE:

GCAGTTCTGCAGCAGACTCAGGTGAGTGGGAGGTGTGGACATAGTGGGAGGAACCCCAGAGAGTGATAGAGCCTCGGCGTA
AGGAGGGGCTGGGCAGGAGAAGGGGGTTATTTTTCTCCTGAAGATGCCTCTGTGGGGAATGTTGGGGGTCTTGTGGGAGGCC
CCAAGCTCTGACTCCTCTGTCTTTCCAGCT**S**TCTATGAAAGGCCCCCCTGGTCCAGTGGGGCTCACTGGGCGCCCAGGCCCTG
TGGTGAGTAAGGCGACTGCGTGGAGGGGGCGGTGCATGCATTAGTGAAGAATCATCTACCCACACTCCTCCATCTCTCC

Underlined sequences indicate oligonucleotide primer sequences; Bolded base pairs indicate SNP location; Highlighted sequences indicate nuclease cutting sites.

PCR CONDITIONS:	
Polymerase	Super Therm DNA Polymerase
Cycling	Standard PCR for 35 cycles
Annealing temperature	58.0˚C
Mg^{2+} concentration	1.5Mm
Oligonucleotides	5'-GCAGTTCTGCAGCAGACTCAGG-3' 5'-GGAGAGATGGAGGAGTGTGG-3'
Amplicon size	325bp
RESTRICTION CONDITIONS	
Nuclease	*PvuII*
Cutting site	CAG'CTG
Reaction Buffer	Fermentas Buffer G
BSA concentration	0%
Incubation temperature	37˚C
Fragment sizes	C allele - 325bp G allele - 192 and 133bp

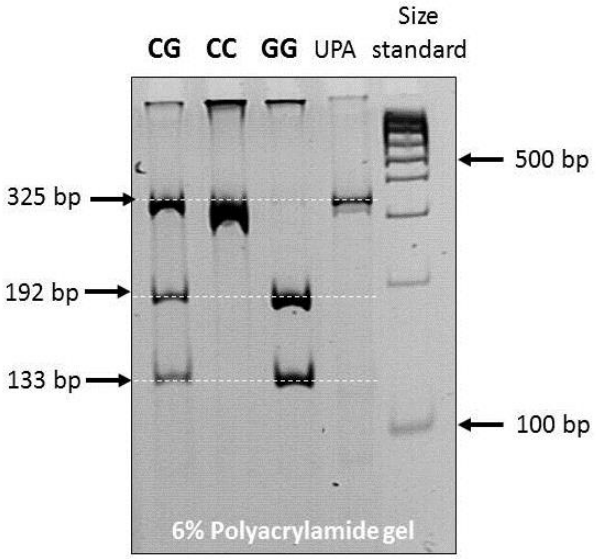

Figure B.14: A typical PAGE image discriminating restriction fragments from individuals who are heterozygous CG (lane 1), homozygous CC (lane 2) and homozygous GG (lane 3) for the rs1559186 polymorphism, together with an uncut PCR amplicon (UPA)(lane 4) and a 100bp size standard (lane 5)

Fragment sizes are indicated with arrows on the sides of the image. Samples were electrophoresed at 120V for 2 hours.

University of Cape Town

dbSNP: **rs2303099**
CHROMOSOMAL POSITION: 19:10116375
BASE PAIR CHANGE: M; A<C
MINOR ALLELE FREQUENCY: 46%
HETEROZYGOSITY: 0.50

FASTA SEQUENCE:

TCCCCTTGCAGGTGGCACCGTGTGGC**M**GTCAGCATAGATGGTGAGATGGTGA

Bolded base pair indicates SNP location

PCR CONDITIONS: TaqMan® SNP Genotyping Assays (Applied Biosystems™)
Assay ID C__2884948_1_

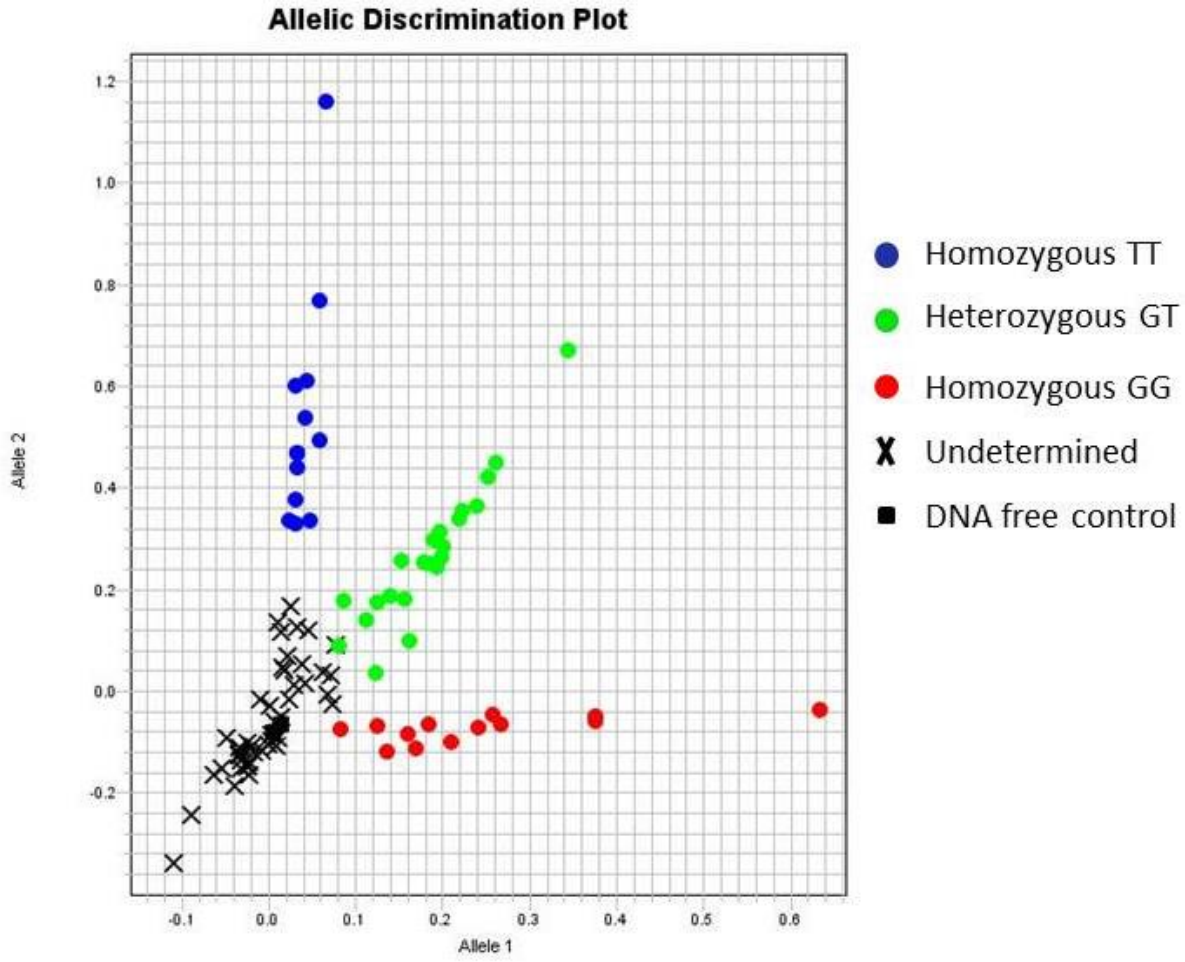

Figure B.15: A typical allelic discrimination plot using the TaqMan® SNP Genotyping Assay for rs2303099 on the StepOnePlus™ Real-Time PCR System

dbSNP:	**rs2056156**
CHROMOSOMAL POSITION:	2:189848468
BASE PAIR CHANGE:	Y; T>C
MINOR ALLELE FREQUENCY:	38.80%
HETEROZYGOSITY:	0.490

FASTA SEQUENCE:

CACCATCTCACAAAAATATAAATTAC**Y**GAACAGTACATCATTTGTTGTTTGC

Bolded base pair indicates SNP location

PCR CONDITIONS: TaqMan® SNP Genotyping Assays (Applied Biosystems™)
Assay ID C___1729165_20

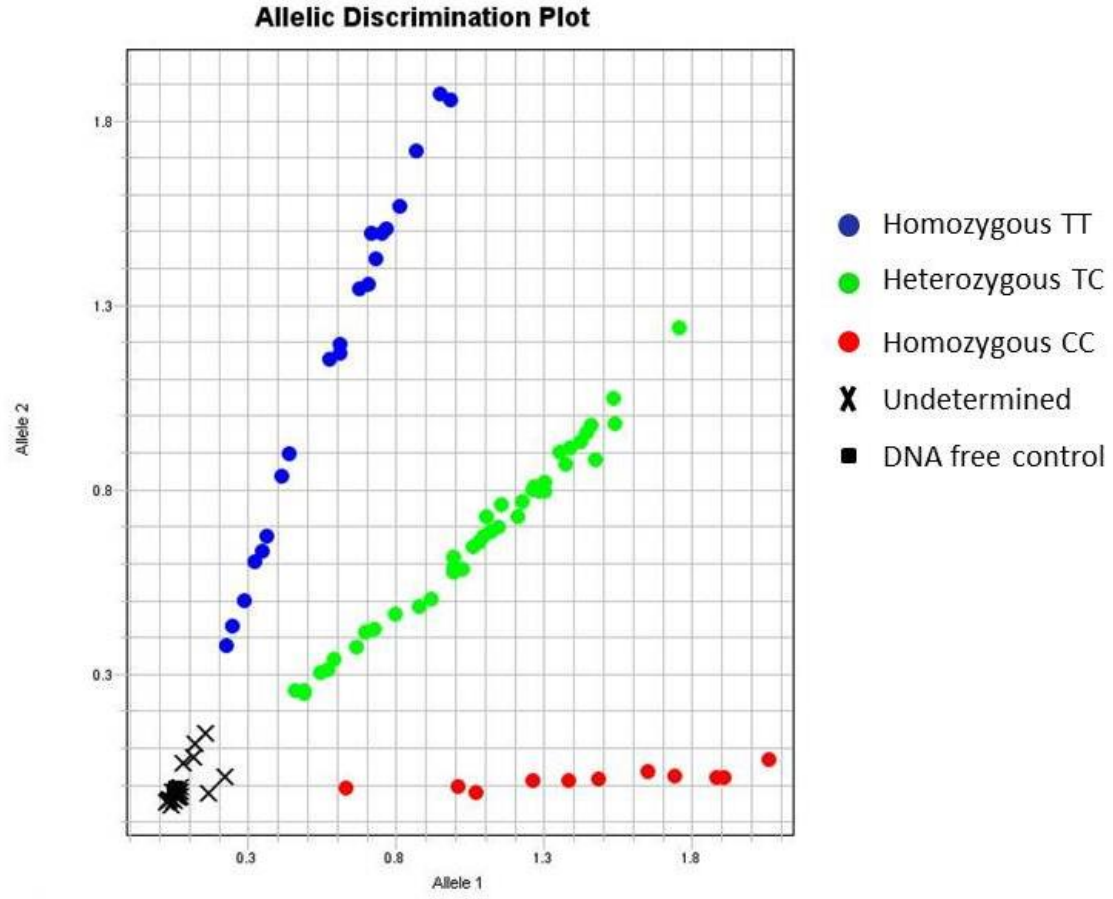

Figure B.16: A typical allelic discrimination plot using the TaqMan® SNP Genotyping Assay for rs2056156 on the StepOnePlus™ Real-Time PCR System

dbSNP:	**rs3106796**
CHROMOSOMAL POSITION:	2:189849773
BASE PAIR CHANGE:	R; A>G
MINOR ALLELE FREQUENCY:	38.2%
HETEROZYGOSITY:	0.490

FASTA SEQUENCE:

CCTATATCATAGGAGCCTAAAAGGGA**R**TGAAAGTCATGTTCATCAAATAGCC

Bolded base pair indicates SNP location

PCR CONDITIONS: TaqMan® SNP Genotyping Assays (Applied Biosystems™)
Assay ID C___1729164_10

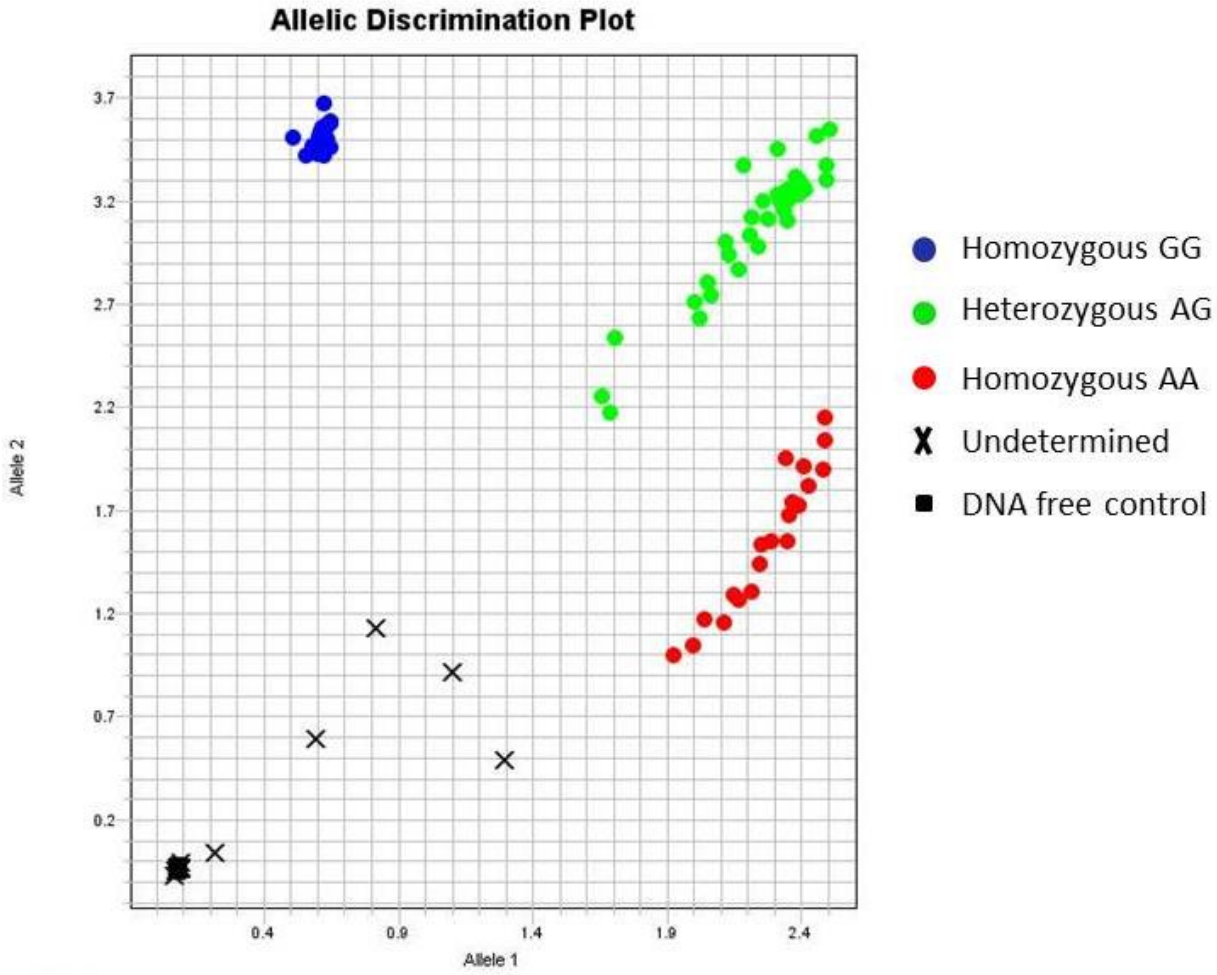

Figure B.17: A typical allelic discrimination plot using the TaqMan® SNP Genotyping Assay for rs3106796 on the StepOnePlus™ Real-Time PCR System

dbSNP:	**rs1800255**
CHROMOSOMAL POSITION:	2:189864080
BASE PAIR CHANGE:	R; G>A
MINOR ALLELE FREQUENCY:	20.9%
HETEROZYGOSITY:	0.380

FASTA SEQUENCE:

CAGGGGCCCCAGGACTTAGAGGTGGA**R**CTGGTCCCCCTGGTCCCGAAGGAGG

Bolded base pair indicates SNP location

PCR CONDITIONS: TaqMan® SNP Genotyping Assays (Applied Biosystems™)
Assay ID C___7477926_10

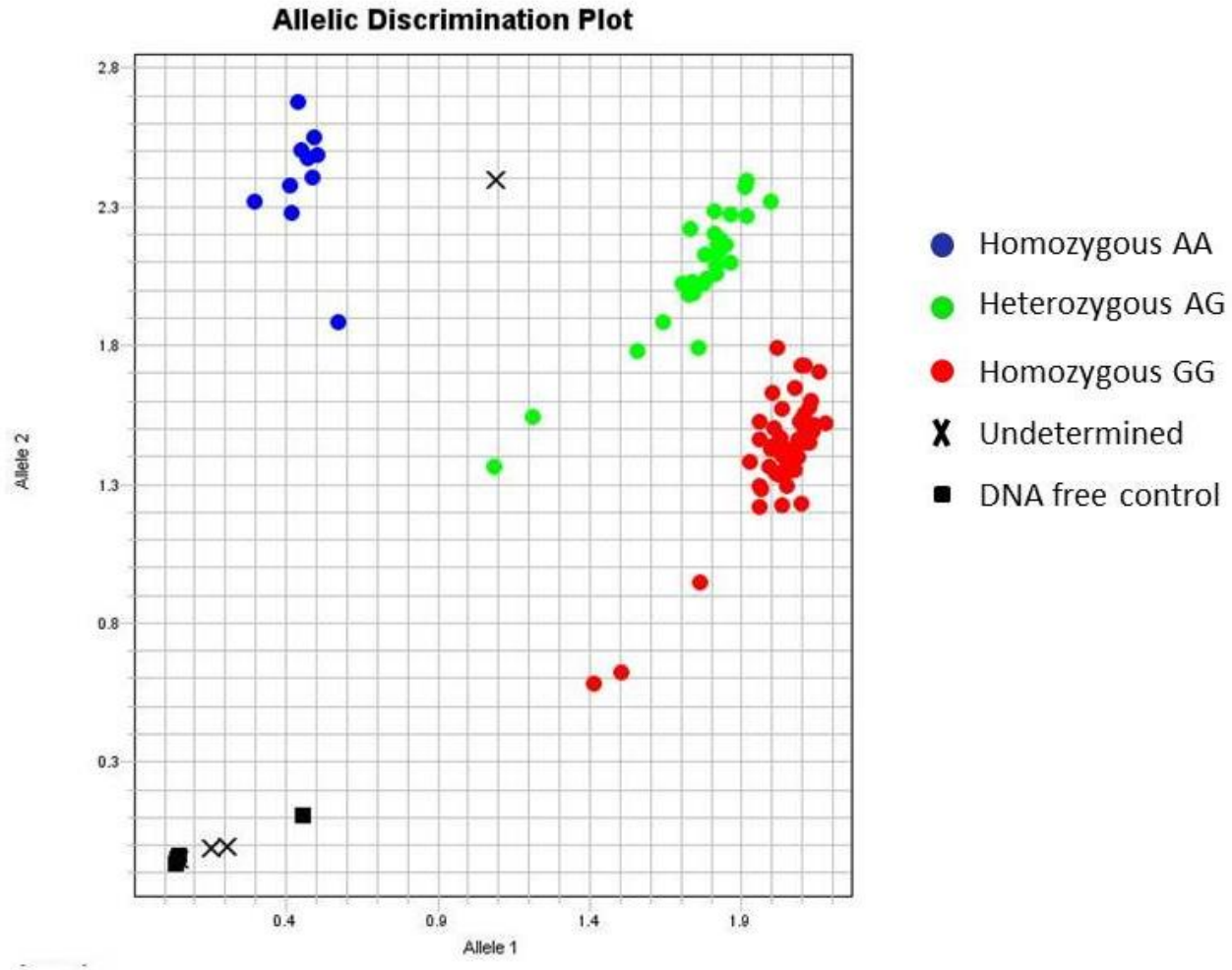

Figure B.18: A typical allelic discrimination plot using the TaqMan® SNP Genotyping Assay for rs1800255 on the StepOnePlus™ Real-Time PCR System

dbSNP: **rs13031549**
CHROMOSOMAL POSITION: 2:190030816
BASE PAIR CHANGE: K; T>G
MINOR ALLELE FREQUENCY: 48.6%
HETEROZYGOSITY: 0.480

FASTA SEQUENCE:

ACTAAATAAATCCCCCAAGACATCAC**K**TAGATATGGTCATACTAAATTAGCA

Bolded base pair indicates SNP location

PCR CONDITIONS: TaqMan® SNP Genotyping Assays (Applied Biosystems™)
Assay ID C__30744562_10

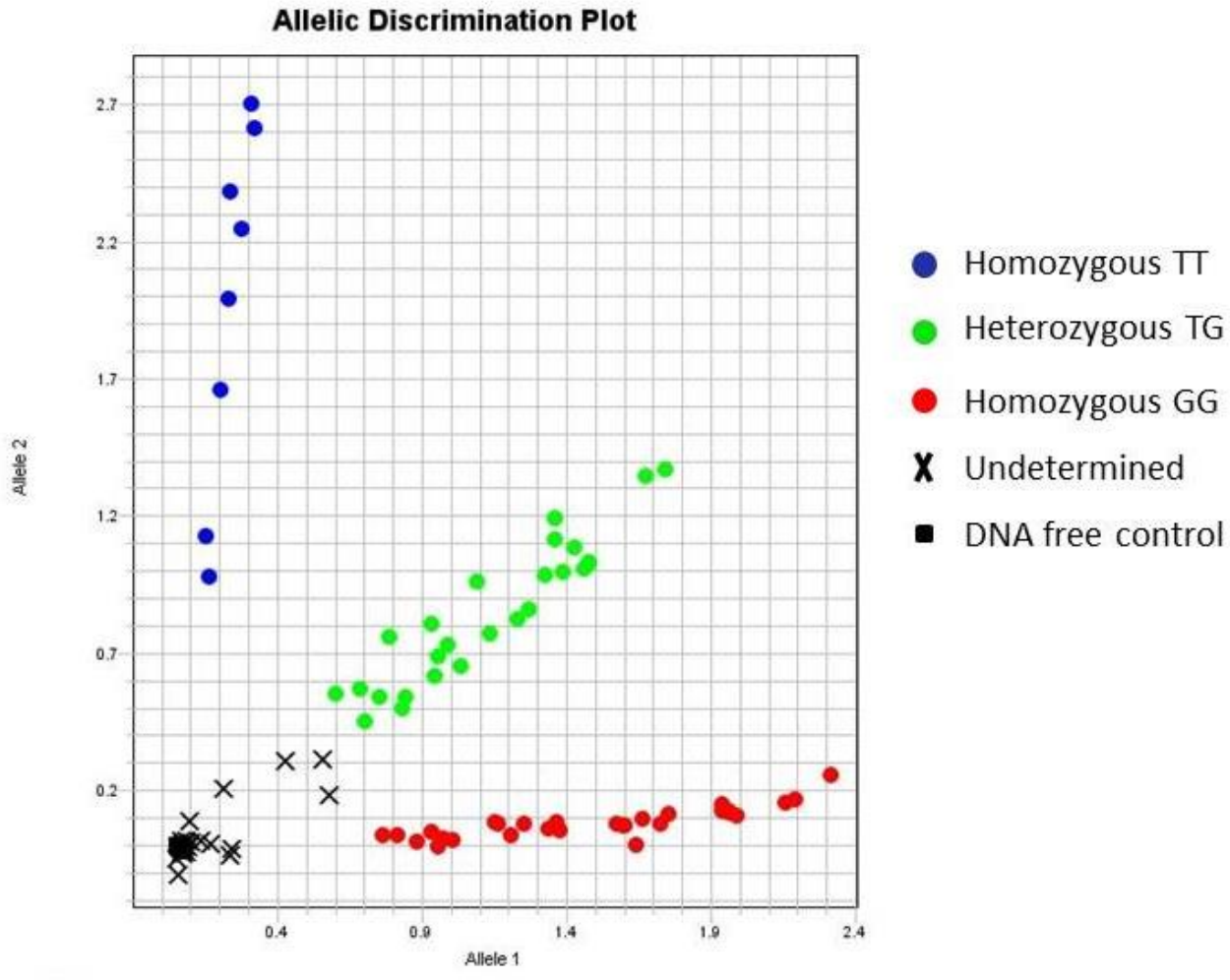

Figure B.19: A typical allelic discrimination plot using the TaqMan® SNP Genotyping Assay for rs13031549 on the StepOnePlus™ Real-Time PCR System

dbSNP:	**rs4667264**
CHROMOSOMAL POSITION:	2:190009718
BASE PAIR CHANGE:	S; C>G
MINOR ALLELE FREQUENCY:	39.1%
HETEROZYGOSITY:	0.460

FASTA SEQUENCE:

GTTCCAATGATCCTTTCTGAACTAGC**S**CTTCTTAGAGTTTAGAACCTCATGG

Bolded base pair indicates SNP location

PCR CONDITIONS: TaqMan® SNP Genotyping Assays (Applied Biosystems™)
Assay ID C__27944619_10

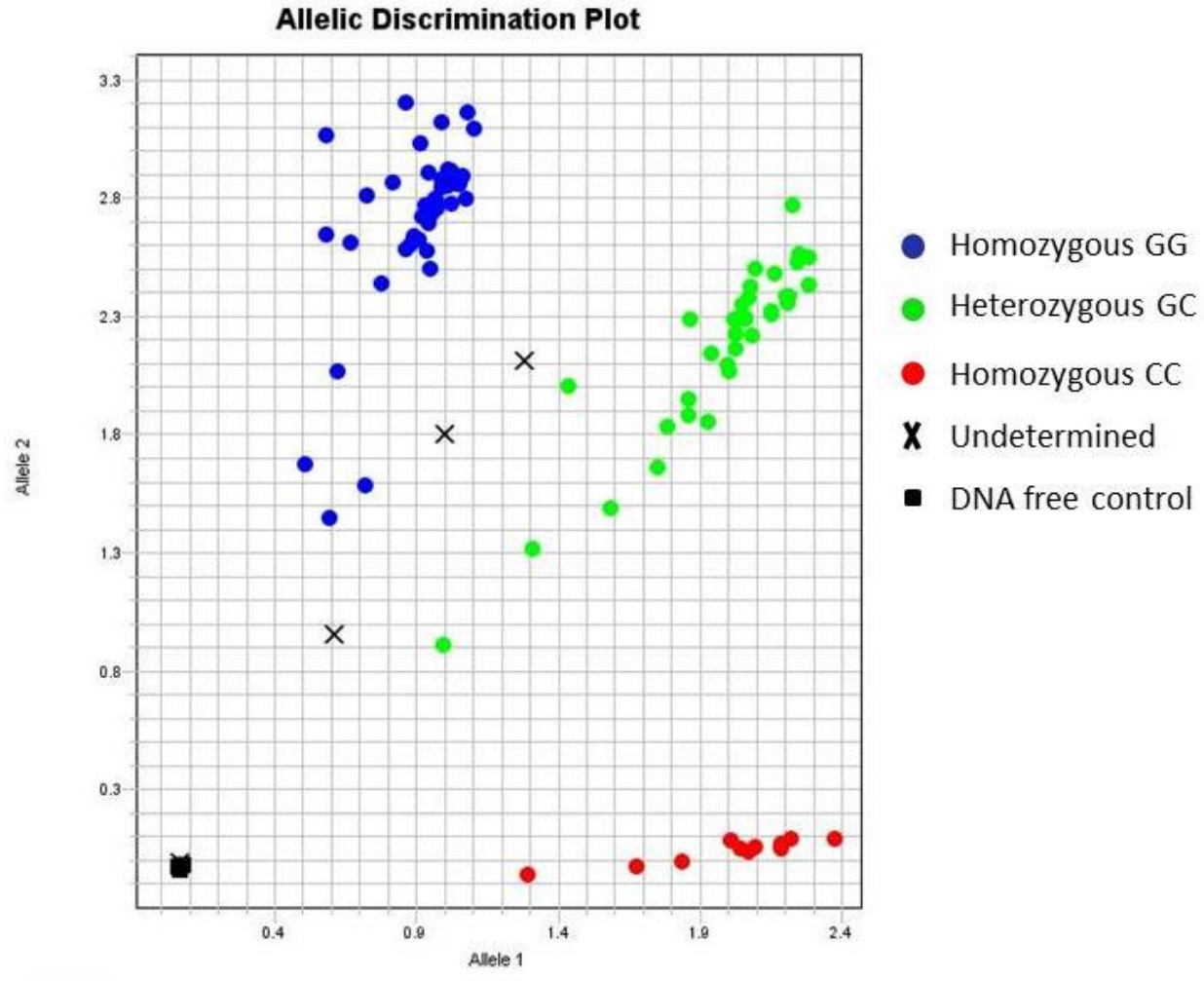

Figure B.20: A typical allelic discrimination plot using the TaqMan® SNP Genotyping Assay for rs4667264 on the StepOnePlus™ Real-Time PCR System

(C) SUPPLEMENTARY RESULTS

Table C.1: Allele combination frequency distributions in CON and TEN participants for significant interactions in the modulation of risk of AT

Values are frequencies with P-values for specific differences in allele combinations between CON and TEN

rs1249744	rs1800795	CON	TEN	P
A	C	0.33	0.25	**0.002**
A	G	0.40	0.41	0.760
G	C	0.13	0.13	0.761
G	G	0.14	0.21	**0.001**

rs1249744	rs16944	CON	TEN	P
A	C	0.41	0.40	0.952
A	T	0.32	0.26	**0.010**
G	C	0.13	0.21	**0.011**
G	T	0.14	0.13	0.316

rs1249744	rs1143627	CON	TEN	P
A	C	0.28	0.26	0.134
A	T	0.45	0.40	0.284
G	C	0.13	0.11	0.583
G	T	0.14	0.23	**0.005**

rs1249744	rs1045485	CON	TEN	P
A	C	0.14	0.09	**0.031**
A	G	0.59	0.57	0.253
G	C	0.07	0.05	0.982
G	G	0.20	0.29	**0.005**

rs1249744	rs3834129	CON	TEN	P
A	CTTACT	0.34	0.27	**0.007**
A	del	0.38	0.39	0.945
G	CTTACT	0.15	0.19	0.156
G	del	0.12	0.16	**0.022**

rs946053	rs1800795	CON	TEN	P
T	C	0.24	0.15	**0.002**
G	C	0.22	0.22	0.682
T	G	0.25	0.28	0.704
G	G	0.29	0.35	**0.012**

rs946053	rs16944	CON	TEN	P
T	T	0.23	0.15	**0.005**
T	C	0.26	0.28	0.806
G	T	0.23	0.25	0.649
G	C	0.28	0.32	0.072

rs13321	rs1800795	CON	TEN	P
C	C	0.33	0.27	**0.020**
G	C	0.13	0.11	0.345
C	G	0.38	0.40	0.320
G	G	0.16	0.22	**0.018**

University of Cape Town

rs2104772	rs1800795	CON	TEN	P
T	C	0.21	0.12	**0.000**
A	C	0.25	0.26	0.887
T	G	0.23	0.26	0.825
A	G	0.31	0.37	**0.010**

rs2104772	rs16944	CON	TEN	P
T	T	0.19	0.13	**0.007**
T	C	0.24	0.25	0.547
A	T	0.26	0.27	0.796
A	C	0.30	0.36	**0.014**

rs2104772	rs1045485	CON	TEN	P
T	C	0.08	0.04	**0.047**
T	G	0.35	0.33	0.115
A	C	0.13	0.10	0.514
A	G	0.44	0.52	**0.004**

rs2104772	rs3834129	CON	TEN	P
T	CTTACT	0.21	0.18	0.052
T	del	0.22	0.19	0.213
A	CTTACT	0.28	0.27	0.891
A	del	0.29	0.35	**0.007**

rs1330363	rs1800795	CON	TEN	P
A	C	0.28	0.20	**0.003**
G	C	0.18	0.18	0.823
A	G	0.30	0.32	0.777
G	G	0.24	0.30	**0.004**

rs1330363	rs1045485	CON	TEN	P
A	C	0.10	0.07	0.081
A	G	0.49	0.45	0.208
G	C	0.11	0.07	0.362
G	G	0.30	0.41	**0.005**

rs1330363	rs3831429	CON	TEN	P
A	CTTACT	0.31	0.22	**0.004**
A	del	0.28	0.30	0.704
G	CTTACT	0.18	0.24	0.147
G	del	0.23	0.24	0.148

rs2303099	rs1800795	CON	TEN	P
G	C	0.22	0.14	**0.017**
T	C	0.24	0.23	0.256
T	G	0.28	0.32	0.233
G	G	0.26	0.30	0.062

University of Cape Town

Reference List

1. (1999) Freely associating. Nat Genet 22(1): 1-2.

2. (2013) Athleticode. http://athleticode.com

3. (2013) DNAFit. https://www.dnafit.com

4. Abate, M; Schiavone, C; Salini, V, and Andia, I. (2013) Occurrence of tendon pathologies in metabolic disorders. Rheumatology 52(4): 599-608.

5. Abate, M; Silbernagel, KG; Siljeholm, C; Di Iorio, A; De Amicis, D; Salini, V; Werner, S, and Paganelli, R. (2009) Pathogenesis of tendinopathies: inflammation or degeneration? Arthritis Res Ther 11(3): 235

6. Abrahams, Y; Laguette, MJ; Prince, S, and Collins, M. (2013) Polymorphisms within the COL5A1 3'-UTR That Alters mRNA Structure and the MIR608 Gene are Associated with Achilles Tendinopathy. Ann Hum Genet doi: 10.1111/ahg.12013. [Epub ahead of print]

7. Adams, JC and Lawler, J. (2004) The thrombospondins. Int J Biochem Cell Biol 36(6): 961-968.

8. Adzhubei, IA; Schmidt, S; Peshkin, L; Ramensky, VE; Gerasimova, A; Bork, P; Kondrashov, AS, and Sunyaev, SR. (2010) A method and server for predicting damaging missense mutations. Nat Methods 7(4): 248-249.

9. Ahmed, IM; Lagopoulos, M; McConnell, P; Soames, RW, and Sefton, GK. (1998) Blood supply of the Achilles tendon. J Orthop Res 16(5): 591-596.

10. Akke, M; Liu, J; Cavanagh, J; Erickson, HP, and Palmer, AG, III. (1998) Pervasive conformational fluctuations on microsecond time scales in a fibronectin type III domain. Nat Struct Biol 5(1): 55-59.

11. Alfredson, H; Thorsen, K, and Lorentzon, R. (1999) In situ microdialysis in tendon tissue: high levels of glutamate, but not prostaglandin E2 in chronic Achilles tendon pain. Knee Surg Sports Traumatol Arthrosc 7(6): 378-381.

12. Altarac, M; Gardner, JW; Popovich, RM; Potter, R; Knapik, JJ, and Jones, BH. (2000) Cigarette smoking and exercise-related injuries among young men and women. Am J Prev Med 18(3 Suppl): 96-102.

University of Cape Town

13. Altman, AR and Davis, IS. (2012) Barefoot running: biomechanics and implications for running injuries. Curr Sports Med Rep 11(5): 244-250.

14. Amanatullah, DF; Lu, J; Hecht, J; Posey, K; Yik, J; Di Cesare, PE, and Haudenschild, DR. (2012) Identification of a 3Kbp mechanoresponsive promoter region in the human cartilage oligomeric matrix protein gene. Tissue Eng Part A 18(17-18): 1882-1889.

15. Angrist, M. (2011) Here is a human being: At the dawn of personal genomics.

16. Aroen, A; Helgo, D; Granlund, OG, and Bahr, R. (2004) Contralateral tendon rupture risk is increased in individuals with a previous Achilles tendon rupture. Scand J Med Sci Sports 14:30-33.

17. Astrom, M and Rausing, A. (1995) Chronic Achilles tendinopathy. A survey of surgical and histopathologic findings. Clin Orthop Relat Res 316:151-164.

18. Astrom, M and Westlin, N. (1994) Blood flow in chronic Achilles tendinopathy. Clin Orthop Relat Res 308:166-172.

19. Astrom, M and Westlin, N. (1994) Blood flow in the human Achilles tendon assessed by laser Doppler flowmetry. J Orthop Res 12(2): 246-252.

20. Azevedo, LB; Lambert, MI; Vaughan, CL; O'Connor, CM, and Schwellnus, MP. (2009) Biomechanical variables associated with Achilles tendinopathy in runners. Br J Sports Med 43(4): 288-292.

21. Backman, LJ; Andersson, G; Wennstig, G; Forsgren, S, and Danielson, P. (2011) Endogenous substance P production in the Achilles tendon increases with loading in an in vivo model of tendinopathy-peptidergic elevation preceding tendinosis-like tissue changes. J Musculoskelet Neuronal Interact 11(2): 133-140.

22. Backman, LJ; Fong, G; Andersson, G; Scott, A, and Danielson, P. (2011) Substance P is a mechanoresponsive, autocrine regulator of human tenocyte proliferation. PLoS One 6(11): e27209-

23. Bank, RA; TeKoppele, JM; Oostingh, G; Hazleman, BL, and Riley, GP. (1999) Lysylhydroxylation and non-reducible crosslinking of human supraspinatus tendon collagen: changes with age and in chronic rotator cuff tendinitis. Ann Rheum Dis 58(1): 35-41.

24. Banos, CC; Thomas, AH, and Kuo, CK. (2008) Collagen fibrillogenesis in tendon development: current models and regulation of fibril assembly. Birth Defects Res C Embryo Today 84(3): 228-244.

University of Cape Town

25. Battista, V; Combs, J, and Warme, WJ. (2003) Asynchronous bilateral achilles tendon ruptures and androstenediol use. Am J Sports Med 31(6): 1007-1009.

26. Beason, DP; Abboud, JA; Kuntz, AF; Bassora, R, and Soslowsky, LJ. (2011) Cumulative effects of hypercholesterolemia on tendon biomechanics in a mouse model. J Orthop Res 29(3): 380-383.

27. Beason, DP; Hsu, JE; Marshall, SM; McDaniel, AL; Temel, RE; Abboud, JA, and Soslowsky, LJ. (2012) Hypercholesterolemia increases supraspinatus tendon stiffness and elastic modulus across multiple species. J Shoulder Elbow Surg S1058-2746(12):00297-2

28. Behrens, G; Fischer, B; Kohler, S; Park, Y; Hollenbeck, AR, and Leitzmann, MF. (2013) Healthy lifestyle behaviors and decreased risk of mortality in a large prospective study of U.S. women and men. Eur J Epidemiol [Epub ahead of print]

29. Beri, A and Khattri, S. (2008) Tendon rupture and statin therapy: is there a link? Comment on the article by Marie et al. Arthritis Rheum 59(8): 1202

30. Bertelli, R; Valenti, F; Oleggini, R; Caridi, G; Altieri, P; Coviello, DA; Botti, G; Ravazzolo, R, and Ghiggeri, GM. (1998) Cell-specific regulation of alpha1(III) and alpha2(V) collagen by TGF-beta1 in tubulointerstitial cell models. Nephrol Dial Transplant 13(3): 573-579.

31. Birch, HL; Wilson, AM, and Goodship, AE. (1997) The effect of exercise-induced localised hyperthermia on tendon cell survival. J Exp Biol 200(Pt 11): 1703-1708.

32. Birk, DE. (2001) Type V collagen: heterotypic type I/V collagen interactions in the regulation of fibril assembly. Micron 32(3): 223-237.

33. Birkedal-Hansen, H; Moore, WG; Bodden, MK; Windsor, LJ; Birkedal-Hansen, B; DeCarlo, A, and Engler, JA. (1993) Matrix metalloproteinases: a review. Crit Rev Oral Biol Med 4(2): 197-250.

34. Bodmer, W and Bonilla, C. (2008) Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet 40(6): 695-701.

35. Boot-Handford, RP; Tuckwell, DS; Plumb, DA; Rock, CF, and Poulsom, R. (2003) A novel and highly conserved collagen (pro(alpha)1(XXVII)) with a unique expression pattern and unusual molecular characteristics establishes a new clade within the vertebrate fibrillar collagen family. J Biol Chem 278(33): 31067-31077.

University of Cape Town

36. Bornstein, P; Agah, A, and Kyriakides, TR. (2004) The role of thrombospondins 1 and 2 in the regulation of cell-matrix interactions, collagen fibril formation, and the response to injury. International Journal of Biochemistry & Cell Biology 36(6): 1115-1125.

37. Bornstein, P; Armstrong, LC; Hankenson, KD; Kyriakides, TR, and Yang, ZT. (2000) Thrombospondin 2, a matricellular protein with diverse functions. Matrix Biology 19(7): 557-568.

38. Boudko, SP; Engel, J, and Bachinger, HP. (2008) Trimerization and triple helix stabilization of the collagen XIX NC2 domain. J Biol Chem 283(49): 34345-34351.

39. Boudko, SP; Zientek, KD; Vance, J; Hacker, JL; Engel, J, and Bachinger, HP. (2010) The NC2 domain of collagen IX provides chain selection and heterotrimerization. J Biol Chem 285(31): 23721-23731.

40. Brown, JC; Miller, CJ; Posthumus, M; Schwellnus, MP, and Collins, M. (2011) The COL5A1 gene, ultra-marathon running performance, and range of motion. Int J Sports Physiol Perform 6(4): 485-496.

41. Bryant, AL; Clark, RA; Bartold, S; Murphy, A; Bennell, KL; Hohmann, E; Marshall-Gradisnik, S; Payne, C, and Crossley, KM. (2008) Effects of estrogen on the mechanical behavior of the human Achilles tendon in vivo. J Appl Physiol 105(4): 1035-1043.

42. Buchanan, CI and Marsh, RL. (2002) Effects of exercise on the biomechanical, biochemical and structural properties of tendons. Comp Biochem Physiol A Mol Integr Physiol 133(4): 1101-1107.

43. Buratti, E and Baralle, FE. (2004) Influence of RNA secondary structure on the pre-mRNA splicing process. Mol Cell Biol 24(24): 10505-10514.

44. Carlson, CB; Lawler, J, and Mosher, DF. (2008) Structure of thrombospondins. Cellular and Molecular Life Sciences 65(5): 672-686.

45. Carstensen, B; Plummer, M; Laara, E, and Hills, M. (2013) Epi: A Package for Statistical Analysis in Epidemiology R package version 1.1.44.

46. Casparian, JM; Luchi, M; Moffat, RE, and Hinthorn, D. (2000) Quinolones and tendon ruptures. South Med J 93(5): 488-491.

47. 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;

University of Cape Town

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, and Collins, FS. (2007) Replicating genotype-phenotype associations. Nature 447(7145): 655-660.

48. Chazerain, P; Hayem, G; Hamza, S; Best, C, and Ziza, JM. (2001) Four cases of tendinopathy in patients on statin therapy. Joint Bone Spine 68(5): 430-433.

49. Chen, H; Deere, M; Hecht, JT, and Lawler, J. (2000) Cartilage oligomeric matrix protein is a calcium-binding protein, and a mutation in its type 3 repeats causes conformational changes. J Biol Chem 275(34): 26538-26544.

50. Chen, HC; Shah, SH; Li, YJ; Stabler, TV; Jordan, JM, and Kraus, VB. (2008) Inverse association of general joint hypermobility with hand and knee osteoarthritis and serum cartilage oligomeric matrix protein levels. Arthritis Rheum 58(12): 3854-3864.

51. Chen, HY; Chung, YW; Lin, WY; Wang, JC; Tsai, FJ, and Tsai, CH. (2008) Collagen type 3 alpha 1 polymorphism and risk of pelvic organ prolapse. Int J Gynaecol Obstet 103(1): 55-58.

52. Cheung, DT; DiCesare, P; Benya, PD; Libaw, E, and Nimni, ME. (1983) The presence of intermolecular disulfide cross-links in type III collagen. J Biol Chem 258(12): 7774-7778.

53. Childs, SG. (2007) Pathogenesis of tendon rupture secondary to fluoroquinolone therapy. Orthop Nurs 26(3): 175-182.

54. Chou, HT; Hung, JS; Chen, YT; Wu, JY, and Tsai, FJ. (2004) Association between COL3A1 collagen gene exon 31 polymorphism and risk of floppy mitral valve/mitral valve prolapse. Int J Cardiol 95(2-3): 299-305.

55. Cilli, F; Khan, M; Fu, F, and Wang, JH. (2004) Prostaglandin E2 affects proliferation and collagen synthesis by human patellar tendon fibroblasts. Clin J Sport Med 14(4): 232-236.

56. Collins, M; Mokone, GG; September, AV; van der Merwe, L, and Schwellnus, MP. (2009) The COL5A1 genotype is associated with range of motion measurements. Scand J Med Sci Sports 19(6): 803-810.

57. Collins, M and Posthumus, M. (2011) Type V Collagen Genotype and Exercise-Related Phenotype Relationships: A Novel Hypothesis. Exercise and Sport Sciences Reviews 39(4): 191-198.

University of Cape Town

58. Collins, M; Posthumus, M, and Schwellnus, MP. (2009) The COL1A1 gene and acute soft tissue ruptures. Br J Sports Med 44(14): 1063-1064.

59. Collins, M and Raleigh, SM. (2009) Genetic Risk Factors for Musculoskeletal Soft Tissue Injuries. Med Sport Sci 54:136-149.

60. Connizzo, BK; Yannascoli, SM, and Soslowsky, LJ. (2013) Structure-function relationships of postnatal tendon development: A parallel to healing. Matrix Biol 32(2):106-16

61. Cook, JL; Khan, KM, and Purdam, C. (2002) Achilles tendinopathy. Man Ther 7(3): 121-130.

62. Cook, JL and Purdam, CR. (2009) Is tendon pathology a continuum? A pathology model to explain the clinical presentation of load-induced tendinopathy. Br J Sports Med 43(6): 409-416.

63. Coombes, BK; Bisset, L, and Vicenzino, B. (2010) Efficacy and safety of corticosteroid injections and other injections for management of tendinopathy: a systematic review of randomised controlled trials. Lancet 376(9754): 1751-1767.

64. Corps, AN; Robinson, AH; Movin, T; Costa, ML; Hazleman, BL, and Riley, GP. (2005) Increased expression of aggrecan and biglycan mRNA in Achilles tendinopathy. Rheumatology 45(3): 291-294.

65. Dang, CV; O'Donnell, KA; Zeller, KI; Nguyen, T; Osthus, RC, and Li, F. (2006) The c-Myc target gene network. Semin Cancer Biol 16(4): 253-264.

66. Davison, PF. (1978) Bovine tendons. Aging and collagen cross-linking. J Biol Chem 253(16): 5635-5641.

67. de Fraipont, F; Nicholson, AC; Feige, JJ, and Van Meir, EG. (2001) Thrombospondins and tumor angiogenesis. Trends Mol Med 7(9): 401-407.

68. de Jonge, S; van den, BC; de Vos, RJ; van der Heide, HJ; Weir, A; Verhaar, JA; Bierma-Zeinstra, SM, and Tol, JL. (2011) Incidence of midportion Achilles tendinopathy in the general population. Br J Sports Med 45(13): 1026-1028.

69. Del Buono, A; Chan, O, and Maffulli, N. (2012) Achilles tendon: functional anatomy and novel emerging models of imaging classification. Int Orthop 37(4):715-21

University of Cape Town

70. Delot, E; King, LM; Briggs, MD; Wilcox, WR, and Cohn, DH. (1999) Trinucleotide expansion mutations in the cartilage oligomeric matrix protein (COMP) gene. Hum Mol Genet 8(1): 123-128.

71. DiCesare, P; Hauser, N; Lehman, D; Pasumarti, S, and Paulsson, M. (1994) Cartilage oligomeric matrix protein (COMP) is an abundant component of tendon. FEBS Lett 354(2): 237-240.

72. Ding, Y; Chan, CY, and Lawrence, CE. (2005) RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble. RNA 11(8): 1157-1166.

73. Ding, Y and Lawrence, CE. (2003) A statistical sampling algorithm for RNA secondary structure prediction. Nucleic Acids Res 31(24): 7280-7301.

74. Dressler, MR; Butler, DL, and Boivin, GP. (2005) Effects of age on the repair ability of mesenchymal stem cells in rabbit tendon. J Orthop Res 23(2): 287-293.

75. Dressler, MR; Butler, DL; Wenstrup, R; Awad, HA; Smith, F, and Boivin, GP. (2002) A potential mechanism for age-related declines in patellar tendon biomechanics. J Orthop Res 20(6): 1315-1322.

76. El Khoury, L; Posthumus, M; Collins, M; Handley, CJ; Cook, J, and Raleigh, SM. (2013) 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 [Epub ahead of print doi: 10.1016/j.jsams.2013.02.006]

77. Erickson, HP. (1994) Evolution of the tenascin family--implications for function of the C-terminal fibrinogen-like domain. Perspect Dev Neurobiol 2(1): 9-19.

78. Eriksen, HA; Pajala, A; Leppilahti, J, and Risteli, J. (2002) Increased content of type III collagen at the rupture site of human Achilles tendon. J Orthop Res 20(6): 1352-1357.

79. European Society of Human Genetics. (2003) Data storage and DNA banking for biomedical research: technical, social and ethical issues. Eur J Hum Genet 11(Suppl. 2): S8-S10.

80. Farris, DJ; Trewartha, G, and McGuigan, MP. (2011) Could intra-tendinous hyperthermia during running explain chronic injury of the human Achilles tendon? J Biomech 44(5): 822-826.

81. 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, and Eynon, N. (2012) Gene variants within the COL1A1 gene are

University of Cape Town

associated with reduced anterior cruciate ligament injury in professional soccer players. J Sci Med Sport [Epub ahead of print doi: 10.1016/j.jsams.2012.10.004.]

82. Finch, C. (2006) A new framework for research leading to sports injury prevention. J Sci Med Sport 9(1-2): 3-9.

83. Florkowski, CM. (2008) Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. Clin Biochem Rev 29 Suppl 1:S83-S87.

84. Fong, G; Backman, LJ; Hart, DA; Danielson, P; McCormack, B, and Scott, A. (2013) Substance P enhances collagen remodeling and MMP-3 expression by human tenocytes. J Orthop Res 31(1): 91-98.

85. Ford, KR; Manson, NA; Evans, BJ; Myer, GD; Gwin, RC; Heidt, RS, Jr., and Hewett, TE. (2006) Comparison of in-shoe foot loading patterns on natural grass and synthetic turf. J Sci Med Sport 9(6): 433-440.

86. Forsgren, S; Danielson, P, and Alfredson, H. (2005) Vascular NK-1 receptor occurrence in normal and chronic painful Achilles and patellar tendons: studies on chemically unfixed as well as fixed specimens. Regul Pept 126(3): 173-181.

87. Fox, AJ; Bedi, A; Deng, XH; Ying, L; Harris, PE; Warren, RF, and Rodeo, SA. (2011) Diabetes mellitus alters the mechanical properties of the native tendon in an experimental rat model. J Orthop Res 29(6): 880-885.

88. Fredberg, U and Bolvig, L. (2002) Significance of ultrasonographically detected asymptomatic tendinosis in the patellar and achilles tendons of elite soccer players: a longitudinal study. Am J Sports Med 30(4): 488-491.

89. Fredberg, U and Stengaard-Pedersen, K. (2008) Chronic tendinopathy tissue pathology, pain mechanisms, and etiology with a special focus on inflammation. Scand J Med Sci Sports 18(1): 3-15.

90. Frey, H; Schroeder, N; Manon-Jensen, T; Iozzo, RV, and Schaefer, L. (2013) Biological interplay between proteoglycans and their innate immune receptors in inflammation. FEBS J [Epub ahead of print doi: 10.1111/febs.12145.]

91. Fu, SC; Rolf, C; Cheuk, YC; Lui, PP, and Chan, KM. (2010) Deciphering the pathogenesis of tendinopathy: a three-stages process. Sports Med Arthrosc Rehabil Ther Technol 2:30

92. Fuller, CW; Dick, RW; Corlette, J, and Schmalz, R. (2007) Comparison of the incidence, nature and cause of injuries sustained on grass and new generation

University of Cape Town

artificial turf by male and female football players. Part 1: match injuries. Br J Sports Med 41 Suppl 1:i20-i26.

93. Fuller, CW; Dick, RW; Corlette, J, and Schmalz, R. (2007) Comparison of the incidence, nature and cause of injuries sustained on grass and new generation artificial turf by male and female football players. Part 2: training injuries. Br J Sports Med 41 Suppl 1:i27-i32.

94. Gaida, JE; Alfredson, H; Kiss, ZS; Bass, SL, and Cook, JL. (2010) Asymptomatic Achilles tendon pathology is associated with a central fat distribution in men and a peripheral fat distribution in women: a cross sectional study of 298 individuals. BMC Musculoskelet Disord 11:41

95. Gaida, JE; Alfredson, L; Kiss, ZS; Wilson, AM; Alfredson, H, and Cook, JL. (2009) Dyslipidemia in Achilles Tendinopathy Is Characteristic of Insulin Resistance. Medicine and Science in Sports and Exercise 41(6): 1194-1197.

96. Gaida, JE; Ashe, MC; Bass, SL, and Cook, JL. (2009) Is adiposity an under-recognized risk factor for tendinopathy? A systematic review. Arthritis Rheum 61(6): 840-849.

97. Gauderman, WJ. (2002) Sample size requirements for matched case-control studies of gene-environment interaction. Stat Med 21:35-50.

98. Gauderman, WJ and Morrison, JM. (2006) QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies. http://hydra.usc.edu/gxe

99. Ghatnekar, A and Trojanowska, M. (2008) GATA-6 is a novel transcriptional repressor of the human Tenascin-C gene expression in fibroblasts. Biochim Biophys Acta 1779(3): 145-151.

100. Gold, L and Igra, H. (2003) Levofloxacin-induced tendon rupture: a case report and review of the literature. J Am Board Fam Pract 16(5): 458-460.

101. Gudbjartsson, DF; Walters, GB; Thorleifsson, G; Stefansson, H; Halldorsson, BV; Zusmanovich, P; Sulem, P; Thorlacius, S; Gylfason, A; Steinberg, S; Helgadottir, A; Ingason, A; Steinthorsdottir, V; Olafsdottir, EJ; Olafsdottir, GH; Jonsson, T; Borch-Johnsen, K; Hansen, T; Andersen, G; Jorgensen, T; Pedersen, O; Aben, KK; Witjes, JA; Swinkels, DW; den Heijer, M; Franke, B; Verbeek, AL; Becker, DM; Yanek, LR; Becker, LC; Tryggvadottir, L; Rafnar, T; Gulcher, J; Kiemeney, LA; Kong, A; Thorsteinsdottir, U, and Stefansson, K. (2008) Many sequence variants affecting diversity of adult human height. Nat Genet 40(5): 609-615.

University of Cape Town

102. Hansen, M; Kongsgaard, M; Holm, L; Skovgaard, D; Magnusson, SP; Qvortrup, K; Larsen, JO; Aagaard, P; Dahl, M; Serup, A; Frystyk, J; Flyvbjerg, A; Langberg, H, and Kjaer, M. (2009) Effect of estrogen on tendon collagen synthesis, tendon structural characteristics, and biomechanical properties in postmenopausal women. J Appl Physiol 106(4): 1385-1393.

103. Hansen, M; Koskinen, SO; Petersen, SG; Doessing, S; Frystyk, J; Flyvbjerg, A; Westh, E; Magnusson, SP; Kjaer, M, and Langberg, H. (2008) Ethinyl oestradiol administration in women suppresses synthesis of collagen in tendon in response to exercise. J Physiol 586(Pt 12): 3005-3016.

104. Hansen, M; Miller, BF; Holm, L; Doessing, S; Petersen, SG; Skovgaard, D; Frystyk, J; Flyvbjerg, A; Koskinen, S; Pingel, J; Kjaer, M, and Langberg, H. (2009) Effect of administration of oral contraceptives in vivo on collagen synthesis in tendon and muscle connective tissue in young women. J Appl Physiol 106(4): 1435-1443.

105. Hardy, GH. (1908) Mendelian proportions in a mixed population. Science 28(706): 49-50.

106. Hart, DA; Archambault, JM; Kydd, A; Reno, C; Frank, CB, and Herzog, W. (1998) Gender and neurogenic variables in tendon biology and repetitive motion disorders. Clin Orthop Relat Res 351: 44-56.

107. Harvie, P; Ostlere, SJ; Teh, J; McNally, EG; Clipsham, K; Burston, BJ; Pollard, TC, and Carr, AJ. (2004) Genetic influences in the aetiology of tears of the rotator cuff. Sibling risk of a full-thickness tear. J Bone Joint Surg Br 86(5): 696-700.

108. Hecht, JT; Hayes, E; Haynes, R, and Cole, WG. (2005) COMP mutations, chondrocyte function and cartilage matrix. Matrix Biology 23(8): 525-533.

109. Hess, GW. (2010) Achilles tendon rupture: a review of etiology, population, anatomy, risk factors, and injury prevention. Foot Ankle Spec 3(1): 29-32.

110. Hildebrand, KA; Frank, CB, and Hart, DA. (2004) Gene intervention in ligament and tendon: current status, challenges, future directions. Gene Ther 11(4): 368-378.

111. Hiller, M; Zhang, Z; Backofen, R, and Stamm, S. (2007) Pre-mRNA secondary structures influence exon recognition. PLoS Genet 3(11): e204

112. Hirose, Y; Chiba, K; Karasugi, T; Nakajima, M; Kawaguchi, Y; Mikami, Y; Furuichi, T; Mio, F; Miyake, A; Miyamoto, T; Ozaki, K; Takahashi, A; Mizuta, H; Kubo, T; Kimura, T; Tanaka, T; Toyama, Y, and Ikegawa, S. (2008) A functional polymorphism in THBS2 that affects alternative splicing and MMP binding is associated with lumbar-disc herniation. Am J Hum Genet 82(5): 1122-1129.

University of Cape Town

113. Hirschmuller, A; Frey, V; Konstantinidis, L; Baur, H; Dickhuth, HH; Sudkamp, NP, and Helwig, P. (2012) Prognostic value of Achilles tendon Doppler sonography in asymptomatic runners. Med Sci Sports Exerc 44(2): 199-205.

114. Hjorten, R; Hansen, U; Underwood, RA; Telfer, HE; Fernandes, RJ; Krakow, D; Sebald, E; Wachsmann-Hogiu, S; Bruckner, P; Jacquet, R; Landis, WJ; Byers, PH, and Pace, JM. (2007) Type XXVII collagen at the transition of cartilage to bone during skeletogenesis. Bone 41(4): 535-542.

115. Hoffmann, A and Gross, G. (2007) Tendon and ligament engineering in the adult organism: mesenchymal stem cells and gene-therapeutic approaches. Int Orthop 31(6): 791-797.

116. Holmes, GB and Lin, J. (2006) Etiologic factors associated with symptomatic achilles tendinopathy. Foot Ankle Int 27(11): 952-959.

117. Hong, Y; Wang, L; Li, JX, and Zhou, JH. (2012) Comparison of plantar loads during treadmill and overground running. J Sci Med Sport 15(6): 554-560.

118. Horn, S; Gregory, P, and Guskiewicz, KM. (2009) Self-reported anabolic-androgenic steroids use and musculoskeletal injuries: findings from the center for the study of retired athletes health survey of retired NFL players. Am J Phys Med Rehabil 88(3): 192-200.

119. Hosking, L; Lumsden, S; Lewis, K; Yeo, A; McCarthy, L; Bansal, A; Riley, J; Purvis, I, and Xu, CF. (2004) Detection of genotyping errors by Hardy-Weinberg equilibrium testing. Eur J Hum Genet 12(5): 395-399.

120. Hou, J; Putkey, JA, and Hecht, JT. (2000) Delta 469 mutation in the type 3 repeat calcium binding domain of cartilage oligomeric matrix protein (COMP) disrupts calcium binding. Cell Calcium 27(6): 309-314.

121. Houshian, S; Tscherning, T, and Riegels-Nielsen, P. (1998) The epidemiology of Achilles tendon rupture in a Danish county. Injury 29(9): 651-654.

122. Ikegawa, S; Ohashi, H; Nishimura, G; Kim, KC; Sannohe, A; Kimizuka, M; Fukushima, Y; Nagai, T, and Nakamura, Y. (1998) Novel and recurrent COMP (cartilage oligomeric matrix protein) mutations in pseudoachondroplasia and multiple epiphyseal dysplasia. Human Genetics 103(6): 633-638.

123. Imamura, Y; Scott, IC, and Greenspan, DS. (2000) The pro-alpha3(V) collagen chain. Complete primary structure, expression domains in adult and developing tissues, and comparison to the structures and expression domains of the other types V and XI procollagen chains. J Biol Chem 275(12): 8749-8759.

University of Cape Town

124. Inhofe, PD; Grana, WA; Egle, D; Min, KW, and Tomasek, J. (1995) The effects of anabolic steroids on rat tendon. An ultrastructural, biomechanical, and biochemical analysis. Am J Sports Med 23(2): 227-232.

125. Ireland, D; Harrall, R; Curry, V; Holloway, G; Hackney, R; Hazleman, B, and Riley, G. (2001) Multiple changes in gene expression in chronic human Achilles tendinopathy. Matrix Biol 20(3): 159-169.

126. Jain, N and Dawson, MJ. (2007) Bilateral spontaneous and simultaneous rupture of the Achilles tendons secondary to exercises for limb ischemia: a case report. J Vasc Nurs 25(2): 26-28.

127. James, SL; Bates, BT, and Osternig, LR. (1978) Injuries to runners. Am J Sports Med 6(2): 40-50.

128. Jarvinen, M; Jozsa, L; Kannus, P; Jarvinen, TL; Kvist, M, and Leadbetter, W. (1997) Histopathological findings in chronic tendon disorders. Scand J Med Sci Sports 7(2): 86-95.

129. Jarvinen, TA; Jozsa, L; Kannus, P; Jarvinen, TL; Hurme, T; Kvist, M; Pelto-Huikko, M; Kalimo, H, and Jarvinen, M. (2003) Mechanical loading regulates the expression of tenascin-C in the myotendinous junction and tendon but does not induce de novo synthesis in the skeletal muscle. J Cell Sci 116(Pt 5): 857-866.

130. Jarvinen, TA; Jozsa, L; Kannus, P; Jarvinen, TL; Kvist, M; Hurme, T; Isola, J; Kalimo, H, and Jarvinen, M. (1999) Mechanical loading regulates tenascin-C expression in the osteotendinous junction. J Cell Sci 112 Pt 18:3157-3166.

131. Jarvinen, TA; Kannus, P; Maffulli, N, and Khan, KM. (2005) Achilles tendon disorders: etiology and epidemiology. Foot Ankle Clin 10(2): 255-266.

132. Jarvinen, TA; Kannus, P; Paavola, M; Jarvinen, TL; Jozsa, L, and Jarvinen, M. (2001) Achilles tendon injuries. Curr Opin Rheumatol 13(2): 150-155.

133. Jelinsky, SA; Rodeo, SA; Li, J; Gulotta, LV; Archambault, JM, and Seeherman, HJ. (2011) Regulation of gene expression in human tendinopathy. BMC Musculoskelet Disord 12(86): doi:10.1186/1471-2474-12-86

134. Jeon, MJ; Chung, SM; Choi, JR; Jung, HJ; Kim, SK, and Bai, SW. (2009) The relationship between COL3A1 exon 31 polymorphism and pelvic organ prolapse. J Urol 181(3): 1213-1216.

135. Jones, FS; Hoffman, S; Cunningham, BA, and Edelman, GM. (1989) A detailed structural model of cytotactin: protein homologies, alternative RNA splicing, and binding regions. Proc Natl Acad Sci U S A 86(6): 1905-1909.

University of Cape Town

136. Jones, FS and Jones, PL. (2000) The tenascin family of ECM glycoproteins: structure, function, and regulation during embryonic development and tissue remodeling. Dev Dyn 218(2): 235-259.

137. Jones, PL and Jones, FS. (2000) Tenascin-C in development and disease: gene regulation and cell function. Matrix Biol 19(7): 581-596.

138. Jozsa, L; Balint, J; Kannus, P; Jarvinen, M, and Lehto, M. (1993) Mechanoreceptors in human myotendinous junction. Muscle Nerve 16(5): 453-457.

139. Jozsa, L; Balint, JB; Kannus, P; Reffy, A, and Barzo, M. (1989) Distribution of blood groups in patients with tendon rupture. An analysis of 832 cases. J Bone Joint Surg Br 71(2): 272-274.

140. Jozsa, L and Kannus, P. (1997) Histopathological findings in spontaneous tendon ruptures. Scand J Med Sci Sports 7(2): 113-118.

141. Kader, D; Saxena, A; Movin, T, and Maffulli, N. (2002) Achilles tendinopathy: some aspects of basic science and clinical management. Br J Sports Med 36(4): 239-249.

142. Kadler, KE; Baldock, C; Bella, J, and Boot-Handford, RP. (2007) Collagens at a glance. Journal of Cell Science 120(12): 1955-1958.

143. Kahai, S; Vary, CP; Gao, Y, and Seth, A. (2004) Collagen, type V, alpha1 (COL5A1) is regulated by TGF-beta in osteoblasts. Matrix Biol 23(7): 445-455.

144. Kannu, P; Bateman, JF; Randle, S; Cowie, S; du, SD; McGrath, S; Edwards, M, and Savarirayan, R. (2010) Premature arthritis is a distinct type II collagen phenotype. Arthritis Rheum 62(5): 1421-1430.

145. Kannus, P. (1997) Etiology and pathophysiology of chronic tendon disorders in sports. Scand J Med Sci Sports 7(2): 78-85.

146. Kannus, P and Natri, A. (1997) Etiology and pathophysiology of tendon ruptures in sports. Scand J Med Sci Sports 7(2): 107-112.

147. Kaufman, KR; Brodine, SK; Shaffer, RA; Johnson, CW, and Cullison, TR. (1999) The effect of foot structure and range of motion on musculoskeletal overuse injuries. Am J Sports Med 27(5): 585-593.

148. Khaliq, Y and Zhanel, GG. (2003) Fluoroquinolone-associated tendinopathy: a critical review of the literature. Clin Infect Dis 36(11): 1404-1410.

University of Cape Town

149. Khan, KM; Cook, JL; Bonar, F; Harcourt, P, and Astrom, M. (1999) Histopathology of common tendinopathies. Update and implications for clinical management. Sports Med 27(6): 393-408.

150. Khan, KM; Cook, JL; Maffulli, N, and Kannus, P. (2000) Where is the pain coming from in tendinopathy? It may be biochemical, not only structural, in origin. Br J Sports Med 34(2): 81-83.

151. Khan, KM and Scott, A. (2009) Mechanotherapy: how physical therapists' prescription of exercise promotes tissue repair. Br J Sports Med 43(4): 247-252.

152. Khoschnau, S; Melhus, H; Jacobson, A; Rahme, H; Bengtsson, H; Ribom, E; Grundberg, E; Mallmin, H, and Michaelsson, K. (2008) Type I collagen alpha1 Sp1 polymorphism and the risk of cruciate ligament ruptures or shoulder dislocations. Am J Sports Med 36(12): 2432-2436.

153. Kjaer, M and Hansen, M. (2008) The mystery of female connective tissue. J Appl Physiol 105(4): 1026-1027.

154. Kjaer, M; Langberg, H; Miller, BF; Boushel, R; Crameri, R; Koskinen, S; Heinemeier, K; Olesen, JL; Dossing, S; Hansen, M; Pedersen, SG; Rennie, MJ, and Magnusson, P. (2005) Metabolic activity and collagen turnover in human tendon in response to physical activity. J Musculoskelet Neuronal Interact 5(1): 41-52.

155. Klemp, P; Halland, AM; Majoos, FL, and Steyn, K. (1993) Musculoskeletal manifestations in hyperlipidaemia: a controlled study. Ann Rheum Dis 52(1): 44-48.

156. Kluivers, KB; Dijkstra, JR; Hendriks, JC; Lince, SL; Vierhout, ME, and van Kempen, LC. (2009) COL3A1 2209G>A is a predictor of pelvic organ prolapse. Int Urogynecol J Pelvic Floor Dysfunct 20(9): 1113-1118.

157. Kraemer, R; Wuerfel, W; Lorenzen, J; Busche, M; Vogt, PM, and Knobloch, K. (2012) Analysis of hereditary and medical risk factors in Achilles tendinopathy and Achilles tendon ruptures: a matched pair analysis. Archives of Orthopaedic and Trauma Surgery 132(6): 847-853.

158. Kujala, UM; Jarvinen, M; Natri, A; Lehto, M; Nelimarkka, O; Hurme, M; Virta, L, and Finne, J. (1992) ABO blood groups and musculoskeletal injuries. Injury 23(2): 131-133.

159. Kujala, UM; Sarna, S, and Kaprio, J. (2005) Cumulative incidence of achilles tendon rupture and tendinopathy in male former elite athletes. Clin J Sport Med 15(3): 133-135.

University of Cape Town

160. Kumar, P; Henikoff, S, and Ng, PC. (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 4(7): 1073-1081.

161. Kvansakul, M; Adams, JC, and Hohenester, E. (2004) Structure of a thrombospondin C-terminal fragment reveals a novel calcium core in the type 3 repeats. EMBO J 23(6): 1223-1233.

162. Kvist, M. (1994) Achilles tendon injuries in athletes. Sports Med 18(3): 173-201.

163. Kyriakides, TR; Zhu, YH; Smith, LT; Bain, SD; Yang, Z; Lin, MT; Danielson, KG; Iozzo, RV; LaMarca, M; McKinney, CE; Ginns, EI, and Bornstein, P. (1998) Mice that lack thrombospondin 2 display connective tissue abnormalities that are associated with disordered collagen fibrillogenesis, an increased vascular density, and a bleeding diathesis. J Cell Biol 140(2): 419-430.

164. Laguette, MJ; Abrahams, Y; Prince, S, and Collins, M. (2011) Sequence variants within the 3'-UTR of the COL5A1 gene alters mRNA stability: implications for musculoskeletal soft tissue injuries. Matrix Biol 30(5-6): 338-345.

165. Lahiri, K and Nurnberger, JI. (1991) A rapid non-enzymatic method for the preperation of HMW DNA from blood for RFLP studies. Nucleic Acids Res 19:5444

166. Langdahl, BL; Lokke, E; Carstens, M; Stenkjaer, LL, and Eriksen, EF. (2000) Osteoporotic fractures are associated with an 86-base pair repeat polymorphism in the interleukin-1--receptor antagonist gene but not with polymorphisms in the interleukin-1beta gene. J Bone Miner Res 15(3): 402-414.

167. Laseter, JT and Russell, JA. (1991) Anabolic steroid-induced tendon pathology: a review of the literature. Med Sci Sports Exerc 23(1): 1-3.

168. Leal, SM. (2005) Detection of genotyping errors and pseudo-SNPs via deviations from Hardy-Weinberg equilibrium. Genet Epidemiol 29(3): 204-214.

169. Leppilahti, J and Orava, S. (1998) Total Achilles tendon rupture. A review. Sports Med 25(2): 79-100.

170. Leppilahti, J; Puranen, J, and Orava, S. (1996) ABO blood group and Achilles tendon rupture. Ann Chir Gynaecol 85(4): 369-371.

171. Liow, RY and Tavares, S. (1995) Bilateral rupture of the quadriceps tendon associated with anabolic steroids. Br J Sports Med 29(2): 77-79.

University of Cape Town

172. Litt, JZ. (2000) Fluoroquinolones. Southern Medical Journal 93(5): 525-526.

173. Liu, SH; Yang, RS; al Shaikh, R, and Lane, JM. (1995) Collagen in tendon, ligament, and bone healing. A current review. Clin Orthop Relat Res 318: 265-278.

174. Liu, Y; Ramanath, HS, and Wang, DA. (2008) Tendon tissue engineering using scaffold enhancing strategies. Trends Biotechnol 26(4): 201-209.

175. Lohmueller, KE; Pearce, CL; Pike, M; Lander, ES, and Hirschhorn, JN. (2003) Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet 33(2): 177-182.

176. Longo, UG; Oliva, F; Denaro, V, and Maffulli, N. (2008) Oxygen species and overuse tendinopathy in athletes. Disabil Rehabil 30(20-22): 1563-1571.

177. Maddox, BK; Mokashi, A; Keene, DR, and Bachinger, HP. (2000) A cartilage oligomeric matrix protein mutation associated with pseudoachondroplasia changes the structural and functional properties of the type 3 domain. J Biol Chem 275(15): 11412-11417.

178. Maffulli, N; Khan, KM, and Puddu, G. (1998) Overuse tendon conditions: time to change a confusing terminology. Arthroscopy 14(8): 840-843.

179. Maffulli, N; Reaper, JA; Waterston, SW, and Ahya, R. (2000) ABO blood groups and Achilles tendon rupture in the Grampian Region of Scotland. Clinical Journal of Sport Medicine 10(4): 269-271.

180. Maffulli, N; Waterston, SW; Squair, J; Reaper, J, and Douglas, AS. (1999) Changing incidence of Achilles tendon rupture in Scotland: a 15-year study. Clin J Sport Med 9(3): 157-160.

181. Magnusson, SP; Hansen, M; Langberg, H; Miller, B; Haraldsson, B; Westh, EK; Koskinen, S; Aagaard, P, and Kjaer, M. (2007) The adaptability of tendon to loading differs in men and women. Int J Exp Pathol 88(4): 237-240.

182. Mahieu, NN; Witvrouw, E; Stevens, V; Van Tiggelen, D, and Roget, P. (2006) Intrinsic risk factors for the development of achilles tendon overuse injury: a prospective study. Am J Sports Med 34(2): 226-235.

183. Malfait, F; Coucke, P; Symoens, S; Loeys, B; Nuytinck, L, and De Paepe, A. (2005) The molecular basis of classic Ehlers-Danlos syndrome: a comprehensive study of biochemical and molecular findings in 48 unrelated patients. Hum Mutat 25(1): 28-37.

University of Cape Town

184. Malfait, F; Wenstrup, RJ, and De Paepe, A. (2010) Clinical and genetic aspects of Ehlers-Danlos syndrome, classic type. Genet Med 12(10): 597-605.

185. Malliaras, P; Cook, JL, and Kent, P. (2006) Reduced ankle dorsiflexion range may increase the risk of patellar tendon injury among volleyball players. J Sci Med Sport 9(4): 304-309.

186. Manolio, TA; Collins, FS; Cox, NJ; Goldstein, DB; Hindorff, LA; Hunter, DJ; McCarthy, MI; Ramos, EM; Cardon, LR; Chakravarti, A; Cho, JH; Guttmacher, AE; Kong, A; Kruglyak, L; Mardis, E; Rotimi, CN; Slatkin, M; Valle, D; Whittemore, AS; Boehnke, M; Clark, AG; Eichler, EE; Gibson, G; Haines, JL; Mackay, TF; McCarroll, SA, and Visscher, PM. (2009) Finding the missing heritability of complex diseases. Nature 461(7265): 747-753.

187. Marie, I; Delafenetre, H; Massy, N; Thuillez, C, and Noblet, C. (2008) Tendinous disorders attributed to statins: a study on ninety-six spontaneous reports in the period 1990-2005 and review of the literature. Arthritis Rheum 59(3): 367-372.

188. Marqueti, RC; Parizotto, NA; Chriguer, RS; Perez, SE, and Selistre-de-Araujo, HS. (2006) Androgenic-anabolic steroids associated with mechanical loading inhibit matrix metallopeptidase activity and affect the remodeling of the achilles tendon in rats. Am J Sports Med 34(8): 1274-1280.

189. Marr, CM; McMillan, I; Boyd, JS; Wright, NG, and Murray, M. (1993) Ultrasonographic and histopathological findings in equine superficial digital flexor tendon injury. Equine Vet J 25(1): 23-29.

190. Matsuda, A; Hirota, T; Akahoshi, M; Shimizu, M; Tamari, M; Miyatake, A; Takahashi, A; Nakashima, K; Takahashi, N; Obara, K; Yuyama, N; Doi, S; Kamogawa, Y; Enomoto, T; Ohshima, K; Tsunoda, T; Miyatake, S; Fujita, K; Kusakabe, M; Izuhara, K; Nakamura, Y; Hopkin, J, and Shirakawa, T. (2005) Coding SNP in tenascin-C Fn-III-D domain associates with adult asthma. Hum Mol Genet 14(19): 2779-2786.

191. McCrory, JL; Martin, DF; Lowery, RB; Cannon, DW; Curl, WW; Read, HM, Jr.; Hunter, DM; Craven, T, and Messier, SP. (1999) Etiologic factors associated with Achilles tendinitis in runners. Med Sci Sports Exerc 31(10): 1374-1381.

192. Metcalfe, D; Achten, J, and Costa, ML. (2009) Glucocorticoid injections in lesions of the achilles tendon. Foot Ankle Int 30(7): 661-665.

193. Michalickova, K; Susic, M; Willing, MC; Wenstrup, RJ, and Cole, WG. (1998) Mutations of the alpha2(V) chain of type V collagen impair matrix assembly and produce ehlers-danlos syndrome type I. Hum Mol Genet 7(2): 249-255.

University of Cape Town

194. Midwood, K; Sacre, S; Piccinini, AM; Inglis, J; Trebaul, A; Chan, E; Drexler, S; Sofat, N; Kashiwagi, M; Orend, G; Brennan, F, and Foxwell, B. (2009) Tenascin-C is an endogenous activator of Toll-like receptor 4 that is essential for maintaining inflammation in arthritic joint disease. Nat Med 15(7): 774-780.

195. Milgrom, C; Finestone, A; Zin, D; Mandel, D, and Novack, V. (2003) Cold weather training: a risk factor for Achilles paratendinitis among recruits. Foot Ankle Int 24(5): 398-401.

196. Millar, NL; Wei, AQ; Molloy, TJ; Bonar, F, and Murrell, GA. (2009) Cytokines and apoptosis in supraspinatus tendinopathy. J Bone Joint Surg Br 91(3): 417-424.

197. Mokone, GC; Gajjar, M; September, AV; Schwellnus, MP; Greenburg, J; Noakes, TD, and Collins, M. (2005) The Guanine-Thymine dinucleotide repeat polymorphism within the Tenascin-C gene is associated with Achilles tendon injuries. Am J Sports Med 33(8): 1016-21

198. Mokone, GG. (2006) Risk factors for Achilles tendon injuries: An emphasis on the identification of specific genetic factors. PhD Thesis. University of Cape Town

199. Mokone, GG; Schwellnus, MP; Noakes, TD, and Collins, M. (2006) The COL5A1 gene and Achilles tendon pathology. Scand J Med Sci Sports 16(1): 19-26.

200. Morrow, JR, Jr.; Defina, LF; Leonard, D; Trudelle-Jackson, E, and Custodio, MA. (2012) Meeting physical activity guidelines and musculoskeletal injury: the WIN study. Med Sci Sports Exerc 44(10): 1986-1992.

201. Mountcastle, SB; Posner, M; Kragh, JF, Jr., and Taylor, DC. (2007) Gender differences in anterior cruciate ligament injury vary with activity: epidemiology of anterior cruciate ligament injuries in a young, athletic population. Am J Sports Med 35(10): 1635-1642.

202. Munnoch, K and Bridger, RS. (2007) Smoking and injury in Royal Marines' training. Occup Med (Lond) 57(3): 214-216.

203. Munteanu, SE and Barton, CJ. (2011) Lower limb biomechanics during running in individuals with achilles tendinopathy: a systematic review. Journal of Foot and Ankle Research 4 DOI: 10.1186/1757-1146-4-15

204. Neidhardt, J; Fehr, S; Kutsche, M; Lohler, J, and Schachner, M. (2003) Tenascin-N: characterization of a novel member of the tenascin family that mediates neurite repulsion from hippocampal explants. Mol Cell Neurosci 23(2): 193-209.

University of Cape Town

205. Nell, EM; van der Merwe, L; Cook, J; Handley, CJ; Collins, M, and September, AV. (2012) The apoptosis pathway and the genetic predisposition to Achilles tendinopathy. J Orthop Res 30(11): 1719-24

206. Nichols, AW. (2005) Complications associated with the use of corticosteroids in the treatment of athletic injuries. Clin J Sport Med 15(5): 370-375.

207. Nyholt, DR. (2004) A Simple Correction for Multiple Testing for Single-Nucleotide Polymorphisms in Linkage Disequilibrium with Each Other. Am J Hum Genet 74(4): 765-769.

208. O'Brien, M. (1997) Structure and metabolism of tendons. Scand J Med Sci Sports 7:55-61.

209. O'Connell, K; Posthumus, M; Schwellnus, MP, and Collins, M. (2012) Collagen Genes and Exercise-Associated Muscle Cramping. Clin J Sport Med 23(1): 64-9

210. Oberhauser, AF; Marszalek, PE; Erickson, HP, and Fernandez, JM. (1998) The molecular elasticity of the extracellular matrix protein tenascin. Nature 393(6681): 181-185.

211. Orsmark-Pietras, C; Melen, E; Vendelin, J; Bruce, S; Laitinen, A; Laitinen, LA; Lauener, R; Riedler, J; von Mutius, E; Doekes, G; Wickman, M; van Hage, M; Pershagen, G; Scheynius, A; Nyberg, F, and Kere, J. (2008) Biological and genetic interaction between tenascin C and neuropeptide S receptor 1 in allergic diseases. Hum Mol Genet 17(11): 1673-1682.

212. Ovcharenko, I; Nobrega, MA; Loots, GG, and Stubbs, L. (2004) ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. Nucleic Acids Res 32(Web Server issue): W280-W286.

213. Ozaras, R; Mert, A; Tahan, V; Uraz, S; Ozaydin, I; Yilmaz, MH, and Ozaras, N. (2003) Ciprofloxacin and Achilles' tendon rupture: a causal relationship. Clin Rheumatol 22(6): 500-501.

214. Ozgurtas, T; Yildiz, C; Serdar, M; Atesalp, S, and Kutluay, T. (2003) Is high concentration of serum lipids a risk factor for Achilles tendon rupture? Clin Chim Acta 331(1-2): 25-28.

215. Paavola, M; Kannus, P; Paakkala, T; Pasanen, M, and Jarvinen, M. (2000) Long-term prognosis of patients with achilles tendinopathy. An observational 8-year follow-up study. Am J Sports Med 28(5): 634-642.

University of Cape Town

216. Pace, JM; Corrado, M; Missero, C, and Byers, PH. (2003) Identification, characterization and expression analysis of a new fibrillar collagen gene, COL27A1. Matrix Biol 22(1): 3-14.

217. Parkinson, J; Samiric, T; Ilic, MZ; Cook, J, and Handley, CJ. (2011) Involvement of proteoglycans in tendinopathy. J Musculoskelet Neuronal Interact 11(2): 86-93.

218. Parssinen, M; Karila, T; Kovanen, V, and Seppala, T. (2000) The effect of supraphysiological doses of anabolic androgenic steroids on collagen metabolism. Int J Sports Med 21(6): 406-411.

219. Pearce, CJ; Ismail, M, and Calder, JD. (2009) Is apoptosis the cause of noninsertional achilles tendinopathy? Am J Sports Med 37(12): 2440-2444.

220. Perneger, TV. (1998) What's wrong with Bonferroni adjustments. BMJ 316(7139): 1236-1238.

221. Pingel, J; Fredberg, U; Qvortrup, K; O Larsen, J; Schjerling, P; Heinemeier, K; Kjaer, M, and Langberg, H. (2012) Local biochemical and morphological differences in human Achilles tendinopathy: a case control study. BMC Musculoskelet Disord 13:53

222. Pollitt, R; McMahon, R; Nunn, J; Bamford, R; Afifi, A; Bishop, N, and Dalton, A. (2006) Mutation analysis of COL1A1 and COL1A2 in patients diagnosed with osteogenesis imperfecta type I-IV. Hum Mutat 27(7): 716

223. Posey, KL; Yang, Y; Veerisetty, AC; Sharan, SK, and Hecht, JT. (2008) Model systems for studying skeletal dysplasias caused by TSP-5/COMP mutations. Cell Mol Life Sci 65(5): 687-699.

224. Posthumus, M; Collins, M; Cook, J; Handley, CJ; Ribbans, WJ; Smith, RKW; Schwellnus, MP, and Raleigh, SM. (2010) Components of the transforming growth factor-beta family and the pathogenesis of human Achilles tendon pathology-a genetic association study. Rheumatology 49(11): 2090-2097.

225. Posthumus, M; September, AV; Keegan, M; O'Cuinneagain, D; van der Merwe, W; Schwellnus, MP, and Collins, M. (2009) Genetic risk factors for anterior cruciate ligament ruptures: COL1A1 gene variant. Br J Sports Med 43(5): 352-356.

226. Posthumus, M; September, AV; O'Cuinneagain, D; van der Merwe, W; Schwellnus, MP, and Collins, M. (2009) The association between the COL12A1 gene and anterior cruciate ligament ruptures. Br J Sports Med 44 (16): 1160-5

227. Posthumus, M; September, AV; O'Cuinneagain, D; van der Merwe, W; Schwellnus, MP, and Collins, M. (2009) The COL5A1 Gene Is Associated With Increased Risk

University of Cape Town

of Anterior Cruciate Ligament Ruptures in Female Participants. Am J Sports Med 37(11): 2234-40

228. Posthumus, M; September, AV; Schwellnus, MP, and Collins, M. (2008) Investigation of the Sp1-binding site polymorphism within the COL1A1 gene in participants with Achilles tendon injuries and controls. J Sci Med Sport 12(1): 184-9

229. Pullatt, RC; Gadarla, MR; Karas, RH; Alsheikh-Ali, AA, and Thompson, PD. (2007) Tendon rupture associated with simvastatin/ezetimibe therapy. Am J Cardiol 100(1): 152-153.

230. R Development Core Team. (2010) R: a language and environment for statistical computing. R Foundation for Statistical Computing www.r-project.org

231. Raleigh, SM; van der Merwe, L; Ribbans, WJ; Smith, RK; Schwellnus, MP, and Collins, M. (2009) Variants within the MMP3 gene are associated with Achilles tendinopathy: possible interaction with the COL5A1 gene. Br J Sports Med 43(7): 514-520.

232. Rees, JD; Stride, M, and Scott, A. (2013) Tendons - time to revisit inflammation. Br J Sports Med. [Epub ahead of print]

233. Rees, JD; Wilson, AM, and Wolman, RL. (2006) Current concepts in the management of tendon disorders. Rheumatology (Oxford) 45(5): 508-521.

234. Rees, SG; Dent, CM, and Caterson, B. (2009) Metabolism of proteoglycans in tendon. Scand J Med Sci Sports 19(4): 470-478.

235. Reule, CA; Alt, WW; Lohrer, H, and Hochwald, H. (2011) Spatial orientation of the subtalar joint axis is different in subjects with and without Achilles tendon disorders. British Journal of Sports Medicine 45(13): 1029-1034.

236. Ribbans, WJ and Collins, M. (2013) Achilles Tendon Problems – Do our Genes Contribute? Journal of Bone and Joint Surgery - British Volume 95-B(3): 305-313.

237. Ricard-Blum, S and Ruggiero, F. (2005) The collagen superfamily: from the extracellular matrix to the cell membrane. Pathol Biol (Paris) 53(7): 430-442.

238. Richards, AJ; Martin, S; Nicholls, AC; Harrison, JB; Pope, FM, and Burrows, NP. (1998) A single base mutation in COL5A2 causes Ehlers-Danlos syndrome type II. J Med Genet 35(10): 846-848.

University of Cape Town

239. Riley, G. (2004) The pathogenesis of tendinopathy. A molecular perspective. Rheumatology (Oxford) 43(2): 131-142.

240. Riley, G. (2005) Chronic tendon pathology: molecular basis and therapeutic implications. Expert Rev Mol Med 7(5): 1-25.

241. Rixe, JA; Gallo, RA, and Silvis, ML. (2012) The barefoot debate: can minimalist shoes reduce running-related injuries? Curr Sports Med Rep 11(3): 160-165.

242. Roach, HI; Yamada, N; Cheung, KS; Tilley, S; Clarke, NM; Oreffo, RO; Kokubun, S, and Bronner, F. (2005) Association between the abnormal expression of matrix-degrading enzymes by human osteoarthritic chondrocytes and demethylation of specific CpG sites in the promoter regions. Arthritis Rheum 52(10): 3110-3124.

243. Rosenberg, K; Olsson, H; Morgelin, M, and Heinegard, D. (1998) Cartilage oligomeric matrix protein shows high affinity zinc-dependent interaction with triple helical collagen. J Biol Chem 273(32): 20397-20403.

244. Rosenbloom, KB. (2011) Pathology-designed custom molded foot orthoses. Clin Podiatr Med Surg 28(1): 171-187.

245. Rowson, S; McNally, C, and Duma, SM. (2010) Can footwear affect achilles tendon loading? Clin J Sport Med 20(5): 344-349.

246. Ryan, M; Grau, S; Krauss, I; Maiwald, C; Taunton, J, and Horstmann, T. (2009) Kinematic analysis of runners with achilles mid-portion tendinopathy. Foot Ankle Int 30(12): 1190-1195.

247. Saunders, CJ; van der Merwe, L; Posthumus, M; Cook, J; Handley, CJ; Collins, M, and September, AV. (2013) Investigation of variants within the COL27A1 and TNC genes and Achilles tendinopathy in two populations. J Orthop Res 31(4): 632-637.

248. Sawaya, GF; Guirguis-Blake, J; LeFevre, M; Harris, R, and Petitti, D. (2007) Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. Ann Intern Med 147(12): 871-875.

249. Schaid, DJ; Rowland, CM; Tines, DE; Jacobson, RM, and Poland, GA. (2002) Score tests for association between traits and haplotypes when linkage phase is ambiguous. American Journal of Human Genetics 70(2): 425-434.

250. Schepsis, AA; Jones, H, and Haas, AL. (2002) Achilles tendon disorders in athletes. Am J Sports Med 30(2): 287-305.

University of Cape Town

251. Schizas, N; Weiss, R; Lian, O; Frihagen, F; Bahr, R, and Ackermann, PW. (2012) Glutamate receptors in tendinopathic patients. J Orthop Res 30(9): 1447-1452.

252. Scott, A; Huisman, E, and Khan, K. (2011) Conservative treatment of chronic Achilles tendinopathy. CMAJ 183(10): 1159-1165.

253. Scott, A; Khan, KM; Heer, J; Cook, JL; Lian, O, and Duronio, V. (2005) High strain mechanical loading rapidly induces tendon apoptosis: an ex vivo rat tibialis anterior model. Br J Sports Med 39(5): e25

254. Scutt, N; Rolf, CG, and Scutt, A. (2006) Glucocorticoids inhibit tenocyte proliferation and Tendon progenitor cell recruitment. J Orthop Res 24(2): 173-182.

255. September, AV; Cook, J; Handley, CJ; van der Merwe, L; Schwellnus, MP, and Collins, M. (2009) Variants within the COL5A1 gene are associated with Achilles tendinopathy in two populations. Br J Sports Med 43(5): 357-365.

256. September, AV; Mokone, GG; Schwellnus, MP, and Collins, M. (2006) Genetic risk factors for Achilles tendon injuries. International SportMed journal 7(3): 201-215.

257. September, AV; Nell, EM; O'Connell, K; Cook, J; Handley, CJ; van der Merwe, L; Schwellnus, M, and Collins, M. (2011) A pathway-based approach investigating the genes encoding interleukin-1beta, interleukin-6 and the interleukin-1 receptor antagonist provides new insight into the genetic susceptibility of Achilles tendinopathy. Br J Sports Med 45(13): 1040-1047.

258. September, AV; Posthumus, M, and Collins, M. (2012) Application of genomics in the prevention, treatment and management of Achilles tendinopathy and anterior cruciate ligament ruptures. Recent Pat DNA Gene Seq 6(3): 216-23

259. September, AV; Posthumus, M; van der Merwe, L; Schwellnus, M; Noakes, TD, and Collins, M. (2008) The COL12A1 and COL14A1 genes and Achilles tendon injuries. Int J Sports Med 29(3): 257-263.

260. September, AV; Schwellnus, MP, and Collins, M. (2007) Tendon and ligament injuries: the genetic component. Br J Sports Med 41(4): 241-246.

261. Sharma, P and Maffulli, N. (2005) Tendon injury and tendinopathy: Healing and repair. J Bone Joint Surg 87A(1): 187-202.

262. Shin, JI and Lee, JS. (2009) Tendinopathy and statin use: the role of matrix metalloproteinases or eicosanoids? Comment on the letter by Beri and Khattri and the article by Marie et al. Arthritis Rheum 61(2): 286-7.

University of Cape Town

263. Shrier, I; Matheson, GO, and Kohl, HW, III. (1996) Achilles tendonitis: are corticosteroid injections useful or harmful? Clin J Sport Med 6(4): 245-250.

264. Shukla, DD. (2002) Bilateral spontaneous rupture of achilles tendon secondary to limb ischemia: a case report. J Foot Ankle Surg 41(5): 328-329.

265. Sicot, FX; Exposito, JY; Masselot, M; Garrone, R; Deutsch, J, and Gaill, F. (1997) Cloning of an annelid fibrillar-collagen gene and phylogenetic analysis of vertebrate and invertebrate collagens. Eur J Biochem 246(1): 50-58.

266. Sinnwell, JP and Schaid, DJ. (2011) Haplo.stats: a package for statistical analysis of haplotypes with traits and covariates when linkage phase is ambiguous. R package(Version 1.5.4): http://cran.r-project.org/package=haplo.stats

267. Slobogean, G and Bhandari, M. (2012) Introducing levels of evidence to the Journal of Orthopaedic Trauma: implementation and future directions. J Orthop Trauma 26(3): 127-128.

268. Smith, LT; Schwarze, U; Goldstein, J, and Byers, PH. (1997) Mutations in the COL3A1 gene result in the Ehlers-Danlos syndrome type IV and alterations in the size and distribution of the major collagen fibrils of the dermis. J Invest Dermatol 108(3): 241-247.

269. Smith, RK; Birch, HL; Goodman, S; Heinegard, D, and Goodship, AE. (2002) The influence of ageing and exercise on tendon growth and degeneration--hypotheses for the initiation and prevention of strain-induced tendinopathies. Comp Biochem Physiol A Mol Integr Physiol 133(4): 1039-1050.

270. Smith, RK; Gerard, M; Dowling, B; Dart, AJ; Birch, HL, and Goodship, AE. (2002) Correlation of cartilage oligomeric matrix protein (COMP) levels in equine tendon with mechanical properties: a proposed role for COMP in determining function-specific mechanical characteristics of locomotor tendons. Equine Vet J Suppl 34): 241-244.

271. Smith, RK; Zunino, L; Webbon, PM, and Heinegard, D. (1997) The distribution of cartilage oligomeric matrix protein (COMP) in tendon and its variation with tendon site, age and load. Matrix Biol 16(5): 255-271.

272. Sodersten, F; Hultenby, K; Heinegard, D; Johnston, C, and Ekman, S. (2012) Immunolocalization of collagens (I and III) and cartilage oligomeric matrix protein (COMP) in the normal and injured equine superficial digital flexor tendon. Connect Tissue Res 54(1): 62-9

273. Speed, CA. (2001) Fortnightly review: Corticosteroid injections in tendon lesions. BMJ 323(7309): 382-386.

University of Cape Town

274. Stone, D; Green, C; Rao, U; Aizawa, H; Yamaji, T; Niyibizi, C; Carlin, G, and Woo, SL. (1999) Cytokine-induced tendinitis: a preliminary study in rabbits. J Orthop Res 17(2): 168-177.

275. Suchak, AA; Bostick, G; Reid, D; Blitz, S, and Jomha, N. (2005) The incidence of Achilles tendon ruptures in Edmonton, Canada. Foot Ankle Int 26(11): 932-936.

276. Sullo, A; Maffulli, N; Capasso, G, and Testa, V. (2001) The effects of prolonged peritendinous administration of PGE1 to the rat Achilles tendon: a possible animal model of chronic Achilles tendinopathy. J Orthop Sci 6(4): 349-357.

277. Swingler, TE; Wheeler, G; Carmont, V; Elliott, HR; Barter, MJ; Abu-Elmagd, M; Donell, ST; Boot-Handford, RP; Hajihosseini, MK; Munsterberg, A; Dalmay, T; Young, DA, and Clark, IM. (2012) The expression and function of microRNAs in chondrogenesis and osteoarthritis. Arthritis Rheum 64(6): 1909-1919.

278. Tessutti, V; Ribeiro, AP; Trombini-Souza, F, and Sacco, IC. (2012) Attenuation of foot pressure during running on four different surfaces: asphalt, concrete, rubber, and natural grass. J Sports Sci 30(14): 1545-1550.

279. Thampatty, BP; Li, H; Im, HJ, and Wang, JH. (2007) EP4 receptor regulates collagen type-I, MMP-1, and MMP-3 gene expression in human tendon fibroblasts in response to IL-1 beta treatment. Gene 386(1-2): 154-161.

280. Theobald, P; Benjamin, M; Nokes, L, and Pugh, N. (2005) Review of the vascularisation of the human Achilles tendon. Injury 36(11): 1267-1272.

281. Tsuzaki, M; Guyton, G; Garrett, W; Archambault, JM; Herzog, W; Almekinders, L; Bynum, D; Yang, X, and Banes, AJ. (2003) IL-1 beta induces COX2, MMP-1, -3 and -13, ADAMTS-4, IL-1 beta and IL-6 in human tendon cells. J Orthop Res 21(2): 256-264.

282. Tuite, DJ; Renstrom, PA, and O'Brien, M. (1997) The aging tendon. Scand J Med Sci Sports 7(2): 72-77.

283. University of Washington. (2012) Genome Variation Server. http://gvs.gs.washington.edu/GVS137/

284. Valdes, AM; Loughlin, J; Oene, MV; Chapman, K; Surdulescu, GL; Doherty, M, and Spector, TD. (2007) Sex and ethnic differences in the association of ASPN, CALM1, COL2A1, COMP, and FRZB with genetic susceptibility to osteoarthritis of the knee. Arthritis Rheum 56(1): 137-146.

285. Valkkila, M; Melkoniemi, M; Kvist, L; Kuivaniemi, H; Tromp, G, and Ala-Kokko, L. (2001) Genomic organization of the human COL3A1 and COL5A2 genes:

University of Cape Town

COL5A2 has evolved differently than the other minor fibrillar collagen genes. Matrix Biol 20(5-6): 357-366.

286. van der Linden, PD; Sturkenboom, MC; Herings, RM; Leufkens, HG, and Stricker, BH. (2002) Fluoroquinolones and risk of Achilles tendon disorders: case-control study. BMJ 324(7349): 1306-1307.

287. van der Worp, H; van, AM; Roerink, S; Pepping, GJ; van, dA-S, I, and Zwerver, J. (2011) Risk factors for patellar tendinopathy: a systematic review of the literature. Br J Sports Med 45(5): 446-452.

288. van Dijk, CN; van Sterkenburg, MN; Wiegerinck, JI; Karlsson, J, and Maffulli, N. (2011) Terminology for Achilles tendon related disorders. Knee Surg Sports Traumatol Arthrosc 19(5): 835-841.

289. Van Ginckel, A; Thijs, Y; Hesar, NGZ; Mahieu, N; De Clercq, D; Roosen, P, and Witvrouw, E. (2009) Intrinsic gait-related risk factors for Achilles tendinopathy in novice runners: A prospective study. Gait & Posture 29(3): 387-391.

290. Verrecchia, F; Chu, ML, and Mauviel, A. (2001) Identification of novel TGF-beta /Smad gene targets in dermal fibroblasts using a combined cDNA microarray/promoter transactivation approach. J Biol Chem 276(20): 17058-17062.

291. Villiger, PM; Kusari, AB; ten, DP, and Lotz, M. (1993) IL-1 beta and IL-6 selectively induce transforming growth factor-beta isoforms in human articular chondrocytes. J Immunol 151(6): 3337-3344.

292. Visuri, T and Lindholm, H. (1994) Bilateral distal biceps tendon avulsions with use of anabolic steroids. Med Sci Sports Exerc 26(8): 941-944.

293. Waldecker, U; Hofmann, G, and Drewitz, S. (2012) Epidemiologic investigation of 1394 feet: coincidence of hindfoot malalignment and Achilles tendon disorders. Foot Ankle Surg 18(2): 119-123.

294. Wall, ME and Banes, AJ. (2005) Early responses to mechanical load in tendon: role for calcium signaling, gap junctions and intercellular communication. J Musculoskelet Neuronal Interact 5(1): 70-84.

295. Wang, L; Hong, Y; Li, JX, and Zhou, JH. (2012) Comparison of plantar loads during running on different overground surfaces. Res Sports Med 20(2): 75-85.

296. Warnes, G; Gorjanc, G; Leisch, F, and others. (2011) Genetics: a package for population genetics. R Package(Version 1.3.6): http://cran.r-project.org/package=genetics

University of Cape Town

297. Wenstrup, RJ; Smith, SM; Florer, JB; Zhang, G; Beason, DP; Seegmiller, RE; Soslowsky, LJ, and Birk, DE. (2011) Regulation of collagen fibril nucleation and initial fibril assembly involves coordinate interactions with collagens V and XI in developing tendon. J Biol Chem 286(23): 20455-20465.

298. Westh, E; Kongsgaard, M; Bojsen-Moller, J; Aagaard, P; Hansen, M; Kjaer, M, and Magnusson, SP. (2008) Effect of habitual exercise on the structural and mechanical properties of human tendon, in vivo, in men and women. Scand J Med Sci Sports 18(1): 23-30.

299. Whitting, JW; Steele, JR; McGhee, DE, and Munro, BJ. (2011) Dorsiflexion capacity affects achilles tendon loading during drop landings. Med Sci Sports Exerc 43(4): 706-713.

300. Williams, IF; McCullagh, KG; Goodship, AE, and Silver, IA. (1984) Studies on the pathogenesis of equine tendonitis following collagenase injury. Res Vet Sci 36(3): 326-338.

301. Williams, S; Hume, PA, and Kara, S. (2011) A review of football injuries on third and fourth generation artificial turfs compared with natural turf. Sports Med 41(11): 903-923.

302. Wilson, AM and Goodship, AE. (1994) Exercise-induced hyperthermia as a possible mechanism for tendon degeneration. J Biomech 27(7): 899-905.

303. Wise, BL; Peloquin, C; Choi, H; Lane, NE, and Zhang, Y. (2012) Impact of Age, Sex, Obesity, and Steroid Use on Quinolone-associated Tendon Disorders. Am J Med 125(12): 1228-1228.

304. Wren, TA; Lindsey, DP; Beaupre, GS, and Carter, DR. (2003) Effects of creep and cyclic loading on the mechanical properties and failure of human Achilles tendons. Ann Biomed Eng 31(6): 710-717.

305. Wyndow, N; Cowan, SM; Wrigley, TV, and Crossley, KM. (2012) Triceps surae activation is altered in male runners with Achilles tendinopathy. J Electromyogr Kinesiol 23(1): 166-72

306. Xu, Y and Murrell, GA. (2008) The basic science of tendinopathy. Clin Orthop Relat Res 466(7): 1528-1538.

307. Yamamoto, F; Cid, E; Yamamoto, M, and Blancher, A. (2012) ABO research in the modern era of genomics. Transfus Med Rev 26(2): 103-118.

University of Cape Town

308. Ye, S; Dhillon, S; Ke, X; Collins, AR, and Day, IN. (2001) An efficient procedure for genotyping single nucleotide polymorphisms. Nucleic Acids Res 29(17): E88-E88.

309. Yoon, JH and Halper, J. (2005) Tendon proteoglycans: biochemistry and function. J Musculoskelet Neuronal Interact 5(1): 22-34.

310. Yuan, HY; Chiou, JJ; Tseng, WH; Liu, CH; Liu, CK; Lin, YJ; Wang, HH; Yao, A; Chen, YT, and Hsu, CN. (2006) FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. Nucleic Acids Res 34(Web Server issue): W635-W641.

311. Yuan, J; Murrell, GA; Wei, AQ, and Wang, MX. (2002) Apoptosis in rotator cuff tendonopathy. J Orthop Res 20(6): 1372-1379.

312. Zafar, MS; Mahmood, A, and Maffulli, N. (2009) Basic science and clinical aspects of achilles tendinopathy. Sports Med Arthrosc 17(3): 190-197.

313. Zondervan, KT and Cardon, LR. (2007) Designing candidate gene and genome-wide case-control association studies. Nat Protoc 2(10): 2492-2501.

University of Cape Town