GENETIC RISK FACTORS FOR OVERUSE AND ACUTE MUSCULOSKELETAL INJURIES

By

Lee-Devlin Hill

HLLLEE004

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
Division of Physiological Sciences
University of Cape Town
February 2024

Supervisors:

Prof. Malcolm Collins

Assoc Prof. Michael Posthumus

Health through Physical Activity, Lifestyle and Sport Research Centre (HPALS)
Sports Science Institute of South African
Boundary Road, Newlands, 7700, Cape Town, South Africa

Doctoral Degrees Board
UNIVERSITY OF CAPE TOWN
Private Bag X3,
Rondebosch 7701,
South Africa
Tel: +27 (0) 21 650 9111

PhD Thesis title: Genetic basis of overuse and acute musculoskeletal injury

DECLARATIONS

I, **Lee-Devlin Hill** 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.

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

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

Signed: _____ Date:_____

This thesis is dedicated to Valérie, Claire, Lilly, my mom, Anni and my brother Oscar, for their love and support.

'*Eagles may soar, but weasels don't get sucked into jet engines*.' – Steven Wright.

TABLE OF CONTENTS

DECLARATIONS iv
ACKNOWLEDGEMENTS xii
LIST OF SCIENTIFIC OUTPUTS FROM THIS THESIS xiii
ABBREVIATIONS xiv
LIST OF FIGURES FROM THIS THESIS xvi
LIST OF TABLES FROM THIS THESIS xix
ABSTRACT xxi
Background xxi
Aims xxii
Methods xxii
Results xxiv
Conclusion xxvi
CHAPTER 1: INTRODUCTION AND SCOPE OF THESIS 1
CHAPTER 2: LITERATURE REVIEW 3
2.1 EPIDEMIOLOGY OF MUSCULOSKELETAL INJURIES 3
2.2 The Shoulder 5
2.2.1 Gross Anatomy of the Shoulder 5
2.2.2 Glenohumeral joint and ligaments 6
2.2.3 The Rotator Cuff Complex 7
2.2.4 Role of Tendons and Ligaments in Glenohumeral Joint Motion and Stability 8
2.2.5 Pathology of Rotator Cuff 8
2.2.6 Mechanism of Injury in Rotator Cuff 10
2.2.7 Risk factors for shoulder pathology 11
2.2.8 The Rotator Cuff and Swimming 15
2.2.9 Rotator Cuff Overuse injuries in Swimming 15
2.3 The knee 19
2.3.1 Gross Anatomy of the Knee 19
2.3.2 Anatomy of ACL 20
2.3.3 Mechanism of Injury in ACL 21
2.3.4 Risk Factors for ACL Injury 22

2.4 TENDONS 23
2.4.1 Molecular Structure of Tendons 23
2.4.2 Tendon Biomechanics 24
2.5 LIGAMENTS 25
2.5.1 Molecular Biology of Ligaments 25
2.5.2 Ligament Biomechanics 26
2.6 COLLAGEN BIOLOGY 27
2.6.1 Overview of Collagens and their structure 27
2.6.2 Fibrillar Collagens 28
2.6.3 Fibril-Associated Collagen with Interrupted Triple Helices (FACITs) 29
2.6.4 Non-Fibrillar Beaded Collagens 30
2.6.5 Other Non-Fibrillar Collagens 31
2.6.6 Collagen Synthesis 32
2.6.7 Collagen Fibril Organization 33
2.7. GENETIC VARIANTS AND COLLAGEN TYPES 35
2.7.1 Type I Collagen 35
2.7.2 Type V Collagen 37
2.7.3 Type VI Collagen 41
2.7.4 Type XI Collagen 42
2.7.5 Type XII and XIV Collagen 43
2.8 AIMS OF THESIS 49
CHAPTER 3: NON-GENETIC RISK FACTORS FOR SHOULDER INJURY IN SWIMMERS 51
3.1 INTRODUCTION 51
3.2 METHODS 52
3.2.1 Search Strategy 52
3.2.2 Level of Evidence and Certainty 55
3.3 RESULTS AND DISCUSSION 57
3.3.1 Shoulder Joint Function and Strength 60
3.3.1.1 Internal/External Rotation 60
3.3.1.2 Clinical Joint Laxity and Instability 61
3.3.1.3 Internal/External Rotation Strength 62
3.3.1.4 Shoulder Flexibility 63

3.3.1.5 Glenohumeral Translation 64
3.3.2 Activity History 64
3.3.2.1 Training load, Volume and Intensity 64
3.3.2.2 Stroke Distance and Stroke Specialty 65
3.3.2.3 Years swimming experience 66
3.3.2.4 Breathing Side 67
3.3.2.5 Swimming Training Equipment 67
3.3.2.6 Cross Training and Stretching 68
3.3.3 Demographics 68
3.3.3.1 Previous history of pain and injury 68
3.3.3.2 Sex 69
3.3.3.3 Age 70
3.3.3.4 Competitive Level 70
3.3.4 Musculoskeletal Factors 71
3.3.4.1 Scapular Kinematics, strength and dyskinesis 71
3.3.4.2 Core Stability 72
3.3.4.3 Pectoral Length 72
3.3.4.4 Sub-acromial bursa thickness 73
3.4 CONCLUSION 73
CHAPTER 4: GENETIC RISK FACTORS FOR ROTATOR CUFF TENDINOPATHY IN SWIMMERS 77
4.1 INTRODUCTION 77
4.2 METHODS 79
4.2.1 Study Design 79
4.2.2 Participants 80
4.2.3 Diagnosis of Rotator Cuff Tendinopathy 81
4.2.4 Blood Collection and Extraction 81
4.2.5 Statistics 92
4.3 RESULTS 93
4.3.1 General Characteristics 93
4.3.2 Characteristics of Clinical Examination 94
4.3.3 Participant and Familial History of Soft Tissue Injuries 94
4.3.4 Swimming Training Variables 95

4.3.5 Independent Genotype Analysis 96
4.3.6 Collagen Gene-Gene Interactions 99
4.3.6.1 Type V and XI Collagen Polymorphism and Gene-Gene Interactions 99
4.3.6.2 Types V and VI Collagen Gene-Gene Interactions 106
4.3.6.3 Types V and XII Collagen Gene-Gene Interactions 107
4.3.6.4 Types VI and XII Collagen Gene-Gene Interactions 108
4.3.6.5 Types V, VI and XII Collagen Gene-Gene Interactions 109
4.4 DISCUSSION 110
CHAPTER 5: COLLAGEN GENE RISK FACTORS FOR ACL RUPTURE IN A SWEDISH AND COMBINED POPULATIONS 118
5.1 INTRODUCTION 118
5.2 METHODS 120
5.2.1 Swedish population Participants 120
5.2.2 DNA Extraction and genotyping 121
5.2.3 Combined Analysis 121
5.2.4 Statistics 122
5.3 RESULTS 123
5.3.1 Participant Characteristics 123
5.3.2 Participant and Family History of Musculoskeletal Soft Tissue Injuries 123
5.3.3 Independent Genotype Analysis 125
5.3.4 Combined Genotype Analysis 127
5.3.5 Inferred Genotype-Genotype Interaction Analysis 133
5.3.5.1 Type V and XI Collagen Gene-Gene Interactions 134
5.3.5.2 Types V, VI and XII Collagen Gene-Gene Interactions 140
5.3.5.3 Types V and XII Collagen Gene-Gene Interactions 141
5.3.5.4 Types V and VI Collagen Gene-Gene Interactions 142
5.3.5.5 Types VI and XII Collagen Gene-Gene Interactions 144
5.4 DISCUSSION 145
CHAPTER 6: GENETIC RISK FACTORS FOR ACL RUPTURE IN A MIXED ANCESTRY COHORT 154
6.1 INTRODUCTION 154
6.2 METHODS 155
6.2.1 Participants 156

6.2.2 Blood Collection, DNA Extraction, and collagen gene polymorphism selection 157
6.2.3 Statistics 157
6.3 RESULTS 157
6.3.1 Participant Characteristics 157
6.3.2 Participant and Family History of Soft Tissue Injuries 158
6.3.3 Genotyping 159
6.3.4 Collagen Gene-Gene Interactions 161
6.3.4.1 Type V and XI Collagen Gene-Gene Interactions 161
6.3.4.2 Types V and VI Collagen Gene-Gene Interactions 167
6.3.4.3 Types V and XII Collagen Gene-Gene Interactions 168
6.3.4.4 Types VI and XII Collagen Gene-Gene Interactions 169
6.3.4.5 Types V, VI and XII Collagen Gene-Gene Interactions 170
6.4 DISCUSSION 170
CHAPTER 7: SUMMARY AND PERSPECTIVES 176
7.1 NOVEL FINDINGS OF THIS THESIS 178
7.1.1 Genetic Risk factors for rotator cuff tendinopathy in swimmers 178
7.1.1.1 Independent collagen variant interactions 178
7.1.1.2. Collagen gene-gene interactions 181
7.1.2 Genetic Risk factors for ACL ruptures 183
7.1.2.1 Independent collagen variant interactions 184
7.1.2.2 Collagen gene-gene interactions 187
7.1.3 Non-genetic Risk factors for Shoulder injuries in swimmers 191
7.2 LIMITATIONS AND FUTURE STUDIES 194
7.3 PRACTICAL IMPLICATIONS 196
7.3.1 Critical Implications for Swimmers and Coaches 196
7.3.1.1 Injury Prevention and Management 196
7.3.1.2 Training Program Design 196
7.3.1.3 Monitoring and Assessment 197
7.3.1.4 Rehabilitation and Recovery 197
7.3.1.5 Educational Implications 197
7.3.2 Implications for Genetic profiling of athletes and the general population 198
7.3.2.1 Personalized Injury Prevention and Training 198

7.3.2.2 Enhanced Understanding of Injury Risks 199
7.3.2.3 Ethical and Privacy Considerations 199
7.3.2.4 Future Research and Methodological Improvements 199
7.4 SUMMARY AND CONCLUSIONS 200
REFERENCES 203
SUPPLEMENTARY MATERIAL 253
A – Ethics Approval and Recruitment Forms 253
B – Supplementary Results 253
Appendix B1: Chapter 2 Summary of investigated genotypes and specific MSK phenotypes I
Appendix B2: Chapter 3 Summary of included studies V
Appendix B3 – RCT Chapter 4 Supplementary Table and Figures XVIII
Appendix B4 – Chapter 5 Swedish ACL and Combined supplementary graphs and tables XXVIII
Appendix B5 – Combined ACL Injuries supplementary tables XLIII
Appendix B6 – Male only Combined ACL cohort Injuries supplementary tables LVII
Appendix B7 – Female only Combined ACL cohort Injuries supplementary tables LXIII
Appendix B8: Mixed ancestry ACL Chapter 6 Supplementary Table and Figures LXIX

ACKNOWLEDGEMENTS

I'd like to thank my supervisors, Prof Malcolm Collins and Assoc Prof Michael Posthumus for their patience and support during this thesis. I understand that I made some difficult decisions over the last few years, and I appreciate the encouragement you have given me to see this thesis.

A special thank you to Mr Neezaam Karim and Mr Trevino Larry for their assistance and guidance in the lab. Your input in everything laboratory-related and computer-related is greatly appreciated. Thank you to Prof Alison September for your expertise and troubleshooting when I was lost.

Id like to thank Dr Nancy Laguette, Dr Masouda Rahim for their unparalleled support with the lab work. Id also like to thank (in no particular order) Dr Sarah McFie, Dr Shameemah Abrahams, Dr Marilize Burger and Dr Kevin O'Connell for their guidance, support and banter.

A special thank you to Dr Sarah Chantler, coffee breaks, ocean swims and banter from across the world. We have come a long way since 2012.

Thank you to each and every participant, without you all, there would be no thesis. Thank you to the admin staff for assisting, even when I made mistakes.

I'd like to thank my family for putting up with me these many years. Anni and Oscar, never failing in their support.

Thank you to Valérie, je ne pensais pas que cela prendrait autant de temps. Sans toi, je ne sais pas où j'en serais arrivé. Tu as changé ma vie d'une manière que je ne peux pas comprendre. Pour cela, je t'aimerai toujours.

Lilly, thank you for brainstorming with me on our long walks. "*Bark, bark bark"*.

Claire, this is for you. Tu es le soleil, la lune et les étoiles.

And finally, this thesis was supported, in part, by funds from the National Research Foundation (NRF) of South Africa and the University of Cape Town.

LIST OF SCIENTIFIC OUTPUTS FROM THIS THESIS

ARTICLES IN INTERNATIONAL PEER-REVIEWED JOURNALS

1. Hill, L., Posthumus, M., Collins, M. (2015). Risk Factors Associated with Shoulder Pain and Injury in Swimmers: A Critical Systematic Review. The Physician and Sports Medicine. 43 (4):412-420.

ABBREVIATIONS

ACL, Anterior cruciate ligament

ANOVA, One-way analysis of variance

TEN, Achilles tendinopathy

ATR, Achilles tendon rupture

BMI, Body mass index

Bp, Base pairs

CI, Confidence interval

COL1A1, The gene encoding the α1 chain of type I collagen

COL5A1, The gene encoding the α1 chain of type V collagen

COL6A1, The gene encoding the α1 chain of type VI collagen

COL11A1, The gene encoding the α1 chain of type XI collagen

COL11A2, The gene encoding the α2 chain of type XI collagen

COL12A1, The gene encoding the α1 chain of type XII collagen

CON, Control group

CTS, Carpal tunnel syndrome

DNA, Deoxyribonucleic acid

ECM, Extracellular matrix

EDS, Ehlers-Danlos syndrome

EDTA, Ethylenediaminetetraacetic acid

FACITs, Fibril-associated collagens with interrupted triple helices

GDF5, The gene encoding growth differentiation factor-5

HWE, Hardy-Weinberg Equilibrium

LD, Linkage disequilibrium

MSK, Musculoskeletal

NCBI, National Centre for Biotechnology Information

NON, Non-contact mechanism of injury

OA, Osteoarthritis

PAGE, Polyacrylamide gel electrophoresis

PC, Prospective cohort

PCR, Polymerase chain reaction

RCT, Rotator Cuff Tendinopathy

RFLP, Restriction Fragment Length Polymorphism

SNP(s), Single Nucleotide Polymorphism(s)

STREGA, Strengthening the Reporting of Genetic Association Studies

TNC, The gene encoding tenascin-C

UK, United Kingdom

UTR, Untranslated region

USA, United States of America

VEGFA, Vascular endothelial growth factor

VEGFA, The gene encoding Vascular endothelial growth factor

WBC, White blood cells

LIST OF FIGURES FROM THIS THESIS

Figure 2.1. An anterior schematic view of the right shoulder with the rib cage resected.

Figure 2.2. An anterior view of the right knee.

Figure 2.3. A schematic diagram of the collagen fibril showing the positions of the various collagen types.

Figure 2.4. An example diagram of the fibril assembly of type I collagen.

Figure 2.5. A schematic example of the hierarchical organization of tendon and ligaments.

Figure 2.6. A general genetic continuum for collagen genes as previously proposed by Collins and Posthumus (1) using Type V collagen as an example.

Figure 3.1. Schematic Diagram of Systematic Review Process.

Figure 4.1. A) Schematic representation of the 18 Kb *COL1A1* gene on chromosome 17q21.33.

Figure 4.2. A) Schematic representation of the 201 Kb *COL5A1* gene *1* gene on chromosome 17q21.33.

Figure 4.3. A) Schematic representation of the 29 Kb *COL6A1* gene on chromosome 21q22.3.

Figure 4.4. A) Schematic representation of the 232 Kb *COL11A1* gene on chromosome 1p21.1.

Figure 4.5. Schematic representation of the 29 Kb *COL11A2* gene **A)** on chromosome 6p21.3.

Figure 4.6. Schematic representation of the 121 Kb *COL12A1* gene **A)** on chromosome 6p12.

Figure 4.7. Typical non-denaturing visualising gels showing the genotypes of **A)** *COL12A1* rs970547 and **B)** *COL5A1* rs12722 restriction fragment length polymorphisms (RFLPs).

Figure 4.8. Typical allelic discrimination plots using the Taqman® Genotyping assay..

Figure 4.9. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) and *COL5A1* rs10628678 (AGGG/-) polymorphisms…

Figure 4.10 A) Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and rs1676486 (C/T), *COL11A2* rs1799907 (A/T) polymorphisms…

Figure 4.11 A) Inferred haplotype frequency distributions constructed from *COL5A1* rs10628678 (AGGG/-), *COL11A1* rs3753841 (T/C) and rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) polymorphisms.

Figure 4.12. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) *COL6A1* rs35796750 (T/C) polymorphisms…

Figure 4.13 Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) and *COL12A1* rs970547 (G/A) polymorphisms…

Figure 4.14. Inferred haplotype frequency distributions constructed from *COL6A1* rs35796750 (T/C) and *COL12A1* rs970547 (G/A) polymorphisms…

Figure 4.15. A schematic diagram of the proposed mechanisms that predispose a swimmer to rotator cuff tendinopathy.

Figure 5.1 The *COL1A1* rs1800012 (G/T) genotype distributions…

Figure 5.2 The *COL5A1* rs12722 genotype distributions…

Figure 5.3 The *COL6A1* rs35796750 genotype distributions…

Figure 5.4 The *COL11A1* rs3753841 genotype distributions…

Figure 5.5 *COL12A1* rs970547 genotypes of combined ACL…

Figure 5.6. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) and *COL5A1* rs10628678 (AGGG/-)…

Figure 5.7. Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and rs1676486 (C/T), *COL11A2* rs1799907 (A/T) polymorphisms…

Figure 5.8. Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) polymorphisms…

Figure 5.9. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C), *COL6A1* rs35796751 (T/C) and *COL12A1* rs970547 polymorphisms…

Figure 5.10. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C), and *COL12A1* rs970547 polymorphisms…

Figure 5.11. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) and *COL6A1* rs35796751 (T/C) polymorphisms…

Figure 5.12. Inferred haplotype frequency distributions constructed from *COL6A1* rs35796751 (T/C) and *COL12A1* rs970547 polymorphisms…

Figure 6.1. The distribution of the inferred haplotypes constructed from the *COL5A1* rs12722 (T/C) and *COL5A1* rs10628678 (AGGG/-) polymorphisms…

Figure 6.2. Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and rs1676486 (C/T), *COL11A2* rs1799907 (A/T) polymorphisms.

Figure 6.3 A) Inferred haplotype frequency distributions constructed from *COL5A1* rs10628678 (AGGG/-), *COL11A1* rs3753841 (T/C) and rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) polymorphisms.

Figure 6.4. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) *COL6A1* rs35796750 (T/C) polymorphisms…

Figure 6.5. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) and *COL12A1* rs970547 (G/A) polymorphisms…

Figure 6.6. Inferred haplotype frequency distributions constructed from *COL6A1* rs35796750 (T/C) and *COL12A1* rs970547 (G/A) polymorphisms…

Figure 6.7. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C), *COL6A1* rs35796750 (T/C) and *COL12A1* rs970547 (G/A) polymorphisms…

Figure 7.1. A schematic diagram, adapted from Meeuwisse's (2) proposed injury model of the dynamic and multifactorial aetiology of injury.

LIST OF TABLES FROM THIS THESIS

Table 2.1. Summary of independent genetic associations in several phenotypes.

Table 2.2. A summary of the candidate collagen gene variants and injury phenotypes that will be investigated in this thesis.

Table 3.1. A three-step method was followed to identify the articles that were included in the systematic review of risk factors associated with shoulder pain or injury in swimmers.

Table 3.2. The identified articles were included in the systematic review of risk factors associated with shoulder pain or injury in swimmers if they met the following inclusion criteria.

Table 3.3. Level of evidence (I–V) definitions used for study evaluation…

Table 3.4. Level of Certainty (high-low) definitions used for risk assessment…

Table 3.5. Summary of included studies from the periods 1985-2015 and 2015-2021.

Table 3.6. Summary of the level of certainty of risk factors associated with shoulder injuries in swimmers

Table 4.1. Details of custom designed Fluorescence-based Taqman® polymerase chain reaction (PCR) assays

Table 4.2: General participant characteristics for the control (CON), Rotator Cuff Tendinopathy (RCT) groups and the Supraspinatus Tendinopathy (SST) sub-group.

Table 4.3: Participant and family history of soft tissue injuries for the participants in the control group (CON), Rotator Cuff Tendinopathy (RCT) group as well as the Supraspinatus Tendinopathy (SST) sub-group.

Table 4.4: Summary of Swimming Training Variables control group (CON), Rotator Cuff Tendinopathy (RCT) group Supraspinatus Tendinopathy (SST) sub-group.

Table 4.5: Genotype frequency distributions for *COL1A1* rs1800012, *COL5A1* rs12722, *COL5A1* rs10628678, *COL6A1* rs35796750, *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 and *COL12A1* rs970547 in the control (CON), Rotator Cuff Tendinopathy (RCT) group and Supraspinatus Tendinopathy (SST) sub-group.

Table 4.6. A continuing summary of the genotype results from each chapter in this thesis.

Table 4.7. A continuing summary of the haplotype results from each chapter in this thesis.

Table 5.1: Summary of previously investigated European Ancestry Anterior Cruciate Ligament (ACL) rupture and related healthy Control (CON) populations.

Table 5.2: General participant characteristics for the Swedish CON and NON groups.

Table 5.3. Medical history and family injury for CON and NON groups of the Swedish cohort.

Table 5.4. Genotype frequency distributions for *COL1A1* rs1800012, *COL5A1* rs10628678, *COL6A1* rs35796750, *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 and *COL12A1* rs970547 in the Swedish control (CON) and Swedish non-contact anterior cruciate ligament injuries (NON).

Table 5.5. A continuing summary of the genotype results from each chapter in this thesis.

Table 5.6. A continuing summary of the haplotype results from each chapter in this thesis.

Table 6.1: General participant characteristics and sport participation for the controls (CON), ACL Rupture (ACL) group and non-contact ACL rupture (NON) sub-group.

Table 6.2. Genotype frequency distributions for COL1A1 rs1800012, COL5A1 rs12722, COL5A1 rs10628678, COL6A1 rs35796750, COL11A1 rs3753841, COL11A1 rs1676486, COL11A2 rs1799907 and COL12A1 rs970547 in the control (CON), anterior cruciate ligament rupture (ACL) and non-contact anterior cruciate ligament injuries (NON).

Table 6.3. A continuing summary of the genotype results from each chapter in this thesis.

Table 6.4. A continuing summary of the haplotype results from each chapter in this thesis.

Table 7.1. Summary of independent genetic associations of several injury phenotypes.

Table 7.2 Main Inferred haplotypes for chronic and overuse injuries.

Table 7.3 Main Inferred haplotypes of ACL ruptures.

Table 7.4: The minor allele and genotype frequencies of the *COL1A1* rs1800012 (G>T), variant within the five main population groups listed in Ensembl.

ABSTRACT

Background

Both acute and chronic tendon and ligament injuries are multifactorial phenotypes, which are the result of a combination of a poorly understood and involve the interaction between multiple intrinsic and extrinsic risk elements. An expanding array of evidence indicates that inherited genetic components might incline an individual towards injury risk, thus highlighting their significance as crucial intrinsic risk factors to be taken into account. Previous studies have investigated the association of several collagen gene (*COL1A1*, *COL5A1*, *COL6A1*, *COL11A1*, *COL11A2* and *COL12A1)* DNA sequence variants with chronic lower limb tendinopathies, such as Achilles tendinopathy, and other exercise-associated phenotypes involving the musculoskeletal system. These genes encode for important structural components of both tendons and ligaments and have been proposed to influence the inter-individual variation in the biomechanical properties of these tissues and by implication, predisposition to injury. The association of these collagen gene variants with rotator cuff tendinopathy (RCT), more specifically supraspinatus tendinopathy (SST), has not been extensively investigated. Except for *COL1A1* variants, the association of the remaining collagen gene variants with an acute injury, such as anterior cruciate ligament (ACL) ruptures, has also not been extensively investigated. In addition, the association of these gene variants with musculoskeletal soft tissue injures has predominately been investigated in populations of European ancestry.

Aims

Therefore, the primary aim of this thesis was to investigate the independent association of the *COL1A1* rs1800012 (G/T), *COL5A1* rs12722 (T/C), *COL5A1* rs10628678 (AGGG/-), *COL6A1* rs35796750 (T/C), *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T), *COL11A2* rs1799907 (A/T) and *COL12A1* rs970547 (G/A) gene variants with RCT in a South African cohort of swimmers (Chapter 4) of European ancestry, as well as ACL rupture in a combined cohort of European ancestry (Swedish, South African, Polish and Australian) (Chapter 5) and a South African Mixed Ancestry cohort (Chapter 6) using a case-control genetic association study approach.

A secondary aim of the thesis was to investigate hypothesis-driven gene-gene interactions between the investigated collagen gene variants in modulating the risk of RCT injury (Chapter 4) and the two ACL (Chapters 5 and 6) cohorts. Finally, a systematic review of the non-genetic risk factors associated with RCT in swimmers was also included in this thesis (Chapter 3).

Methods

Chapter 3 investigated the non-genetic risk factors in swimmers using a systematic review methodology according to the Preferred Reporting for Systematic Reviews and Meta Analyses (PRISMA) guidelines. An electronic database search of SpringerLink, Science Direct and PubMed/Medline was conducted during 2015 and with an updated search in 2022 following the same protocol. Risk factors were grouped to together into common themes and the level of evidence and the strength of evidence was critically appraised.

For Chapter 4, 103 (49 females, 54 males) swimmers with clinically diagnosed rotator cuff tendinopathy (RCT group) were recruited, of which 84.5% (n=87) were diagnosed with a supraspinatus tendinopathy (SST sub-group). In addition, 101 (55 females,46 males) apparently healthy swimmers with no previous history of shoulder pathology (including RCT, trauma, bursitis, or adhesive capsulitis) (CON group) were recruited. All participants were unrelated, of self-reported European ancestry and recruited between 2013 and 2016.

For Chapter 5, 195 physically active and unrelated participants of self-reported European ancestry were recruited between 2011 and 2013 from the University Hospital in Umeå and orthopaedic clinics in Luleå, Sweden. These participants within this cohort comprised of 79 individuals who had been clinically diagnosed ACL ruptures with a non-contact mechanism of injury (ACL group) and 116 apparently healthy, asymptomatic individuals with no history of ACL ruptures (CON group). The Swedish cohort was included in a larger combined analysis consisting of 661 participants with ACL rupture and 378 uninjured controls from previously published cohorts of self-reported European ancestry from South Africa, Poland, and Australia.

For Chapter 6, 209 unrelated South African self-reported mixed ancestry participants were recruited and included in this study. Ninety-four participants (77 males and 17 females) were included in the ACL group, of which 51 had sustained their ACL rupture through a non-contact mechanism of injury (NON sub-group). Furthermore, 100 (81 males and 19 females) apparently healthy, individuals with no history of ACL rupture or injury were recruited from gyms and local sports clubs within the Cape Town area of South Africa.

All participants were genotyped for the following collagen gene polymorphisms: *COL1A1* rs1800012 (G/T), *COL5A1* rs12722 (T/C) and rs10628678 (AGGG/-), *COL6A1* rs35796750 (T/C),

COL11A1 rs3753841 (T/C) and rs1676486 (C/T), *COL11A2* rs1799907 (A/T) and *COL12A1* rs970547 (G/A).

Results

As presented in a systematic review of non-genetic risk factors (Chapter 3), a total of 1050 articles were reviewed between 1985 and 2021. The risk factors identified from the systematic literature search were categorized into four categories based on factors identified during the review process, namely shoulder joint function and strength, activity history, demographics, musculoskeletal anatomy and sub-acromial bursa thickness. Only four risk factors for shoulder injuries were determined to be of moderate certainty, with the remaining 25 risk factors being appraised as low certainty. Moderate level of certainty was determined for (i) previous history of pain and injury, (ii) internal/external rotation range of motion, (iii) clinical joint laxity and instability and (iv) internal/external rotation strength.

Although previously associated in some studies investigating other overuse musculoskeletal soft tissue injuries, none of the investigated collagen gene variants were independently associated with rotator cuff tendinopathy (RCT) or supraspinatus tendinopathy (SST) risk (Chapter 4). A novel finding of this thesis was that the C-A-(-) inferred haplotype constructed from *COL11A1* rs3753841(T/C), *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) was significantly (p=0.034) under-represented in the SST (0.4 %) sub-group compared to the CON group (6.0 %). However, none of the other inferred haplotypes constructed from (i) the two *COL5A1* variants, (ii) the two *COL11A1* variants, (iii) the three *COL11A1* and *COL11A2* variants, as well as all (iv) the *COL11A1, COL11A2* and *COL5A1* variants were associated with RCT or

SST risk. Similarly inferred haplotypes constructed from (i) *COL5A1* and *COL6A1*, (ii) *COL5A1* and *COL12A1*, as well as (iii) *COL6A1* and *COL12A1* were also not associated with RCT or SST.

The *COL1A1* rs1800012 TT genotype was found to be significantly (p=0.027) under-represented in the ACL group (1.1%) of European ancestry during the combined analysis compared to the CON group (TT 4.0 %). Interestingly this significant association was only observed in females (0.0% ACL vs 4.8% CON, p = 0.045) but not males (1.5% ACL vs 3.6% CON, p =0.299). Although independently associated with ACL rupture in European populations, the *COL1A1* rs1800012 TT genotype was however not significantly associated with ACL rupture in the Mixed Ancestry cohort (1.0% ACL vs 0.0 % CON, p=0.204). An additional novel finding was that the *COL12A1* rs970547 GG genotype was significantly associated with increased risk of ACL rupture the South African Mixed Ancestry cohort (20.9% ACL vs 5.4% CON, p=0.021). This genotype was however not independently associated with ACL rupture in the combined European cohort (5.1 % ACL vs 4.3% CON, p=0.423). None of the other investigated collagen gene variants were independently associated with ACL rupture in the European or mixed ancestry cohorts.

Within participants of European ancestry, the C-AGGG (31.2% ACL vs 20.6% CON, p=0.001) and T-(-) (14.4% ACL vs 5.7% CON, p=0.010) inferred haplotypes constructed from *COL5A1* rs12722 (C/T) and rs10628678 (AGGG/-) were significantly over-represented in the ACL rupture group, while the T-AGGG inferred haplotype was significantly under-represented (36.6% ACL vs 50.5 CON, p<0.001). None of the *COL5A1* inferred haplotypes were however significantly associated with ACL rupture in the South African mixed ancestry cohort.

A further novel finding was a significant interaction between the *COL6A1 rs35796750* and *COL12A1* rs970547 variants, the T-A (11.0% ACL vs 7.1% CON, p=0.030) and T-G (23.0% ACL

vs 33.7% CON, p=0.010) inferred haplotype was significantly over- and under- represented, respectively in the ACL group when the combined participants of European ancestry were analysed. Within the South African mixed ancestry population, the different C-G inferred haplotype constructed from the *COL6A1* and *COL12A1* variants was significantly (p=0.029) over-represented in the ACL group (37.2%) compared to the CON group (31.1%). This haplotype remained significantly (p=0.027) associated when only the participants with a non-contact mechanism of injury (34.3% NON sub-group) was analysed.

Finally, the inferred T-A haplotype constructed from *COL5A1* rs12722 and *COL12A1* rs970547 was significantly (p=0.039) over-represented with ACL group (27.3% ACL vs 17.9% CON) in the combined European, but not the mixed ancestry, cohorts.

Conclusion

Several non-genetic risk factors were identified during the systematic review. The aetiology of shoulder injuries in swimmers, whilst well defined, is still poorly understood. Further, the review revealed no studies have incorporated genetic predisposition into the risk factor assessment for swimmers. Secondly, even though only two collagen gene variants were found to be independently associated with modulating the risk of ACL ruptures, namely *COL1A1* rs1800012 and *COL12A1* rs970547 in the European and mixed ancestry cohorts respectively, several distinct collagen gene variants inferred haplotypes associated with RCT and ACL ruptures. An inferred haplotype constructed from the *COL11A1*, *COL11A2* and *COL5A1* variants were associated with SST. Distinct differences in the associated haplotypes with ACL rupture in the two population groups were also observed. Specifically, *COL5A1* inferred haplotypes modulated the risk of ACL rupture in the combined European, but not the South African mixed ancestry, cohorts. Different *COL6A1*

and *COL12A1* inferred haplotypes modulated the risk of ACL rupture in the two ACL populations, while a *COL5A1* and *COL12A1* inferred haplotypes modulated the risk of ACL rupture in only the European cohort. Taken together these findings highlight the possibility that the association collagen and other gene variants are injury and population specific

CHAPTER 1: INTRODUCTION AND SCOPE OF THESIS

Physical activity is an integral part for growth and both physical and mental development and is an essential aspect of maintaining one's health (1–4). Even though there are several health benefits to regular physical activity, participation in exercise has been shown to increase the risk of injury, specifically to the musculoskeletal system (MSK) (5–8). Furthermore, these injuries can be subdivided into either chronic (tendinopathies and other overuse injuries) or acute (partial or complete ruptures and tears) injuries depending on the causative mechanisms (9–11). Despite an insufficient understanding of the biological mechanisms of these injuries, several intrinsic, which includes genetic factors, and extrinsic factors have been associated with risk (reviewed in Chapter 2, sections 2.2.5 – 2.3.4) (12–20).

The recognition of risk factors, particularly genetic ones (Chapter 2, section 2.7), will further enhance the understanding of the underlying mechanisms and aetiology of tendon and ligament injuries (16,21,22). Since tendons and ligaments are collagenous tissues, the main objective of this thesis was therefore to further explore the genetic contribution of the collagen encoding genes, specifically *COL1A1*, *COL5A1*, *COL6A1*, *COL11A1* and *COL11A2*, in modulating the risk of these injuries. Previous research has predominately focused on the association of collagen gene variations with chronic lower limb tendinopathies, with only a limited number of studies investigating the association of these variants with upper body tendinopathies or acute injuries (21,23–32). Therefore, the primary objective of this thesis was to investigate if rotator cuff tendon (an overuse injury) and anterior cruciate ligament (ACL) (an acute injury) injuries are associated with previously identified polymorphisms. Candidate genetic loci were selected based on their

previous implication in MSK phenotypes, in addition to their potential role in the structural and biological function of tendons and ligaments.

This genetic contribution to rotator cuff injuries was investigated using a case-control genetic association study design within swimmers with symptomatic shoulder pain and injury and asymptomatic swimmer controls of self-reported European ancestry (Chapter 4). The genetic contribution of ACL ruptures was also investigated using the same study design in a combined European ancestry cohort consisting of Swedish, South African, Polish, and Australian participants (Chapter 5). Furthermore, since previous studies have primarily investigated populations of European ancestry, the association of these variants with ACL rupture was investigated in a South African mixed ancestry cohort (Chapter 6) consisting of ACL ruptures which were surgically confirmed and healthy age and sex matched controls.

In preparation for exploration and discussion of the collagen gene association studies (Chapters 4-6) of this thesis, Chapter 2 will provide a brief review of the epidemiology, gross anatomy, potential mechanisms of injury including non-genetic risk factors and previously identified genetic risk factors of tendon and ligament injuries with specific focus on the rotator cuff and ACL. The non-genetic risk factors for ACL rupture, but not shoulder injuries in swimmers, have been extensively reviewed(15,17,20,33). Therefore, the non-genetic risk factors for shoulder injuries in swimmers was systematically reviewed in Chapter 3. Finally, the main novel findings and limitations of this thesis are discussed in the final chapter (Chapter 7).

CHAPTER 2: LITERATURE REVIEW

2.1 EPIDEMIOLOGY OF MUSCULOSKELETAL INJURIES

The MSK system comprises of different tissues that provide the body with various structures and functions, including movement, stability, and support which include tendons and ligaments (34). Tendons connect muscle to bone and convert muscle contraction to joint motion (35). Ligaments connect two articulating bones across a joint and are responsible for maintaining joint congruency and guiding joint movement (36). Although functionally different, tendons and ligaments are nevertheless very similar in their compositions, with only certain minor variations (37).

Tendons and ligaments, including their surrounding structures, as with other MSK tissues, are commonly injured as a result of regular and/or repetitive physical activity participation, including various activities in the workplace (38–41). MSK soft tissue injuries are divided into either acute (resulting from a single macro-trauma event e.g., acute ACL rupture) or chronic (overuse e.g., RCT) injuries. It has been reported that tendon injuries account for approximately 30-50% of all sporting injuries (17,41,42)of which, injuries to the rotator cuff tendons being the most common source of pain and dysfunction in the shoulder (43–46). Tendon injuries often result in chronic pain, discomfort and long-term functional impairment, remain a challenge to both clinicians and surgeons and are often 1) recurrent 2) resistant to current treatment protocols and 3) risk factors for other orthopaedic injuries and medical conditions (47). The recurrent motion involved in swimming strokes and the rigorous training regimes frequently employed raise numerous concerns

regarding the occurrence and seriousness of injuries throughout a competitive season (48–55). The shoulder stands out as the most frequent site of injury, with supraspinatus tendinopathy being the most commonly cited pathology (56–59).

Similarly, injuries to ligaments are common and account for about 22-30% of all athletic injuries (60) and can vary from a mild ligament sprain to complete ruptures (tears) to a single or to multiple ligaments (61). Injuries to ankle ligaments are the most common lower extremity injury (62), however injuries to the knee complex, specifically, to the various ligaments in the knee that support the joint and limit excess movement are the most severe (60,62). These ligaments include the anterior cruciate ligament (ACL), medial collateral ligament (MCL), lateral collateral ligament (LCL) and posterior cruciate ligament (PCL) (63–65). Of which, the ACL being the most injured, with ruptures occurring in about 1 in every 3500 people (66). Furthermore, surgical repair of ACL can be costly, not including time lost from work and the subsequent rehabilitation (67). Furthermore, ACL injury has long term implications and personal consequences, in terms of disability, pain and discomfort, reduced knee function, impact on physical activity and an increased risk of osteoarthritis later in life (68–70).

Injuries to MSK soft tissues are multifactorial and although they have different aetiologies, they do share several intrinsic and extrinsic risk factors (37,71,72). Common extrinsic risk factors include sporting factors, protective equipment, and environment. Intrinsic risk factors include age, sex, health, anatomy, physiological factors, and genetic profile (17,71,73). Several genetic and/or familial risk factors have been identified for various specific MSK soft tissue injuries, such Achilles (74–82), ACL (83–91), rotator cuff (27,29,44,92,93), carpal tunnel (94–98), tennis elbow (24,25).

Broadly the genetic factors that modulate the risk of MSK injuries may be classified as those coding for 1) structural components of connective tissue (collagens, elastin and non-fibre forming glycoproteins); 2) extracellular matrix (ECM) proteases (MMPs); 3) cytokines and growth factors; and 4) other regulatory molecules (37,99). Most of the research has investigated the collagen encoding genes. Collagen is the main structural protein in connective tissue and is the most abundant protein in mammals, making up 25-35% of total protein content (100–103). As is shown in Table S.S1, many significant collagen gene associations have been reported with the ACL and Achilles Tendinopathy (TEN). Several collagen genetic loci have also been associated with tennis elbow(24,25), carpal tunnel syndrome (94,98) and shoulder dislocation (27).

2.2 THE SHOULDER

2.2.1 Gross Anatomy of the Shoulder

The shoulder, or glenohumeral joint, ranks among the most intricate joints in the human body. The complex structure of the joint allows for the most mobility of any major joint in the body (104–106). This is in part due to the ball-and-socket type joint that confers a limited interface between the humerus and the scapula. Yet, to ensure effective joint movement, a vast network of ligaments, tendons, and connective tissues offers stability and enables functional motion.

The shoulder girdle itself is comprised of several joints that connect the head of the humerus to the scapula and clavicle(104,107,108). The scapula serves as a landmark for the coracoid process superiorly (Figure 2.1), the glenoid cavity laterally, the subscapular process superiorly and the supraspinous and infraspinous fossae divided by the scapular spine posteriorly (104,107,108). The

scapular spine extends laterally to a free end, the acromion, which articulates with the lateral end of the clavicle (104,108). Medially to laterally, in the anterior proximal section of the humerus, are the lesser tuberosity and, bicipital groove and the greater tuberosity landmarks. The medial side of the humeral head is comprised of articular cartilage which is critical to providing a smooth articular service between the humeral head and the glenoid fossa (104,109). The articular cartilage is predominantly composed of type II collagen and the proteoglycan aggrecan that can be divided into two zones from the surface articular cartilage to the subchondral bone (110–112).

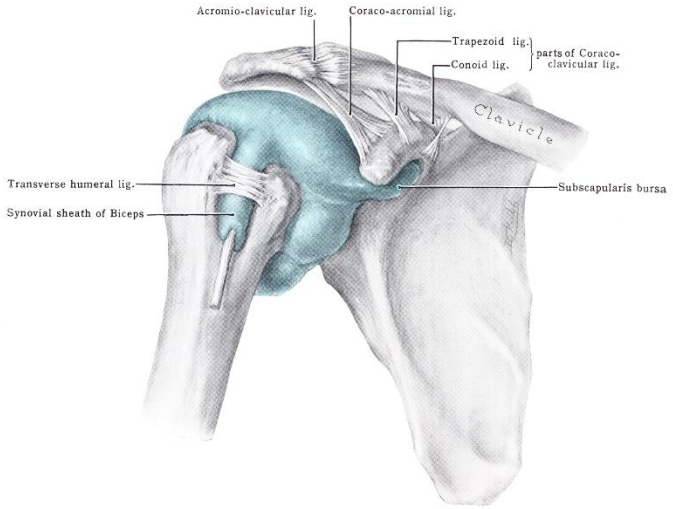


Figure 2.1. An anterior schematic view of the right shoulder with the rib cage resected. The locations of the primary skeletal structures (clavicle, coracoid process, acromion, scapula, humeral shaft, rib glenohumeral bursa) including the growth plate and biceps tendon. Not shown on the diagram are the rib cage, glenohumeral joint and tendon of the short head of the bicep. *Wikimedia Commons, the free media repository*. Retrieved 10:09am, 12 February 2024 from [commons.wikimedia.org].

2.2.2 Glenohumeral joint and ligaments

Numerous ligaments are involved in providing support to the shoulder including the superior and inferior acromioclavicular ligaments, the coracoclavicular ligament, the coracohumeral ligament,

the coracoacromial ligament, and the superior, middle, and inferior glenohumeral ligaments (113). The coracoacromial ligament functions as the ceiling of the subscapular space and forms the coracoacromial arch in conjunction with coracoid process and the acromion (108,114,115). The labrum envelopes the glenoid cavity which is composed of a band of fibrocartilaginous tissue that varies in form, thickness, and density around the margin of the glenoid cavity (116–118). The articular capsule of the glenohumeral joint, therefore forms multiple attachments to the labrum, scapular neck, and anatomical neck of the humerus (119,120). Stabilisation of the glenohumeral joint occurs through the glenohumeral ligaments, located anteriorly and coracohumeral ligament located superiorly (121). However, the inferior section of the joint capsule is not reinforced by fibrocartilage, thus resulting in the axillary recess and allowing for greater freedom of movement, in addition to anterior and superior translation of the humerus (108,114,122).

2.2.3 The Rotator Cuff Complex

As a prelude to this section, it must be noted that although the rotator cuff consists of four muscles and their respective tendons that may appear to be separate superficially, however in their deeper regions they are associated with each other, the capsule underneath, and the long head of the biceps tendon (123,124). The rotator cuff is comprised of four muscles which includes the supraspinatus, infraspinatus, subscapularis and teres minor and their tendons (123–125). The supraspinatus muscle originates from the supraspinous fossa of the scapula where the supraspinatus tendon passes into the subscapular space and has its insertion on the superior and middle facets of the greater tuberosity (119,126). The infraspinatus and teres minor both originate from the infraspinous fossa and fibrous septum, and their tendons insert on the middle and inferior facets of the greater tuberosity, respectively (119). The subscapularis originates from the subscapular

fossa, and its tendon inserts on the lesser tuberosity of the humerus (115). The rotator cuff is exceptionally unique in that its tendons are fused at their insertions to form a continuous structure. The bursal surface of the glenoid cavity is covered by the deep extensions of coracohumeral ligament, while the articular surface articulates with the joint capsule (127,128).

Components of the rotator cuff tendons have unique properties, conferring specific biomechanics and strain distributions across its various regions and surfaces (129–131). The supraspinatus tendon, for example consists of two sub-regions, anterior and posterior, which have a significantly higher modulus of elasticity occurring within the anterior sub-region compared to the posterior sub-region (104). These properties serve to preserve the shape of the tendon during shoulder movements (131)

2.2.4 Role of Tendons and Ligaments in Glenohumeral Joint Motion and Stability

The inherent stability of the shoulder joint is poor, as the articular surface of the proximal humerus and the glenoid are not equal in size (104,106,132). The cumulative effect of the fibrocartilaginous labrum in addition to the joint capsule and ligaments provides the shoulder with stability (118,121,133). These specific tendon and ligament connections confer a wide range of motion and create clearly defined roles for tendons and ligaments as either specifically active or passive restraints during movement of the arms.

2.2.5 Pathology of Rotator Cuff

Clinically, the rotator cuff is difficult to diagnose when a pathology begins to occur. Several clinical tests have been developed in order to assess each individual tendon (134). To do this, the

shoulder and arm are moved and manipulated into different defined positions and then a force is applied to the arm requiring the muscle tissue to contract (135). But as stated previously, rotator cuff tendons are fused into a single structure with supraspinatus and infraspinatus fused almost completely together at their insertions (119). The subscapularis and supraspinatus fuse to form a sheath which envelopes the biceps tendon. Further, the muscular portions of the infraspinatus and teres minor are also fused near to the proximal section of the musculotendinous junction (119). The paradox is that although the rotator cuff forms a tightly adherent structure around the glenohumeral joint capsule, creating a complex that is interwoven and improved resistance to failure underload, it is near impossible to test a singular structural unit (18,120,133,134,136). Clinical presentation of pain and impairment of the shoulder during movement and function usually occurs shoulder elevation and rotation (134,136,137). Numerous factors including hormonal influences (138), genetics (44,139–141), lifestyle factors (142,143), biomechanical (104,107,117), patho-anatomical (18,107,120), sensory and motor cortex changes (134) even psycho-social factors (144) have the potential to contribute to pain of the rotator cuff. Excessive and mal-adaptive load imposed on the tissues are most certainly a major factor in themselves (40,145,146).

To date, there is no clear consensus as the cause of pain in the rotator cuff, the mechanism of injury, the connection between symptoms and observed structural failure within the rotator cuff tendons and the exact role and extent of inflammatory response (18,46,147,148). The symptoms are commonly termed tendinopathy which implicates the tendon as the source of the pathology, however, due to the complex nature of the shoulder, it is difficult to directly state that the only the tendons are the source of the pain experienced. Therefore, to accurately diagnose a rotator cuff

pathology, a number of factors must be considered including structural abnormalities, patient feedback, psycho-social factors in order to accurately define and treat the injury or pathology.

2.2.6 Mechanism of Injury in Rotator Cuff

In terms of defining mechanism of injury, Neer (149,150) had postulated that 95% of all rotator cuff related pathologies were because of irritation and inflammation of the rotator cuff tendons and the subacromial bursa against the over-lying anterior aspect of the acromion terming the condition as subacromial impingement syndrome (151,152). The relevance of the acromion in the development of painful symptoms and rotator cuff tears remains unclear. The argument that the irritation and inflammation caused by friction against the acromion is not supported by observational studies (153–156).

It has been repeatedly and consistently shown that the tears to the rotator cuff are predominately located within the tendon itself or on the articular side which does not support the acromial impingement model. Interestingly, Hashimoto and colleagues (157) had observed variations within rotator cuff morphology reporting that greater fibre degeneration and disorientation in the middle and deeper fibres and that that the degeneration occurring directly lead to rotator cuff tears. Further, Nakajima et al. (129) had shown that the deeper fibres of the rotator cuff tendons had a smaller cross-sectional area than the articular side fibres and that the lower fibres failed at approximately half the tensile load of the superiorly located fibres. Moreover, strain within the supraspinatus tendon increased when the arm was elevated between 15 and 60 degrees of abduction with no significant difference in tensile strain between upper and lower fibres (158,159). However, as shown by Nakajima et al (129), the lower fibres are relatively weaker and therefore suggests that

the joint side fibres are more vulnerable to tensile load than the bursal side fibres making them more susceptible to failure during elevation. Therefore, failure may occur independent of acromion irritation. The importance of the acromion is further challenged in a longitudinal study by Henkus et al. (160). The authors randomised 2 groups of patients (n=57) who had subacromial impingement syndrome into either acromioplasty and bursectomy or bursectomy only. At the 2.5 year follow up after the surgery, both groups had reported similar results with no significant differences occurring any of the should test variables (Constant score, Simple Shoulder test, and visual analogue scores for pain and function). The authors argued that the outcome scores supported an intrinsic degenerative model (40,46,161–163) for rotator cuff pathology rather than the extrinsic acromial irritation model.

2.2.7 Risk factors for shoulder pathology

The shoulder is a complex ball and socket joint and injuries to the rotator cuff are common (12). Injury rates vary between 13% (164) to 54% (165) and risk increases with age. The literature often refers to the several generic terms for shoulder related pathologies comprising of “rotator cuff disease”, “rotator cuff disorders” and “rotator cuff pathology” which includes but is not limited to impingement syndrome, subacromial or subdeltoid bursal disease, RCT, rotator cuff degeneration and rotator cuff tears (166). Further, rotator cuff pathology is a clinical syndrome that refers to structural changes and clinical presentation in both symptomatic and asymptomatic individuals (167–169). In a recent meta-analysis (12), it was found that increasing age was associated with higher risk of sustaining a rotator cuff tear with prevalence increasing markedly after 50 years of age (167,170).

The aetiology of rotator cuff pathology is multifactorial and revolve around the homeostasis of normal tissues and the apoptosis pathways of damaged tissues (163). Several risk factors have been noted in the literature including smoking (143), hypercholesterolemia (171), dominant hand (168,169), older age (164,169), trauma (172–174), anatomical variation (109,175) and overuse (176). Several studies in recent years have investigated the familial and genetic predisposition to rotator cuff pathology (21,29,109,140,141,177–180) alluding to rotator cuff pathology being a heritable trait (166).

The first of these studies investigating inherited risk was Harvie et al. (181) who retrospectively examined and compared the rates of symptomatic and asymptomatic rotator cuff tears in siblings and their spouses. Using a retrospective cohort, the authors evaluated 205 patients, 150 spouses and 129 siblings with diagnosed full-thickness tears of the rotator cuff tendons. The study showed that siblings within the experimental group had more than twice the risk of developing tears of the rotator cuff relative to the control group and nearly five times the risk of experiencing symptoms. It was determined that a sibling of someone with a rotator cuff tear were more than twice at risk for developing a tear compared to the unrelated spouses. The group followed up several years later and found that the siblings were at higher risk of their tears progressing and requiring and intervention (44). The authors demonstrated that rotator cuff tears in siblings had a higher risk of progressing with the relative risk for cases to have full thickness tears was 2.84% compared to 1.44% in the controls (94% CI, 1.75-4.64). More recently, Tashjian et al. (139) utilised a population database of more than 1889 patients and determined that significantly elevated risk for rotator cuff tears in first and second-degree relatives. Further, it was determined that a person was at significant risk for developing a tendinopathy if a first or second degree relative had a history of tendinopathy. Furthermore, several studies have identified genetic sequence variants that were

associated with rotator cuff pathology risk including *MMP-1* (182), *MMP-3* (183), *ESRRB* (140,184), *SASH1* (185), *TNC* (186), *GDF-5* (30), *COL1A1* (27,30,174) and *COL5A1* (31,33,202) (30,187). Rotator cuff pathology is comprised of several risk factors that may modulate risk. It is difficult to determine its exact aetiology in addition to the individuals often being asymptomatic (164,168–170). The identification and characterisation of these factors is crucial to furthering our understanding of rotator cuff pathology risk, specifically tendons and tendinopathy of the rotator cuff.

More recently, a handful of studies (29,140,179,180,188–190) have investigated potential associations of polymorphisms within candidate genes with rotator cuff disease. One of the first studies by Peach et al. (191) attempted to investigate the association polymorphisms within *ANKH* and *TNAP*, which encode for a multipass transmembrane protein transporting PPi (inorganic pyrophosphate) and Tissue-Nonspecific Alkaline Phosphatase which hydrolyses PPi to Pi, respectively, with cuff tear arthropathy. Six SNPs in the *ANKH* gene were investigated and it was found that the heterozygous genotype of the rs3045 (C/T) 3'UTR single nucleotide polymorphisms (SNP) of *ANKH* was significantly over-represented (Relative Risk= 2.78) in patients with cuff tear arthropathy (44%) compared to controls (20%). Further, 16 SNPs were investigated in *TNAP* where it was found that the heterozygous genotype of the rs4654760 (C/T) was significantly over-represented (Relative Risk=2.85) in patients with cuff tear arthropathy (32%) in comparison to healthy control subjects (9%). These SNPs were found to influence inorganic pyrophosphate concentrations, which may alter inorganic pyrophosphate concentrations leading to calcium crystal depositions.

A study by da Motta et al (29) investigated 23 SNPs within 6 genes (*DEFB1, DENND2C, ESRRB, FGF3, FGF10* and *FGFR1*) involved in repair and the degenerative processes in patients with rotator cuff disease. *DEFB1* encodes for b-defensin which has been shown to be involved with muscle degeneration and is expressed by a wide variety of tissues as acts to regulate immune response. *ESRRB* is a nuclear receptor that has been hypothesised to promote cell survival in hypoxic environments and pluripotency and is proposed to interact with *HIF* (Hypoxic inducing factor) in order to mediate adaption to hypoxic environments (166). This interaction makes *ESRRB* essential for *HIF* to function. In terms of the rotator cuff tears, the rotator cuff insertion is a relatively hypoxic environment and therefore, increased expression of *HIF* may potentially initiate the apoptosis pathway that could result in tears and tendinopathy (140). Further, Fibroblast growth factors (*FGF*s) are an important and crucial mediator in angiogenesis and mesenchymal cell mitogenesis (192). Additionally, *FGF* expression has been shown to stimulates the production of collagen in the in various animal models (193) indicating *FGF* encoding genes are associated with collagen synthesis and turnover.

Tashjian et al. (179) conducted a genome-wide association study of 257,558 SNPs in patients with full-thickness rotator cuff tears. It was found that only two SNPs, *SAP30BP* rs820218 (A/G) and *SASHI* rs12527089 (C/T) were significantly associated with rotator cuff tears. These SNPs have been reported to be involved in cell death and the apoptosis pathway. Overexpression of *SASHI* has been shown to be associated with inhibition of growth and promotion of apoptosis (194). Further, studies suggest that *SAP30BP* is a transcriptional regulator protein, and its primary function is to induce apoptosis and cell death and may act as a co-expressor of genes related to cell survival (195). Moreover, Geiger et al (196) investigated the role of *TNXB* (Tenascin-X B) in recurrent shoulder dislocations. TNX is a member of the tenascin family of extracellular matrix

glycoproteins and has been shown to have an important role in collagen fibril deposition during collagen synthesis and fibrillogenesis (197,198). However, none of the investigated SNPs within the *TNXB* were found to be associated with risk of recurrent shoulder dislocation. These studies have been previously reviewed in the literature (21,177) and are beyond the scope of this thesis.

2.2.8 The Rotator Cuff and Swimming

Swimming is one of the most popular recreational and competitive sporting activities. Elite swimmers involved in rigorous competitive programs typically train an average of 20-30 hours per week with training sessions ranging from 7-18 km daily (199). Richardson et al. (200) estimated that the average collegiate swimmer performs more than 1 million strokes with each arm annually. Approximately 90% of the forward propulsive power is generated from the upper extremities (201) including hands, arms, shoulder and torso rotations. It is therefore unsurprising that the shoulder is the mostly commonly reported site of pain and injury (202).

2.2.9 Rotator Cuff Overuse injuries in Swimming

"Swimmer's shoulder" was first described by Kennedy and Hawkins (203) as a common, painful syndrome of repeated mechanical impingement of the shoulder and related structures. While the occurrence and seriousness of shoulder pain and injury in swimmers are well-documented, there remains a lack of consensus regarding its root cause. Some have proposed that the pervasive and diverse nature of "swimmer's shoulder" makes it improbable for a single cause to fully account for its prevalence (48). It has been proposed that due to the pervasive and varied nature of swimmer's shoulder, it is unlikely that a single pathology can adequately account for its prevalence. Therefore,

shoulder pain can be represented by an overlap of impingement syndrome complex, RCT, biceps tendinopathy, shoulder instability and shoulder subluxation (59,204,205). Wheldon and Richardson (205) suggested the three main factors contributing to shoulder injuries are: increased range of motion, increased internal rotation and adduction strength, and prolonged fatiguing of shoulder stability muscles. It has been shown that reduced stability in the shoulder joint leads to subluxation of the glenohumeral joint which when combined with repetition of overhead movements involved in the swimming stroke, lead to inflammation and pain (50). Sein et al. (59) proposed that repetitive movement that occurs during intensive training may leads to an increase in tendon thickness which is associated with tendinopathy. When the thickened tendon and associated bursa are repeatedly squashed against the bony arch of the acromion during swimming, pain and inflammation of the shoulder ensues. In a study by Tate et al. (50), it was found that swimmers who are involved in a high-level training program were found to have reduced posterior shoulder flexibility and external rotation of glenohumeral joint. The tightening of the posterior capsule produced superior and anterior humeral head translation which has been shown to reduce subacromial space during overhead upper extremity use increasing the likelihood of impingement syndrome and other shoulder pathologies (123).

In a study by Rupp et al. (56) in which 22 elite competitive swimmers were examined for shoulder pain and injuries, it was found that 14 of the 22 (64%) had a previous history of shoulder pain. Moreover, 23% reported interfering pain resulting in cessation of training. Fifty-six % reported pain during isokinetic testing, 50% were found to have impingement sign, 50% had a positive Apprehension sign, 23% had scapular winging and 55% had shoulder protrusion.

In a study by Borsa et al. (202) it was found that 27 of 42 (64%) of national collegiate athletic association (NCAA) Division I swimmers reported a history of unilateral or bilateral shoulder pain. Most these subjects (65% - 17/27) reported pain from RCT, whereas less common causes of pain consisted of biceps tendinopathy (5 subjects), non-specific causes (3 subjects), thoracic outlet syndrome (1 subject) and labral tear (1 subject). Furthermore, Brushoj et al. (206) examined 16 national level swimmers for shoulder pathologies. It was found that 11 of the 16 (61%) had labral pathology, 46% were found to have posterior supraspinatus impingement, 28% had subacromial impingement and 38% had a labral tear. In addition to examination of and diagnosis of shoulder pathology, the authors compiled a retrospective analysis of 18 swimmers who had undergone arthroscopic surgery between 1999 and 2000. The most common procedure carried out by medical staff was debridement (11 of 18), whereas partial release of the coraco-acromial ligaments was carried out in 4 of the swimmers and a bursectomy in four swimmers. It was found of the 18 injured swimmers, nine returned to pre-injury level, seven returned to the sport after a period of rehabilitation and two returned with chronic shoulder pain.

Further, in a study (58) investigating MRI changes in the shoulder of 36 competitive swimmers. The results showed that 27 of the 36 (75%) swimmers demonstrated grade I supraspinatus tendinopathy, 8 (22%) revealed grade II tendinopathy and one showed a grade III tendinopathy. The study further revealed that 14 of the 52 of the swimmers examined showed a supraspinatus thickening and three swimmers had supraspinatus tendon tears (2 swimmers were found to have delaminated instar-substance tears and one swimmer had a partial articular side tear). Less common amongst the swimmers in the study was subscapularis tendinopathy (n=2) and infraspinatus thickening (n=1). In addition, 10 of the swimmers demonstrated labral tears (8 with SLAP lesions, 1 with a SLAP lesion and Bankart lesion and 1 with a Bankhart lesion). Moreover,

changes in acromion shape were observed with 20 of 52 (38%) shoulders examined showed type I changes, 29 of 52 (56%) showing type II changes and only three of 52 (6%) showing type III changes. In addition, 33 of 52 (63%) demonstrated subacromial bursa thickening, 3 of 52 had subscapularis bursa thickening and 21 had an increase in sub acromial and subdeltoid fluid (63). Interestingly, when the swimmers were separated into their competitive level; all four international level swimmers had supraspinatus tendinopathy, 24 of 27 (89%) at the national level and 8 of the 20 (40%) state (provincial) level swimmers had supraspinatus tendinopathy. This is closely linked to the number of hours spent training (53,63). These results are further supported by Chase et al. (48) in a 12-month prospective study of Division 1 Collegiate swimmers. Chase et al reported that shoulder injuries were the most common ailment of swimmers with 46.2% of men and 33.3% of women on the team experiencing pain or injury, with shoulder impingement being the primary shoulder injury (58%). The authors noted that 58.1% of the injuries incurred during the 12-month study were due to overuse injuries. Therefore, it is reasonable to conclude that the large majority of swimmers who participate in swimming will be at risk of developing a shoulder pathology at some point in their careers. This can be linked to the number of hours spent training in the pool and the repetitive nature of the sport which in turn will place a tremendous burden on the shoulder complex. Therefore, “swimmer’s shoulder” represents an interesting phenotype for overuse injury investigation.

In addition to investigations into ligament pathology in the form of ACL ruptures, this thesis aims to add to the growing evidence of single nucleotide polymorphisms which may predispose an individual to rotator cuff pathology. The following section will review the anatomy of the knee, ACL, mechanism of injury and relevant risk factors for ACL injuries.

2.3 THE KNEE

2.3.1 Gross Anatomy of the Knee

The knee joint is a biaxial, hinge-type synovial joint (Figure 2.2) that can move through flexion, extension and in a limited capacity, medial and lateral rotation about the vertical axis (63,65,207). The articular capsule is thin and weak and offers very little support. Its primary function is to prevent friction between the femur and tibia allow smooth articulation of the tibia (63). Conversely, the fibrous capsule is exceptionally strong. It attaches to the femur superior condyles and the articular margin of the tibia (65).

The fibrous capsule is reinforced by several extracapsular and intracapsular ligaments (208). The extracapsular ligaments are located outside the fibrous capsule and include the MCL, the LCL, the patellar ligament, oblique popliteal ligament, and the arcuate popliteal ligament (207,208). The intracapsular ligaments are located within the fibrous capsule and are the ACL and PCL (209). Normal range of motion is guided by the ACL together with the PCL, LCL and MCL, whilst a tensile load is applied (36). This cross-like structure of dense and fibrous connective tissue extends from the femoral attachment on the posterior section of the inner surface of the lateral femoral condyle (63).

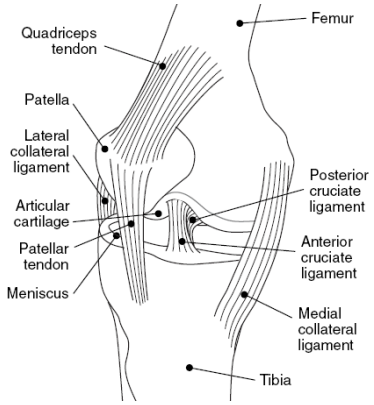


Figure 2.2. An anterior view of the right knee. The patella and patella tendon have been removed to expose the inner structures of the knee. The primary components are the ACL, PCL, MCL, medial meniscus, tibial tuberosity, lateral meniscus, and fibular/lateral collateral ligament. *Wikimedia Commons, the free media repository*. Retrieved 10:14 am, 12 February 2024 from [commons.wikimedia.org].

2.3.2 Anatomy of ACL

The primary function of the ACL is to resist and help prevent anterior translation and displacement of the tibia relative to the femur (209). The ACL is divided into two bundles; the posterolateral and the anteromedial bundle which are functionally distinct (210), comprising of collagens and other ECM proteins, tightly packed in parallel bundles (211). The two bundles work synergistically together to optimize the restraining function resisting anterior translation of the tibia. The posterolateral bundle shortens during flexion whilst the anterolateral bundle lengthens (209,212).

The collagen fibres longitudinally arranged in the ACL and have a crimped or wavy appearance (213,214). These crimps or waves straighten when enough strain is applied the ACL (215,216). Once all the collagen fibres in the ligament have been straightened, a sudden increase in stiffness is occurs to resist the applied strain and prevent tissue failure (217). Further application of strain

to the ligament after the fibres have straightened may result in rupture, either partial or complete (214,215).

2.3.3 Mechanism of Injury in ACL

Injuries to the ACL can occur because of direct or indirect forces acting on the knee during daily activity or sport (212), but ACL ruptures can through intrinsic or non-contact forces (17,20,218) which can occur whilst an athlete is participating in various movements involved in their sport (41,219). The American Orthopaedic Society of Sports Medicine (220) have created a standardized classification system based on external forces that act upon the knee. The classification system divides ACL injuries according to the type and degree of external force applied to the knee during the injury incident. A direct contact mechanism is defined as a force acting directly through direct contact with the knee. Indirect contact is defined as physical contact or force applied to the person and not the knee directly resulting in injury to the knee. A non-contact knee injury is the result of the person's own movement, such as a sudden direction change or cutting movement during running and doesn't involve any contact with another person or object (220). A fourth type of injury can occur to the ACL namely the ski-boot injury or as Boot Induced Anterior Drawer Injury (220). The resilient strain energy of the rear part of the ski generates a rotational movement, which results in the tibia migrating forward relative to the femur and excess tension is placed on the ACL which results in injury (221). For purpose of this thesis, ACL injuries will be classified using this scheme and only participants with diagnosed ACL rupture will be included.

2.3.4 Risk Factors for ACL Injury

Multiple factors have been associated with modulating the risk of ACL injury (15,17,33,38,60,222). These risk factors can be sub-divided into two broad categories: within the body or personal factors (intrinsic) and from outside the body or external factors (extrinsic). Meeuwisse's model (14) proposed that an individual or athlete may be predisposed to injury based on the presence of specific intrinsic risk factors, making them susceptible to injury. In so doing, the individual with the predisposition has a susceptibility to injury which is acted upon or influenced by the extrinsic risk factors (13,15,60,222,223). However, it is important to remember that the risk factors themselves are not causative but merely moderate the risk of injury. Instead, a specific inciting event is required and during this event, the ACL is placed under sufficient strain for it to result in rupture or injury (224,225).

Previously, in a systematic review by Posthumus et al. (17) several intrinsic risk factors that that may predispose an individual to ACL injury were critically reviewed. The authors had broadly grouped the risk factors into anatomical (225–227), neuromuscular (228–230), hormonal (15,231–234) and genetic (23,73,77,84,86,87,89,235–240). The anatomical risk factors comprised of increased quadriceps angle however, the authors found the level of certainty was low as this risk factor was represented by only a single study (17,228). Further, femoral notch geometry was identified as a potential risk factor. It was found that the femoral notch dimensions had a strong relationship with ACL injury risk and was concluded that a decreased notch width and notch width index demonstrated a high level of certainty it was associated with increased ACL rupture risk (17,227,241). Moreover, the review identified several risk factors with moderate certainty of ACL rupture risk. These include Foot pronation (242), Pelvic tilt (226), generalized joint laxity (243),

anterior knee laxity (244,245), phase of menstrual cycle in females (232), dynamic knee valgus in females (246), knee flexor and extensor pre-activation in females (247), familial disposition (237) and the *COL1A1* gene (84,248). Furthermore, the factors that were found to have low certainty of ACL rupture risk were found to be quadriceps angle (249), tibial plateau geometry (250), the *COL5A1* (73) and *COL12A1* (87) genes.

Even though ACL injury represents a multifaceted phenomenon with its precise causes and mechanisms still largely elusive, various intrinsic and extrinsic factors have been recognized as influential in modulating the risk. It is crucial to identify and characterize these factors to gain a comprehensive understanding of ACL injury risk. The forthcoming sections will delve into the molecular structure of tendons and ligaments, particularly emphasizing the significance of collagens within these musculoskeletal soft tissues.

2.4 TENDONS

2.4.1 Molecular Structure of Tendons

Tendons form part of the functional link in both the dynamic and static aspects of the musculoskeletal system by transferring the force generated by muscle contraction to the axial skeletal which results in movement (47). Tendons are predominantly composed of collagen fibres, that have rod- or spindle-shaped fibroblast-like cells known as tenocytes, which are located within a highly organised extra-cellular matrix (ECM) (251). The tenocytes produce collagen molecules which after post-translational modifications aggregate to form the basic structural unit of tendon tissue, the collagen fibril (22). The collagen polypeptide molecules form triple helices which self-

assemble into collagen fibrils that form intermolecular cross-links with adjacent helices that form collagen fibres (112). It is through this mechanism that gives tendons their tensile strength. Tendons are predominantly made up of collagen type I whereas there is a higher proportion of type III collagen in the endo- and epitenon (251). Further, an important component is collagen type V which cross-links with the other collagen types to regulate fibrillogenesis (252) and tendon diameter (103,252,253). Furthermore, tenocytes can respond to different mechanical stresses by regulating the expression of various extracellular matrix components as well as important matrix degrading enzymes (254). The highly organized structural components of tendons are crucial for the viscoelastic and non-linear response to varying tensile loads applied to it (255,256). In addition to collagen synthesis, tenocytes produce ECM components such as proteoglycans including decorin and aggrecan (257). These ECM proteins are crucial for tendon function as they serve to bind water molecules of which approximately 55% of a tendons weight comprised of water (258). The strongly hydrophilic nature of the proteoglycans facilitates the rapid diffusion of water-soluble molecules and aid in the cell migration (258–260).

2.4.2 Tendon Biomechanics

Tendons can demonstrate high mechanical strength, a flexibility as well as some elastic properties due to its molecular composition (261,262). Tendons transmit forces that are generated by muscles to bone to facilitate movement and additionally act as a buffer to limit injury to skeletal muscle by absorbing external forces (258,263,264). Therefore, the properties of the tendon are dependent on its composition which in turn effects its mechanic-biology which is determined by the number of types of intra- and inter-molecular bonds (258). A stress-strain curve aides in demonstrating the basic properties of the tendon. At rest, the tendon fibres are in a crimped or wavy configuration

(257,261). As force is applied to the tendon, the fibres begin to straighten, this first region is known as the toe region. Beyond this point, the tendon begins to deform in a linear fashion due to the collagen triple helices sliding and becoming more parallel (257,261). If the tendon strain remains below 4%, then the tendon still acts in an elastic fashion and when the strain is reduced, it returns to its normal length at approximately 4%. Once the strain progresses beyond 8-10% macroscopic failures begin to occur from intrafibrillar damage done by molecular slipping (255). Previous studies utilizing X-ray diffraction have shown that the collagen fibril elongation initially occurs due to the molecular elongation of the crimped fibres, but as the stress increases on the tendon, the gap between the molecules begin to increase. This results in the slippage of the lateral adjoining molecules and further strain leads to failure and the fibres recoil into a tangled bundle at the ruptured end (258,262). Tensile strength of tendons is therefore influenced by the thickness and collagen content (265–267). During strenuous activity, very high loads are placed on the tendon (251,262). However, the total load placed on the tendon is as important as the rate of loading of the tendon (138) which may influence the risk of rupture or injury (19,46). When tensile stress is applied quickly and obliquely, the tendon is at its highest risk (19,40,66).

2.5 LIGAMENTS

2.5.1 Molecular Biology of Ligaments

Ligaments are a dense fibrous connective tissue in the skeletal system, much denser than tendons and muscles, and serves to provide mechanical support and stability to joints, aiding in guiding the joint within their normal range of motion (215,258,268). Typically, ligaments display a non-linear anisotropic mechanical behaviour and are relatively compliant under phases of low mechanical

loads (269). This is in part due to the nature of the organised collagen fibres and viscoelastic properties of the extra cellular matrix (260). In terms of composition, ligaments are comprised predominantly of collagen molecules (~75%) with the remaining composition comprised of ECM (~25%) including proteoglycans, glycoproteins, elastin, and other proteins (270). The composition of ligamentous collagen varies depending on the specific location and its required function, but it is generally accepted that around 75-85% of the total collagen composition is accounted for by that of type I collagen with the remaining collagens being type III, VI, XI and XIV (216,271). Like tendons, fibroblasts are located between the rows of the collagen fibres where they produce the various collagen molecules and actively maintain the ECM (260,272). The molecular process to produce the collagen fibres are very similar to that of tendons in that the collagen molecules form triple helices with adjacent fibrils eventually forming densely packed fibres (273,274).

2.5.2 Ligament Biomechanics

Ligaments, like tendons, are high load bearing tissues and serve to transfer force from bone to bone in the longitudinal direction of the ligament and act to passively stabilise joints, in addition to guiding the joint through a normal range of motion (209,213,268). Thus, the properties of ligamentous tissue can be observed by executing a uniaxial tensile test of the bone to ligament to bone complex (275). Since ligaments exhibit a nonlinear anisotropic mechanical behaviour, when a low load is applied to them, they are relatively compliant and elongate, in part due to the crimped nature of the collagen fibres (276). This biomechanical feature of ligaments enables them to maintain unhindered movement under normal physiological conditions and to provide resistance and restraint of excessive movements when the joint is placed under high tensile loads (277,278). This function permits the ligament to absorb a large amount of mechanical force placed on it until

tensile failure or disruption leading to injury (279). Finally, ligaments act as sensory organs and play a crucial role in proprioception during the ligamento-muscular reflex (280,281). Histological analysis has shown that ligaments contain various mechanoreceptors; namely Pacinian corpuscles, Golgi tendon organs, and Ruffini endings which respond to various mechanical stimulus produced by movement or external forces (282,283). The stimulus detected by the mechanoreceptors activates the ligamento-muscular reflex which acts to protect the ligament by directly or indirectly modifying the load applied to it (281,284). Collagen is a major component of both tendon and ligaments and is the primary structural protein located in the extracellular space being the predominant component of various connective tissues and is the most abundant protein found in mammals(101,102,285).The following section will provide an overview of the various collagens, their structures, and their roles in musculoskeletal soft tissues.

2.6 COLLAGEN BIOLOGY

2.6.1 Overview of Collagens and their structure

The specific properties of collagen vary depending on its location, degree of mineralisation and function. Therefore, the collagen can range from rigid (bone), compliant (tendon/ligament) or has a compliance gradient as is the case of cartilage (102,285–287) and encompasses several different types collected together in a superfamily (288,289). The collagen superfamily comprises of 28 known types (types I to XXVIII) of functionally and structurally varied proteins (267,289). The most common structural feature shared by all the collagen types is the presence of three alpha polypeptide chains (Figure 2.3) in either an uninterrupted or interrupted triple helical structure (267,290,291) that can range from the majority of their structure (96% in type I collagen) to less

than 10% as is in the case of collagen type XII (267,292). Furthermore, the collagen superfamily can be subdivided into two main sub-groups based on structure and function. These groups include 1) fibril-forming collagens (fibrillar collagens) which are involved in the construction of the fibrillar structural platform of the ECM and (293), 2) the non-fibrillar collagens which incorporate the various fibril associated collagens with interrupted triple helices (FACITS), network forming collagens, short chain collagens and beaded filament collagens (294).

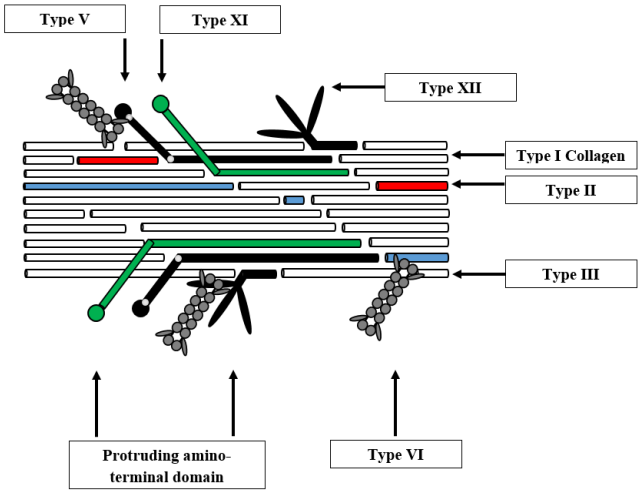


Figure 2.3. A schematic diagram of the collagen fibril showing the positions of the various collagen types. The proteins and fibril are not necessarily drawn to scale. Adapted from Collins and Raleigh (295) with permission from Karger Publishers.

2.6.2 Fibrillar Collagens

Fibrillar collagens can be into one of two sub-types, major or minor depending on their properties (294,296). The primary major type of fibrillar collagen is type I which consists of two α1(I) and one α2(I) chains. Type I collagen, found in various non-cartilage connective tissues, is primarily responsible for the maintaining the hierarchical structure and mechanical strength of the tissues

(266,267,288). Type II collagen is the principal structural protein found in cartilaginous connective tissues and is a homotrimer consisting of three α1(II) chains (296,297). Type III is the third type of major fibrillar collagen and consists of three α1(III) chains and interacts with type I collagen forming heterotypic fibrils (288,289). This interaction between type I and III collagen is believed to regulate the diameter of the fibril by limiting lateral growth during development (fibrillogenesis) and involved with tendon healing (285,294,296).

Type V collagen is a minor fibrillar collagen consisting of several different isoforms with the predominant isoform being a heterotrimer consisting of two α1(V) and one α2(V) chains (252,298,299). Further, it is believed to have a functional role in fibrillogenesis of fibrils (252,261,299,300) by means of preventing the addition of other collagen molecules (285). The second minor fibrillar collagen is type XI which is most often found in cartilaginous tissues (253,296,298). More recently, it is suggested that type V collagen interacts with type XI in a synergistic manner to regulate fibrillogenesis and growth in mature tendons (286,294). Although not classical fibrillar collagens, types XXIV and XXVII are fibrillar and are primarily found and expressed in cartilage, bone growth plates and other sites of bone to cartilage transitions (301,302). Furthermore, type XXVII has also been found in skeletal muscle tissues and interacts with other extracellular matrix proteins and molecules (75,302,303).

2.6.3 Fibril-Associated Collagen with Interrupted Triple Helices (FACITs)

FACITs are a group of significant non-fibrillar collagens that are involved in mediating various interactions between different collagen fibres, facilitating cell surface interactions (285,288,292,303), stabilisation of connected collagen triple helices (304–306) as well as cell-

matrix connections (307). This group includes types IX, XII and XIV (308–312). Type IX collagen has been shown to have a meaningful role in cartilaginous tissues (308,313,314). Type XII collagen is believed to have a similar structure and function to type XIV collagen (311,315–317) mainly expressed in tendons and ligaments (309,318). Type XIV collagen has implicated in skeletal muscle fibre regeneration (311,316), cardiac tissue (319,320) as well as in the connective tissue of blood vessels (320,321).

Other less common FACITs in the collagen superfamily are that of type XVI, XX, XXI, XIX and XXII. Type XVI collagen has been found to be expressed in smooth muscle (such that of intestinal lining and blood vessels) (322) and have been detected in fibroblasts and keratinocytes (323,324). Type XX collagen has been shown to be expressed in foetal and embryonic tendons and skin (325) as well as expressed in high quantities in corneal epithelium (326,327). Both collagen types XIX (315,328,329) and XXI (330) have been detected in in skeletal muscle. Finally, collagen type XXII has been shown to be produced by cells within the myotendinous junctions and acts a site for cell adhesion for fibroblasts and epithelial cells (331,332).

2.6.4 Non-Fibrillar Beaded Collagens

Chains of non-fibrillar beaded collagens are assembled intracellularly during development and then expressed extracellularly as a mature collagen heterotrimer (102,323)which is unique to non-fibrillar collagens. A further difference between non-fibrillar beaded collagens different to fibrillar collagen is that there is no enzymatic activity at the C- and N- terminals of the non-collagenous regions (311,333). Type VI collagen is a type of non-fibrillar beaded collagen and has been shown to be a relatively ubiquitous collagen with important roles in tissue integrity maintenance

(198,334) by creating a microfibrillar network (334,335) and helps to facilitate fibrillar interaction with ECM through various interfaces with other collagen types (336,337). More recently, type III and VI collagen were found to mediate injury risk (89). Type XXVIII collagen is one of the least commonly expressed collagen types as it is restricted to expression in the skin specifically in the peripheral nervous system (301,302).

2.6.5 Other Non-Fibrillar Collagens

Other non-fibrillar collagens include type IV (Table 2.1) which is involved in network forming and has a major role in the development of basement membranes (338,339). It functions to surround and anchor the cellular and molecular components of various tissues including muscle, tendon, ligament, and bone (305,339). Other collagens include XVIII and X which are short chain and form networks of non-cartilaginous and cartilaginous tissues within basement membranes, in addition to type VIII (329,333,340–343). Type VII is only found in the retinal tissues (344). Multiplexins, consisting of collagen types XV and XVIII (340,345–348), are produced mainly from muscle cells and fibroblasts. The function of type XVIII is unclear; however, it has been associated with ocular defects (340,341,349). XIII, XVII, XXIII and XXV are transmembrane domains collagens and have similar structures except for type XVII (350–353). Type XIII collagen is largely found during development and growth of neuromuscular junctions and the basement membrane of myotendinous structures (318). Type XXIII has been identified in the basement membranes of several tissues including the kidney, lung, skin, tendon, and cornea (354,355). Finally, type XVII although not implicated in the MSK, it is involved in melanocytes in the epidermis (340,341,352).

2.6.6 Collagen Synthesis

Fibrillar collagens are typically comprised of roughly 1000 amino acids occurring in uninterrupted repeating triplet structure (270). All collagens share the same structure of the repeating uninterrupted and interrupted triple helical domains which are essential correct formation (291,356). Most collagens form homotrimers of three identical α- chains, however heterotrimers consisting of two or three different α- chains (Figure 2.5) commonly occur (357–359). However, the process of synthesising collagen is beyond the scope of this thesis. Banos et al. (294), Exposito et al. (285), and Ricard-Blum (102) have extensively reviewed collagen synthesis.

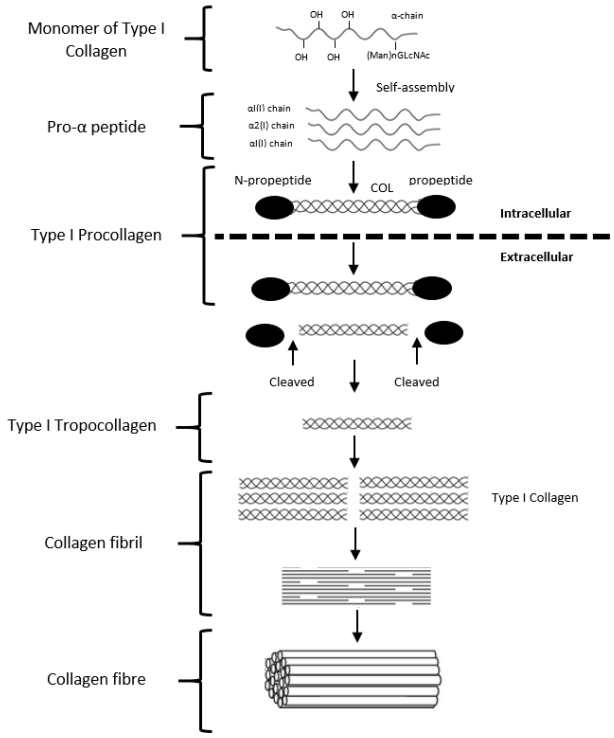


Figure 2.4. An example diagram of the fibril assembly of type I collagen. The molecules and proteins in this diagram are not necessarily drawn to scale. Adapted from Canty et al. (360) with modifications.

2.6.7 Collagen Fibril Organization

As stated previously, collagen fibres form the main structural component (Figure 2.5), acting like a scaffold in many MSK tissues (266,361) including tendons of which collagen accounts for roughly 85% of the total dry weight (266). The remaining components may vary depending on the type and location of the tissue and may include 1-5% proteoglycans and 2% elastin (101,362). There are 28 known collagen types, type I collagen account for nearly 95% of collagen fibres in various tissue structures, followed by type III collagen which accounts for around 1-5% (101,363). In the context of tendons and ligaments, other minor collagens are also found, and these include

collagen types IV, V, VI, XII and (286,293,294). Expression of these collagens vary depending on the specific portion or area of the tendon and differs according intended function, including the myotendinous junction, the tendon midsubstance and the osteotendinous junction (where the tendon connects to bone) (102,288). The collagen fibrils bundle together and form fibres which are then further bundled together to fascicles (363,364). An endotenon, a type of loose connective tissue, in the case of tendons and endoligament in ligaments, surrounds each fascicle (363). The fascicles are then grouped together and further enveloped by an epitenon (a second connective tissue sheath), resulting in a tendon (364,365). The product of this process is a highly organized hierarchical structure comprised of fibres clusters that are orientated uni-directionally along the length of the tendon (251,366).

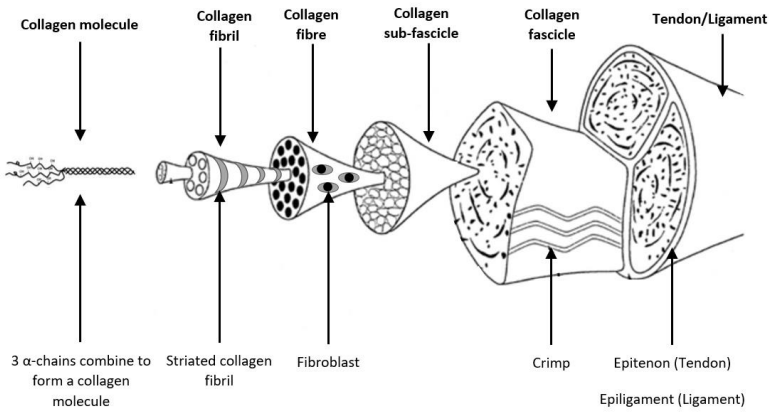


Figure 2.5. A schematic example of the hierarchical organization of tendon and ligaments. Diagram is not drawn to scale. Adapted from (367) with modifications.

It is important to note that, although tendons and ligaments vary with regards to function, both tissues have a very similar hierarchical structure (216,268). The fibres within the ligaments are much less well organized than tendons, resulting in a multi-directional orientation creating a weaved pattern in the tissue (265,288,368). Furthermore, 70-80% of a ligaments total dry mass is comprised of collagen fibres, whilst the remainder is composed of various other extracellular matrix molecules (glycoproteins, proteoglycans and glycosaminoglycans) (366,369,370). Ligaments are composed of approximately 90% of type I collagen, and depending on the location, the remaining 10% is either type III or type V collagen (209,215). An additional difference is that the collagen within a ligament remains consistent over its entire structure, contrasting to the variation collagen seen within tendons (271,356).

2.7. GENETIC VARIANTS AND COLLAGEN TYPES

2.7.1 Type I Collagen

Type I collagen is the most abundant fibrillar collagen in non-cartilaginous MSK tissues (303,365,371,372). It is encoded by two genes, namely the *COL1A1* (17q21.33), which constitutes 52 exons and is 18.34 kb in length and the *COL1A2* (7q21.3) which also is comprised of 52 exons and is 36.67kb in length (370,373). The *COL1A1* gene encodes for the alpha (α) 1 chain, while the *COL1A2* gene encodes for the α(2) chain (248,374). Mutations within the *COL1A1* gene have been linked with several pathologies including osteogenesis imperfecta (375), Ehlers-Danlos Syndrome (376) and osteoporosis (377).

Since collagen type I is the most common type of collagen, it comes as no surprise that polymorphisms within the *COL1A1* gene have been extensively investigated. More specifically, the Sp-1 binding site variant (rs1800012, G/T) (Table 1) (373,378). In a Swedish cohort (27), in which ACL injuries and shoulder dislocations were investigated, the rare TT genotype of the functional *COL1A1* Sp1 binding site polymorphism (rs1800012, G/T) was shown to be significantly under-represented in participants with ACL ruptures (Table 2.3). Only 0.4% (n=1/233) participants with ACL rupture, compared to 3.7% (n=12/325) participants control participants had a TT genotype. Although not significant, a trend was noted with the TT genotype being under-represented in participants with shoulder dislocations (0.8%) compared to controls (3.7%). Similarly, the rare TT genotype was significantly under-represented in a South Africa ACL rupture cohort (0%, p=0.031) compared to uninjured controls (4.6%) (261). In addition, the rs1800012 variant was also investigated in a South African cohort ATRs and chronic TEN (77). To further understand the contribution to risk, Collins et al. (248) combined previously published data of other acute injures (27,379) with the South African Achilles Tendon injury cohort (84) and observed an 11.1x reduction in MSK rupture risk (471,477). A further risk reduction of 15.1x was noted when only ACL ruptures were analysed. As a result, the *COL1A1* rs1800012 seems to be associated with several MSK soft tissue injuries.

Moreover, Ficek et al (90), investigated the *COL1A1* -1997G/T and +1245G/T polymorphisms and the potential risk of ACL rupture in professional soccer players. The cohort consisted of 91 male professional level soccer players and 143 apparently healthy professional soccer player controls. The authors observed that the G-T haplotype (-1997G/T and +1245G/T polymorphism) was found to be significantly under-represented in the ACL group when compared to the CON group (p=0.048). Further, Stepien-Stodkowska et al. (380) examined the +1245G/T polymorphism

in the *COL1A1* gene in Polish Male recreational skiers with ACL ruptures (n=138) and apparently healthy male skier controls (n=183). A significant difference in genotype distribution was found between the injured skiers and controls (p=0.045). However, no statistical difference in allele distribution was found (OR 1.43; 0.91-2.25, p=0.101). The authors concluded that the risk of ACL rupture was 1.43 times lower in carriers with a minor G allele as compared to carriers of the T allele. Whilst the TT genotype seems to provide a reduction in risk for various MSK injuries, it is important to note that has been linked to an increased risk of a to note several other pathologies (381–383).

However, it must also be noted that many studies investigating this specific polymorphism have failed to find any associations with injury risk including tennis elbow (lateral epicondylitis) (25), RCT (30) and Carpal Tunnel syndrome (98). Further, other polymorphisms have been investigated in the *COL1A1* and *COL1A2* genes (Table 1) but were not found to be significantly associated with risk (90,382,384). Highlighting the fact that certain gene variants may have different effects on the aetiology of pathology depending on the specific phenotype investigated. Therefore, the implication of each variant remains largely unknown and may be due to various other factors such as gene interactions with other variants, the influence and interaction with the external environment. Further research is required to fully clarify the role this polymorphism within specific phenotypes.

2.7.2 Type V Collagen

Type V collagen's major isoform is encoded by *COL5A1* and *COL5A2* genes, resulting in a heterotrimer consisting of two α1(V) and one α2(V) chains (102,370). Mutations within *COL5*

genes have been implicated in the classical form of Ehlers-Danlos Syndrome (376). Within the 3′-untranslated region (3′-UTR) of the *COL5A1* gene (rs13946 T/C, rs14776422, rs5574880 W/M, rs12722 T/C, rs3196378 C/A, rs10628678 –/AGGG, rs16399 –/ATCT, rs1134170 A/T) has been one of the most extensively studied (independently or in combination) polymorphisms in musculoskeletal injuries (Table 1).

One of the first genes found to be associated with the modulation of risk of MSK injury was the *COL5A1* gene. More specifically, the rs12722 T/C 3'-UTR variant which has been associated with several MSK phenotypes including Achilles tendon injuries (385). The CC genotype was significantly associated with a risk reduction of Achilles tendon injury in two independent South African and Australian cohorts (80,82). Several other variants, rs10628678 (-/AGGG), rs169399 (ATCT/-) and rs1134170 (A/T) have also been implicated with risk within the above cohort (386). Specifically, the *COL5A1* rs10628678 AGGG/AGGG, rs169399 -/-, and rs1134170 TT genotypes were significantly associated with increased risk of TEN. More recently, Brown et al. (80) failed to find a significant association of *COL5A1* rs12722, rs10628678 and rs3196378 with either TEN or Achilles tendon rupture (ATR). Although there were no independent associations within the *COL5A1* polymorphism, it was found that the C-A-AGGG and T-C-(-) inferred pseudo-haplotypes were significantly over-represented in the TEN and ATR groups when compared to uninjured controls. Further the C-C-(-) inferred pseudo-haplotype was found to be significantly associated with a reduction of tendinopathy risk (80).

Moreover, Burger et al. (94), whilst investigating the genetic risk for carpal tunnel syndrome (CTS), found the *COL5A1* (rs13946 C/T) TT genotype to be significantly over-represented ($p=0.007$) in the control group when compared to the carpal tunnel group. When the authors

combined the other investigated variants within the *COL5A1* gene (rs1477462 C/T, rs55748801 G/A, rs12722 C/T and the rs10628678 -/AGGG), it was found that the combined WW+CC (41.7%, p=0.008) and WW+CT (40.3%, p=0.009) were significantly over- and under-represented in the control group respectively when compared to the CTS group. Furthermore, The T-W-C (52.2%, p<0.001) and C-W-C (15.9%, p=0.005) inferred haplotypes were significantly over- and under-represented in the control groups compared to the carpal tunnel group (94). In addition, Altinsik et al. (24), demonstrated a significant association between the *COL5A1* rs12722 (C/T) and rs13946 (C/T) variants and Tennis Elbow (lateral epicondylitis) risk. The variants were investigated in a total of 349 individuals, including 154 patients with tennis elbow and 195 healthy control participants. There was a significant difference in the frequencies of genotypes (p =0.029) and alleles (p=0.030) between the Tennis Elbow and control groups where the CC genotype was significantly associated with a decreased risk. There was also a significant association in genotype and allele frequencies of rs13946 between Tennis Elbow and control groups (p=0.004 and 0.002, respectively) (24).

The study by Posthumus et al. (73) is the first study to find an association with rs12722 and anterior cruciate ruptures. The CC genotype of rs12722 (p=0.006) were significantly underrepresented in the female ACL group, when compared to the female control (CON) group. No significant effect was found in the male only subgroup. *COL5A1* rs13946 was not found to be significantly associated with risk (73).

Stepien-Stodkowska et al. (91) examined the association between the *COL5A1* rs12722 and rs13946 variants in 138 recreational male Polish Skiers with ACL ruptures and 183 uninjured male Polish Skier controls. Interestingly, the study did not find any significant associations between the

variants and injury risk. The study did find a tendency toward the C-T haplotype to be under-represented in the ACL injury group when compared to controls (91). Although, the study failed to find any significant association, it is possible to speculate the mechanism of injury to the skiers may have influenced the outcome of risk as previous studies had reported non-skiing non-contact injuries to the ACL. However, several looking at Rotator Cuff Disease (30), Lumbar Disk Degeneration (387), Achilles Pathology (80), and TEN (82) failed to find a significant association with any of the investigated *COL5A1* polymorphisms.

Collins and Posthumus (388) suggested that the proportion of type V collagen present in tendons, ligaments, and other tissues modifies the diameters of fibrils and the density at which they are packed within these tissues (refer to Figure 2.6). Consequently, this alteration affects their mechanical characteristics, thereby influencing their vulnerability to injury and other exercise-related traits. Evidence by Laguette et al. (389) demonstrated that the T-form of the *COL5A1* 3'UTR to have increased mRNA stability compared to the C-form, suggesting functional allelic variants may have a significant effect on the on-collagen structure and resultant risk.

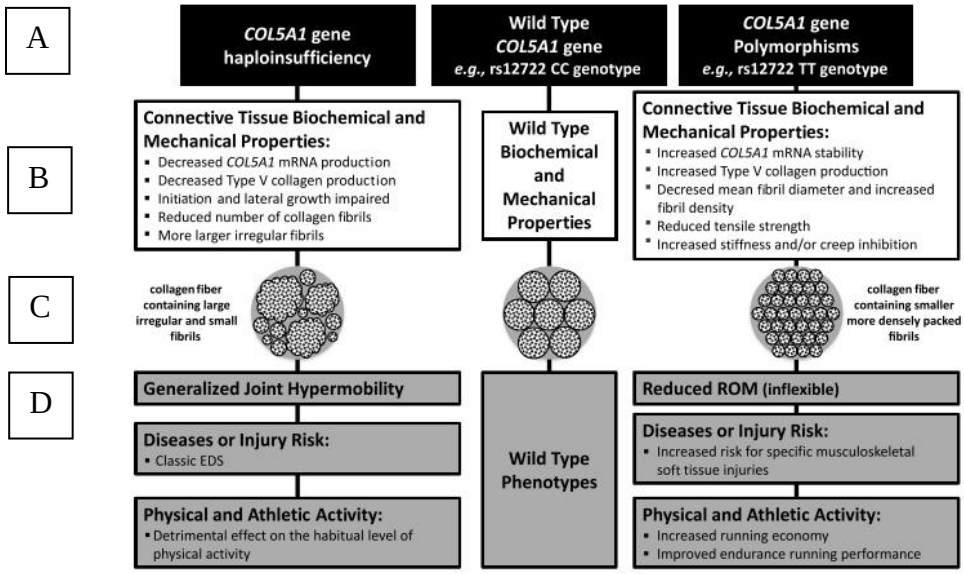


Figure 2.6. A general genetic continuum for collagen genes as previously proposed by Collins and Posthumus (1) using Type V collagen as an example. The black boxes (A) represent lethal or severe phenotypes and inherited conditions because of mutations to the collagen gene. The white boxes (B) the variants or phenotypes that would be the most beneficial. C) represents the proposed changes to the underlying tissues with regards to structure because of mutation. The grey boxes (D) represent the resultant phenotypes. Reprinted from Collins and Posthumus (388) with permission from Wolters Kluwer Health, Inc.

2.7.3 Type VI Collagen

Type VI collagen is a ubiquitous and unusual nonfibrillar collagen encoded by six different genes (*COL6A1*, *COL6A2*, *COL6A3*, *COL6A4*, *COL6A5* and *COL6A6*) (390,391). Type VI collagen serves as a bridging molecule by binding several components of the ECM, and the surrounding skeletal muscle, bone and cartilage connective tissue (390). The collagen chains are organized into a network of microfibrils important in anchoring the basement membrane to the extracellular matrix (ECM) (294,338) and is known to interact with types I, V and XII collagens to facilitate normal collagen fibrillogenesis (198,390,392). Mutations within the *COL6A1*, *COL6A2* and

COL6A3 genes have been shown to associated with congenital muscular dystrophies (393) such as Ulrich's congenital muscular dystrophy (394) and Bethlem myopathy (393), characterized by distal joint laxity and a combination of distal and proximal joint contractures (393,395,396).

Moreover, the *COL6A1* gene, specifically, a variant within intron 32 (rs35796750 T/C), has been previously implicated in several other conditions, including posterior longitudinal ligament (395) ligamentum flavum (397) ossification. The rs35796750 has also been associated with cycling performance (398). More recently, a study by O'Connell et al. (89), investigated if *COL3A1* (rs1800255 G/A), *COL5A1* (rs12722 T/C), *COL6A1* (rs35796750 T/C) and *COL12A1* rs970547 A/G were involved in modulating ACL injury risk in a South African and Polish cohort. Interestingly, the *COL6A1* (rs35796750 T/C) was not independently associated with ACL injury risk in both South African ACL injury participants and therefore was not investigated in the Polish cohort. However, when the rs35796750 variant was shown to have an interaction with the *COL3A1* *COL5A1* genes resulting in a modulation of risk. The A-C-C (*COL3A1* rs1800255 G/A, *COL5A1* rs12722 T/C and (*COL6A1* rs35796750 T/C) and A-C (*COL3A1* rs1800255 G/A and *COL6A1* rs35796750 T/C) inferred pseudo haplotypes were significantly associated with decreased risk of TEN in a combined South African and Australian cohort (89). Further research is required in order to fully understand type VI collagen and risk of musculoskeletal injury.

2.7.4 Type XI Collagen

Type XI collagen is an important minor fibrillar collagen which consists of α1(XI), α2(XI) and α3(XI) chains encoded by the *COL11A1*, *COL11A2* and *COL11A3* respectively. It is believed to

play a role in fibrillogenesis of tendons (103). Further, these proteins share structural and functional with type V collagen in various connective tissues (253). These genes have been previously indicated in several inherited disorders and musculoskeletal injuries including lumbar disc degeneration (399,400). In a study by Mio et al. (400), it was found that the CC genotype was significant associated with decreased risk of Lumbar Disk Degeneration. Recently, Hay et al (84) found no independent associations between *COL11A1*, *COL11A2* genes with TEN, however, when the rs3753841 (T/C), rs1676486 (C/T) and rs1799907 (T/A) SNPs were constructed into a pseudo haplotype, the T-C-T pseudo haplotype was significantly associated with increased risk (p=0.006) of TEN. Additionally, the T-C-T-(AGGG) pseudo haplotypes constructed using these type XI (76) polymorphisms and the functional *COL5A1* rs17146744 (-/AGGG) polymorphism was also found to be significantly associated with increased risk. Although each of the type XI genes were not independently associated with injury risk, the structural and functionally similar genes encoding for type XI and type V collagens interact with one another and can potentially modulate risk for injury. More recently, Dada et al. (98) found that the TT genotype of the *COL11A1* rs3753841 (T/C) was significantly associated with increased risk of CTS. Additionally, the T-C inferred pseudo haplotype was associated with increased risk of CTS.

2.7.5 Type XII and XIV Collagen

Type XII collagen is a member of the FACIT family of collagens (267,401) and functions to mediate the interactions between collagen fibres and other components of the ECM (309,402). Type XII is believed to have a similar function to type V collagen, both believed to regulate

fibrillogenesis (403). The *COL12A1* gene encodes for the α1(XII) chain of which two distinct homotrimeric isoforms exist, namely the long (XIIA) and short (XIIB) isoforms (289). In the context of tendons and ligaments, the short isoform is predominantly expressed in response to mechanical loading (289). Further, type XIV collagen and is encoded by the *COL14A1* which produces which produces a single homotrimer α1(XIV) collagen. It is considered a FACIT due to its interaction with collagen I (404). It plays a similar role to other minor collagens, such as collagen type III, V, XI and XII by regulating the process of linear and lateral growth (261) especially in tissues with mechanical demand (319). Previously it has been shown that type XIV collagen integrates fibrils in fibres during development and is especially prevalent during development of immature tendons (316). It has been hypothesized that type XII collagen replaces type XIV both functionally and structurally in later development as the tendon and ligament matures (316).

September et al. (74) found no significant associations between *COL12A1* rs240736 (T/C), rs970547 (G/A), *COL14A1* rs4870723 (A/C), rs1563392 (T/A) and risk of TEN. Interestingly, the *COL12A1* (rs970547 A/G) gene was implicated with ACL ruptures within a South African European Ancestry population (87) but only after stratification by sex. The AA genotype was significantly (AA vs. GT + GG; OR=2.4; 95% CI 1.0 - 5.5; p=0.048) associated with ACL injury risk in females but no significant association in males. More recently, a study by O'Connell et al. (89), investigated if *COL3A1* (rs1800255 G/A), *COL5A1* (rs12722 T/C), *COL6A1* (rs35796750 T/C) and *COL12A1* rs970547 A/G were involved in modulating ACL injury risk in a South African and Polish cohort, however no significant independent associations were noted between *COL6A1* rs35796750 and *COL3A1* rs1800255 genotypes. Interestingly, only the *COL3A1* AA genotype was significantly associated with ACL injury risk in the Polish cohort (p = 0.036). When inferred

pseudo-haplotype were constructed (Table 2.4), the T-A haplotype constructed from the *COL5A1* rs12722 and *COL12A1* rs970547 was significantly associated with ACL rupture risk when stratified by sex. When females were analysed separately, the T-A haplotype was significantly over-represented in the ACL group in the South African (ACL 50.5%, CON 38.1%, p = 0.022), Polish (ACL 56.3%, CON 36.3%, p = 0.029) and combined (ACL 51.8%, T+A CON 37.5%, p = 0.004) cohorts when compared to the CON group (89). This further emphasises the importance of investigating gene-gene interactions that may modulate risk for musculoskeletal injuries.

Khoschnau et al. (27) investigating *COL1A1* Sp1 binding site polymorphism and the risk of ACL ruptures and shoulder dislocations, is the first to examine the role of collagen genes in shoulder pathology. The TT genotype was significantly associated with reduced risk compared with the controls (OR, 0.15; 95% CI, 0.03-0.68). Previously, the TT genotype has been associated with decreased risk of ACL rupture in three independent cohorts (73,91,385). The *COL5A1* 3'-UTR rs12722 (T/C) had been shown to alter the risk of ACL rupture in females (73), TEN (82) and Tennis elbow (24) where the CC genotype conferring a reduction of risk. This polymorphism also interacted with other polymorphisms within the *COL5A1* 3'-UTR to reduce risk in CTS (94).

Furthermore, it has been demonstrated that the AGGGAGGG genotype of the *COL5A1* rs10628678 (-/AGGG), also in the 3'-UTR, was significantly associated with increased risk of TEN (386). In addition, the *COL6A1* rs35796750 (T/C) has been implicated in lumbar disc herniation whereby the TT genotype was associated with increased risk (397). Further, O'Connell et al. (89) found that the A-C haplotype constructed from *COL3A1* rs1800025 (A/G) and the *COL6A1* rs35796750 (T/C) polymorphisms was associated with decreased risk of TEN. Furthermore, the AA genotype of the *COL12A1* rs970547 (G/A) was shown to be significantly

over-represented in ACL rupture cohort in females (87). Finally, although not independently associated with risk, the haplotypes constructed *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) demonstrate that there is a risk modulating effect. Hay et al. (76) found the T-C-T pseudo haplotype to be associated with increased risk of TEN. This result is further supported by Dada et al (97) in a CTS cohort in which the T-C haplotype constructed from *COL11A1* rs3753841 and *COL11A1* rs167648 was associated with increased risk. A summary of previously investigated phenotypes are presented in Table 2.1 (independent genetic association) and Table 2.2 (gene-gene interactions).

Table 2.1. Summary of independent genetic associations in several phenotypes from the literature

	COL1A1 rs1800012	*COL5A1* rs12722	*COL5A1* rs10628678	*COL11A1* rs3754841	*COL11A1* rs16746744	*COL11A2* rs1799907	*COL6A1* rs35796540	*COL12A1* rs970547
RCT								
SST		CC (30)						
SD	TT (27)							
RC	N.S	N.S						
ACL	TT (84) GG (380)	CC (73)						AA (87,405)
AT	TT (77)	N.S	AGGG/AGGG (389)					
ATEN	N.S	CC (80)		N.S	N.S	N.S	N.S	N.S
CTS	N.S	CC (94)	N.S	TT (98)		N.S		N.S
LDD	TT (383,399,406) GT (387)	N.S			CC (400)	N.S	N.S	
SLL							N.S	
TE	N.S	TT (24)		CT (26)	N.S			
AI		TT (407)						
TMJ		CT (408)						

RCT: Rotator Cuff Tendinopathy, Shoulder dislocation, SST; Supraspinatus Tendinopathy, RC; rotator cuff, ACL; Anterior Cruciate Ligament rupture, AT; Achilles Tendon rupture, ATEN; Achilles Tendinopathy, CTS; Carpal tunnel syndrome, LDD; Lumbar Disk Degeneration, SLL; spinal longitudinal ligament, TE; Tennis elbow; AI; Ankle Instability, TMJ; Temporomandibular Joint, N.S; Not significantly associated. Green highlight indicates reduced risk, red highlight indicates increased risk. Grey highlight indicates variants investigated in this thesis Bracket refers to specific reference.

Table 2.2. Summary of MSK injuries and previously associated haplotypes from the literature.

MSK Injury	*COL11A1* rs3753841 (T/C)	*COL11A1* rs1676486 (C/T)	*COL11A2* rs1799907 (A/T)	*COL5A1* rs12722 (T/C)	*COL5A1* rs10628678 (AGGG/-)	*COL6A1* rs35796750 (C/T)	*COL12A1* rs970547 (A/G)	Ref.
ATP, ATR$^{\gamma}$				C	-			(80)
ATP, ATR$^{\gamma}$				T	-			(80)
ATP, TEN, ATR$^{\gamma}$				C	AGGG			(80)
CONSENSUS				**C**	**-**			
				T/C	**AGGG/-**			
AT	T	C	T					(76)
ACL, NON (A+M)	C	T	A					(409)
ACL, NON (A+M)	C	T						(409)
CTS	C	C						(98)
AT, CTS	T	C						(409)
AT		C	A					(409)
AT		C	T					(409)
AT		T	T					(409)
AT	T		T					(409)
CONSENSUS	**C**	**T/C**	**A/T**					
	T	**C**	**T**					
AT (A+M)	T	C	T		AGGG			(84)
ACL	T	C	T		AGGG			(409)
ACL, NON (M)	T	C	T		-			(409)
NON (A)	T	C	A		AGGG			(409)
AT	C	C	A		-			(409)
AT	C	T	A		-			(409)
ACL	C	T	A		AGGG			(409)
CTS	C	C			AGGG			(98)
CTS	T	C			-			(98)
CTS	T	C			AGGG			(98)
TEN	T				AGGG			(409)
TEN	T				-			(409)
TEN	C				AGGG			(409)
TEN	C				-			(409)
TEN		C			AGGG			(409)
TEN		C			-			(409)
TEN		T			-			(409)
TEN			A		-			(409)
TEN			T		AGGG			(409)
CONSENSUS	**C**	**T/C**	**A**		**AGGG/-**			
	T	**C**	**T/A**		**AGGG**			
ACL (F)				T			A	(89)
EAMC				T		C		(410)
EAMC				C		T		(410)
TEN$^{\Phi}$				C		C		(411)
ROM				C		T		(89)
ROM				T		T		(89)
CONSENSUS				**C**		**T/C**		
				T		**T/C**	**A**	

ATP; Achilles Tendon Pathology, ATR; Achilles Tendon Rupture, TEN; Achilles Tendinopathy, ACL; Anterior Cruciate Ligament Rupture, NON; Non-contact ACL Rupture, CTS; Carpal Tunnel Syndrome, RCT; Rotator Cuff Tendinopathy, SST; Supraspinatus Tendinopathy, EAMC; Exercise Associated Muscle Cramps, ROM; Range of Motion. Green colour indicates protection, Red colour indicates risk. $^{\gamma}$: Haplotype constructed with *COL5A1* rs3196378 (C/A) $^{\Phi}$: Haplotype constructed with *COL3A1* rs1800255 (G/A) A: All participants, F: Female participants, M: Male participants.

In summary, since it has demonstrated that *COL1A1*, *COL5A1*, *COL6A1* and *COL12A1* variants have been independently associated with risk of various pathologies (ACL rupture, TEN, Tennis Elbow, shoulder dislocations, shoulder instability and/or RCT). In addition, various collagen gene-gene interactions such as those within *COL11A1* and *COL11A2* variants within CTS and TEN, it may be hypothesized that these collagen gene polymorphisms independently or interaction to modulator the risk of RCT. Therefore, the objective to this thesis was to investigate the role collagen gene polymorphism and risk. The primary, secondary and tertiary aims of this study within the context of this thesis were therefore to determine if *COL1A1* rs1800012 G/T, *COL5A1* rs12722 T/C, *COL5A1* rs10628678 -/AGGG, *COL6A1* rs35796750 T/C, *COL11A1* rs3753841 T/C, *COL11A1* rs1676486 C/T, *COL11A2* rs1799907 A/T and *COL12A1* rs970547 G/A are independently associated with risk of RCT in a South African swimming cohort (Chapter 4), ACL rupture risk in a Swedish and Combined European Ancestry cohort (Chapter 5), and ACL rupture risk in a South African mixed ancestry cohort (Chapter 6). Further, gene-gene interactions between collagen variants and risk will be explored in each cohort. Based on previous findings (Chapter 2), it is hypothesised that the following genotypes and/or alleles will either be independently associated with increased risk of and/or the allele will contribute to increased risk in gene-gene interactions.

2.8 AIMS OF THESIS

To investigate eight previously identified polymorphisms within six different genes, that encode for structural collagen proteins using a candidate gene case control genetic association study approach within three independent cohorts (Chapter 3, 4 and 5, respectively).

The objectives to address these aims were to investigate whether the *COL1A1* rs1800012, *COL5A1* rs12722, *COL5A1* rs10628678, *COL6A1* rs35796750, *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 and *COL12A1* rs970547 gene variants are risk factors, either independently of via gene-gene interactions, for RCT in a South African cohort (Chapter 4), and/or ACL ruptures in a Swedish and combined European Ancestry cohort (Chapter 5) and a South African mixed ancestry cohort (Chapter 6).

CHAPTER 3: NON-GENETIC RISK FACTORS FOR SHOULDER INJURY IN SWIMMERS

The data presented in this chapter has been published in a condensed form in the peer-reviewed article: Hill, L., Posthumus, M., Collins, M. (2015). Risk Factors Associated with Shoulder Pain and Injury in Swimmers: A Critical Systematic Review. The Physician and Sports Medicine. 43 (4):412-420.

3.1 INTRODUCTION

As discussed in the previous chapter both swimming associated shoulder injuries and ACL ruptures are multi-factorial conditions consisting of genetic, which is the focus of the research chapters of this thesis, and non-genetic components. The non-genetic risk factors for ACL rupture have been extensively reviewed (17,20,212,412) and will therefore not be repeated in this thesis. Although several risk factors have been identified, the level of evidence (413) from individual studies, as well as the level of certainty (414) that these factors predispose a swimmer to pain and injury, to our knowledge has yet to be critically evaluated in a systematic review. Therefore, the primary objective of this chapter is to conduct a systematic review to critically assess the published evidence for risk factors that may predispose a swimmer to shoulder pain and injury.

3.2 METHODS

3.2.1 Search Strategy

Published articles that examined potential risk factors for shoulder injuries in swimmers were reviewed using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (415). Two rounds of reviewing the literature, using the three-step method in each round, was performed. The first round pertained to the review of articles from the search results of the databases and the second round involved reviewing the bibliography lists of identified articles from the first round (55). Articles were excluded if they met the exclusion criteria (Table 3.1). Thereafter, the identified articles were critically appraised according to the inclusion criteria outlined in Table 3.2.

Table 3.1. Exclusion Criteria

1	Commentaries, book chapters, letters, editorials, conference proceedings, case reports, conference, abstracts, or non-peer reviewed articles
2	Studies not conducted in Swimmers
3	Studies examining shoulder/upper limb injuries without reference to shoulder pain or injury
4	Studies of other medical/systemic conditions (e.g., diabetes, amyloidosis) without specific reference to shoulder pain/injuries
5	Animal or cadaver studies

Table 3.2. Inclusion Criteria

1	The article must include original data
2	The article must be published in English
3	The article must include a minimum of one potential risk factor for Shoulder pain or Injury in Swimmers
4	Provide a definition of pain/injury in the shoulder
5	The article must include an association with the 95% CI, $p<0.05$

An electronic database search of SpringerLink, Science Direct and PubMed/Medline was conducted using the search terms "*(Injury OR pain) AND (risk* OR incidence) AND (Swim* OR Triath*)"*, *"(Injury OR pain) AND (Swim*)"* and *"(Shoulder) AND (Swim*)"*. Initially 2731 unique articles (1179 SpringerLink, 909 Science Direct and 643 PubMed) were identified during the search of the electronic databases during 2015. This search was limited to articles between January 1985 and December 2014. After duplicates were removed, 440 articles were screened by title, then 158 abstracts and 132 full texts, resulting in 94 unique articles. The second round consisted of the bibliography stage, whereby 953 articles were identified from the reference lists of the 94 previously identified articles. Thereafter, screening followed the same process and 854 titles, 311 abstracts and 171 full texts were evaluated resulting in only four unique articles being identified combining for a total of 98 articles. Of the 98 potential articles, only 29 were remained after the inclusion criteria was applied (55).

An additional 244 unique articles (121 SpringerLink, 68 Science Direct and 55 PubMed) were identified during a more recent search of the same electronic databases using the same search term during between January 2015 and March 2023 (Figure 3.1). Following the same protocol, after checking for duplication and previously included studies, 97 titles, 52 abstracts and 48 full texts were screened resulting in 41 unique articles. The bibliography round identified 287 articles from the 41 previously identified studies. After removal of duplicates, 262 titles, 104 abstracts, and 18 full texts were screened resulting in only one unique study. Therefore, a total of 42 unique articles were identified. Following application of the inclusion criteria, a final total of 16 unique articles were identified and critically appraised. These articles were either added to already identified themes or resulted in the new themes being generated.

The exclusion criteria for the identified articles are as follows, (1) all commentaries, book chapters, letters, editorials, conference proceedings, case reports, conference, abstracts, or non-peer reviewed articles. (2) studies that were not conducted in swimmers, (3) articles that did not include at least one potential risk factor for shoulder pain or injury in swimmers, (4) studies examining shoulder/upper limb injuries without reference to shoulder pain or injury, (5) that did not provide a definition of pain/injury in the shoulder (6) studies of other medical/systemic conditions (such as diabetes, amyloidosis) without specific reference to shoulder pain/injuries, (7) animal or cadaver studies and (8) not published in English were also excluded.

After inclusion criteria was applied, a total of 45 unique articles were in this systematic review. Identified articles were grouped into specific risk factors. The titles, abstracts and full texts were independently screened by Lee Hill and Dr Michael Posthumus. Articles were excluded if they were unrelated to the topic or met the exclusion criteria outlined in Table 3.1. Only risk factors with greater than one investigating manuscript were included in the results section under a specific subheading. A summary of the systematic review process is shown in Figure 3.1.

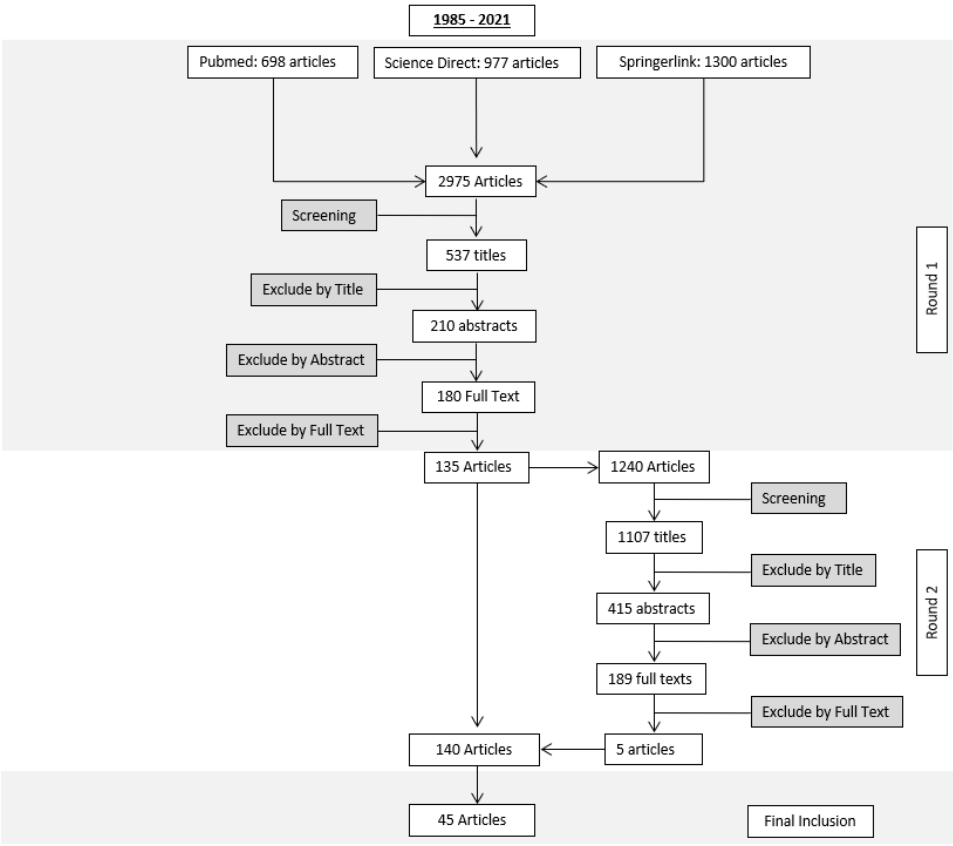


Figure 3.1. Schematic Diagram of Systematic Review Process. The first review included studies identified between 1985 and 2014 including the updated review from 2015 to 2023. All articles underwent the same screening process outline above. A total of 45 unique articles from both periods were included in the study.

3.2.2 Level of Evidence and Certainty

The 45 identified articles, 29 from the original search and 16 from the more recent search, were included in this reviewed and appraised using two established methods, level of evidence and level of certainty (413,414), as recently applied in similar systematic reviews (17,416). According to the hierarchy of evidence, randomized controlled trials and high-quality (All patients were enrolled at the same point in their disease course (inception cohort) with ≥80% follow-up of enrolled patients,

large sample sizes, robust methodology) prospective cohort studies are considered to be level I; lower-quality (small sample sizes and weaker methodology prospective studies and retrospective studies are considered to be level II; case-control and cross sectional studies are considered to be level III; case series studies are level IV; and expert opinions are level V (Table 3.3). Lee Hill and Dr Michael Posthumus independently ranked each included study according to these criteria. We agreed on the level of evidence assigned to each study.

Table 3.3. Level of evidence (I–V) definitions used for study evaluation.

Level of Evidence	Study Types
I	Randomised controlled trials and high-quality (large sample sizes, robust methodology) prospective cohort studies
II	Lower quality (small sample sizes, weaker methodology) prospective and retrospective cohort studies
III	Case–control
IV	Case series
V	Expert opinions

In addition to evaluating each studies level of evidence, a level of certainty to the risk factor being associated with shoulder injuries in swimming, was also incorporated into this review. Lee Hill and Dr Michael Posthumus ranked each identified risk factor according to the US Preventative Service Task Force (414) as either a high, moderate or low level of certainty. We agreed on the classification of each risk factor. This classification method of the risk factors identified in this review allows for a relative measure of strength that an identified risk factor is associated with shoulder pain in swimmers. The levels are defined as follows (Shown in Table 3.4).

Table 3.4. Level of Certainty (high-low) definitions used for risk assessment.

Level of Certainty	Definition
High	The available evidence includes consistent results from level I studies. These studies provide a good estimate of risk and are unlikely to be strongly affected by future studies.
Moderate	The available evidence includes sufficient evidence to determine that there is risk associated with the injury, but confidence in the estimate is constrained by factors such as the sample size and quality of studies, as well as Inconsistency of findings across individual studies. As more information becomes available, the magnitude of risk could change or even alter the conclusion.
Low	The available evidence is insufficient to assess risk. Evidence is insufficient because of the limited number or size of studies, and inconsistency of findings across individual studies. More information may allow an estimation of risk.

3.3 RESULTS AND DISCUSSION

The article selection process is outlined in Figure 3.1. In total, the 45 studies were identified and included in this review. Study designs of the 45 included studies are summarised in Table 3.5. The included critically appraised articles identified 28 risk factors for shoulder injuries in swimmers which are discussed below, summarized in Table 3.6 and supplementary Tables S3.1 in Appendix B. The risk factors identified from the systematic literature search were categorized into four categories based on factors identified during the review process, namely shoulder joint function and strength, activity history, demographics and musculoskeletal anatomy.

Table 3.5. Summary of included studies from the periods 1985-2015 and 2015-2023.

		Number of studies			
Study Design	Study Level	1985-2015	Ref.	2015-2023	Ref.
RCT	Level 1	1	(417)	0	
Prospective	Level 1	2	(48,54)	1	(418)
	Level 2	1	(419)	2	(420,421)
Retrospective	Level 2	3	(62,422)	1	(423)
Cross-Sectional	Level 3	13	(51,53,57,59,424–432)	8	(433–440)
Case Control	Level 3	9	(56,202,441–447)	4	(448–451)
Total		29		16	

Ref. – Reference, RCT – Randomised control trial

Table 3.6. Summary of the level of certainty of risk factors associated with shoulder injuries in swimmers.

Level of Certainty		
High	**Moderate**	**Low**
Shoulder Joint Function and Strength		
None	Internal/External Rotation	Shoulder Flexibility
	Clinical Joint laxity and Instability	Glenohumeral Translation
	Internal/External Rotation Strength	
Training Activity History		
None	None	Training load, volume, and intensity
		Stroke distance and stroke specialty
		Years of Swimming Experience
		Breathing side
		Swimming training equipment
		Cross Training and Stretching
Demographics		
None	Previous history of pain and injury	Sex
		Age
		Competitive level
Musculoskeletal Anatomy		
None	None	Scapular Kinematics, strength and dyskinesis
		Core Stability
		Pectoral Length
		Sub-Acromial Bursa thickness
Single Study Risk Factors		
		Swimming Scholarship Status
		Triceps Length
		Latissimus Length
		Lower Trapezius Thickness
		Inadequate treatment
		Internal/External Rotation Endurance
		Serratus Anterior Thickness
		Arm Span
		Acute: Chronic Workload Ratio

All single study risk factors were automatically assigned a low level of certainty.

3.3.1 Shoulder Joint Function and Strength

3.3.1.1 Internal/External Rotation

Six studies investigated whether there Internal/External Rotation was associated with shoulder pain in swimmers (Appendix B: Table S3.1). Three studies found an association between Internal/External Rotation and shoulder pain (54,424,431) and three studies found no association (56,427,445). In a prospective study by Walker et al (54), increased external range of motion and decreased internal range of motion were associated with significant interfering shoulder pain and significant shoulder injury, after adjusting for swim training exposure. The study divided participants into either low, mid or high external rotation range of motion. The adjusted odds ratios indicated that swimmers in both the low and high external rotation range of motion groups were at increased risk of developing a shoulder injury than those in the mid-range group. Similarly, a cross sectional study (424) reported a significant reduction in internal rotation range of motion in the 8–11-year-old swimmers with shoulder pain compared to those swimmers without pain. However, shoulder range of motion was not associated with shoulder pain in any of the other age groups examined (12-14, 15-18 and Masters). Furthermore, a cross sectional study by Bansal et al (431) found that swimmers with shoulder impingement pain had significantly decreased internal range of motion and a significantly increased external range of motion compared to swimmers without impingement pain. Although the term impingement pain has been abandoned due to its generality and poor correlation to what is occurring in the shoulder, Bansal et al (431) used specific criteria to diagnose impingement pain; this includes 1) history of exercise-related shoulder pain, 2) positive Neer's or Hawkins' impingement test and 3) presence of any of the following – painful arc, tenderness at the greater tuberosity or bicipital groove, painful active contraction or any rotator

cuff muscle. However, a case control (451) and three cross sectional studies (427,440) failed to find a significant association between shoulder range of motion and shoulder pain. Sample sizes were however limited in these studies that failed to report any associations. In summary, there is sufficient evidence that shoulder internal and external rotation range of motion is a risk factor for shoulder pain in swimmers. The confidence in the estimate is however constrained by factors such as sample size, study quality, and inconsistency of findings across individual studies. Therefore, the level of certainty that this risk factor is associated with shoulder pain is moderate.

3.3.1.2 Clinical Joint Laxity and Instability

Six studies examined whether there was an association between Clinical Joint Laxity and Instability and shoulder pain. Three studies found an association between Clinical Joint Laxity and Instability and shoulder pain (424,425,431) and three studies failed to show and association (54,202,443). A cross sectional study by McMaster et al (425) found a significant correlation between a clinical shoulder exam score for joint laxity and shoulder pain in elite swimmers who reported having shoulder pain. Furthermore, Tate et al (424) found that shoulder pain was significantly associated with self-reported shoulder instability in swimmers between the ages of 12 and 19 years old but not in Masters swimmers. In addition, the cross-sectional study by Bansal et al (431) found that atraumatic anterior instability was significantly associated with shoulder impingement pain in swimmers with impingement syndrome, but multidirectional instability was not found to be significantly associated with impingement pain. Bansal et al. (431) defined atraumatic anterior instability as a positive Apprehension sign when the tested arm was abducted to 90 degrees and externally rotated. A positive test suggests that there is increased anterior translation of the humeral head with concurrent posterior capsule tightness leading to anterior

instability. Whereas multidirectional instability is characterized as generalized capsular laxity and is considered to be atraumatic. There is therefore excessive mobility of the glenohumeral joint in all directions (anterior, posterior, superior and inferior). The tests used were the relocation test with a Sulcus sign and this would constitute a positive test for multidirectional instability. However, two case control studies (202,443) and a prospective study (54) failed to show that shoulder joint laxity was significantly associated with shoulder pain. In summary, the available evidence suggests that joint laxity and instability is a risk factor for shoulder pain in swimmers, but confidence is constrained by the quality of studies, as well as inconsistency of findings across individual studies. Therefore, the level of certainty that this risk factor is associated with shoulder pain is moderate.

3.3.1.3 Internal/External Rotation Strength

Ten studies examined whether Internal/External Rotation Strength was associated with shoulder pain. Five studies found an association between Internal/External Rotation Strength and shoulder pain (418–420,424,437) and five studies failed to show an association (56,435,442,445,450). In a large two-year prospective study by Feijen et al (418), highlight the importance of internal/external muscle rotation strength and endurance in shoulder pain and injury. The study found that for each 1-count increase in posterior muscle endurance, the odds of developing shoulder pain decreased by 5%. Furthermore, a prospective study by Drigny et al (420) observed that swimmers with a lower eccentric external rotation: concentric internal rotation ratio were 4.5-fold at risk for developing a shoulder injury. This is supported by McLaine et al (437) that found that a decrease in relative extension strength were significantly at risk than for developing shoulder pain. A cross sectional study (424) reported that reduced internal rotation torque was associated with shoulder

pain and disability in swimmers with shoulder pain. Although it does not provide evidence that decreased internal/external rotation strength is a risk factor for injury, a randomized control trial by Swanik et al (419) found a significant decrease in the incidence of shoulder pain between the experimental group who underwent shoulder functional training to strengthen the internal and external rotators compared to the control group. However, three case control studies (56,442,450)and two cross sectional study (435,445) failed to find an association between internal and external rotation strength and shoulder pain. In summary, there is sufficient evidence that isometric measurements are associated with risk of shoulder pain and injury. As a result, our level of certainty that isometric measurements are associated with risk of shoulder pain and injury is moderate.

3.3.1.4 Shoulder Flexibility

Three studies examined whether shoulder flexibility was associated with shoulder pain. Two studies found an association between shoulder flexibility and shoulder pain (426,447) and one study didn't find an association (428). In a prospective study by Greipp (426) anterior shoulder inflexibility was associated with shoulder pain in male and female swimmers. Similarly, a case control study by Ozcaldiran (447), found that swimmers with shoulder pain had a significant positive correlation between total flexibility index scores and shoulder pain. However, there was no significant association between swimmers who reported shoulder pain and subjective level of joint flexibility in a cross-sectional study (428). There is currently a lack of evidence and consistency between findings to suggest that shoulder flexibility is a risk factor for shoulder pain in swimmers. The level of certainty that this risk factor is associated with shoulder pain is therefore low.

3.3.1.5 Glenohumeral Translation

Only two studies examined whether glenohumeral translation was associated with shoulder pain. Both studies found no association between glenohumeral translation and shoulder pain (202,417). In a case control study by Borsa et al (202), no significant difference in glenohumeral joint displacement between swimmers and controls was found. In addition, no difference between joint displacement and swimmers with a history of shoulder pain was found. Furthermore, in a randomized intervention trial by Lynch et al (417), a significant decrease in forward shoulder translation and forward head angle was found in the intervention group but not in the control group. In the intervention group, a significant increase in shoulder girdle strength was found following the intervention. However, no significant interaction was found in pain score following the intervention and reduction of forward shoulder translation and forward head angle. In summary, there is currently insufficient evidence to suggest that glenohumeral displacement is a risk factor for shoulder pain in swimmers; therefore, the level of certainty is low.

3.3.2 Activity History

3.3.2.1 Training load, Volume and Intensity

Ten studies examined whether Training load, Volume and Intensity was associated with shoulder pain. Four studies found an association between Training load, Volume and Intensity and shoulder pain (59,422,424,429) and six studies didn't find an association (48,54,429,444–446). A cross sectional MRI study by Sein et al (59) reported a significant correlation between training load (number of hours swum per week and weekly mileage) and increased risk of shoulder impingement

pain. This study also showed that supraspinatus tendon thickening was significantly correlated with increased hours of training per week, weekly mileage and cumulative shoulder use. Furthermore, two cross sectional studies (50,429) showed that swimmers who swam a greater number of hours and mileage reported a significantly higher incidence of shoulder pain. Moreover, the retrospective study by Ristolainen et al (422)showed that injured swimmers had reported swimming significantly more during the 12-month period than the uninjured swimmers and the mean number of kilometres swum was higher in swimmers with at least one injury joint compared to swimmers without an injury. However, the case control study by Su and Colleagues (444) found no significant difference between practice duration (h/wk.) or practice distance (km/wk.) and shoulder impingement and pain in swimmers. Additionally, two recent prospective studies (48,54) reported that mileage and hours of practice were not associated with shoulder pain. In contrast, eight studies including two prospective (418,421), a retrospective (423), five cross sectionals (436–438,445,446) and two case control (450,451) found no association between training load and risk. In summary, the studies which investigated training load, volume, and intensity as a risk factor for shoulder pain are inconsistent. Therefore, the level of certainty that this risk factor associated with shoulder pain is low.

3.3.2.2 Stroke Distance and Stroke Specialty

Eight studies examined whether Stroke Distance and Stroke Specialty was associated with shoulder pain. One study found an association between Stroke Distance and Stroke Specialty and shoulder pain (53) and seven studies found no association (54,59,424,426,429,439,452). It has been reported that of all strokes, the butterfly stroke was the most likely to cause shoulder pain in males and females (53). However, in a prospective study by Walker et al (54) and a retrospective

study by Wolf et al (452), stroke specialty or stroke distance were not significantly associated with shoulder pain. Furthermore, a prospective study (426), five cross sectional studies (59,424,429,430,450) and a case control (431) failed to find a significant association between stroke specialty or stroke distance with shoulder pain in swimmers. In summary, based on the available evidence it is unlikely that stroke distance and stroke specialty is a risk factor for shoulder pain in swimmers. Therefore, the level of certainty that this factor is associated with shoulder pain is low.

3.3.2.3 Years swimming experience

Six studies examined whether years of swimming experience was associated with shoulder pain. Two studies found an association between Years swimming experience and shoulder pain (59,424) and four found no association (48,445,446). In a cross-sectional MRI study by Sein et al (59), increased supraspinatus tendon thickness was associated with years of swimming training experience demonstrated impingement pain and supraspinatus tendon thickening. Furthermore, a cross sectional study by Tate et al (424), found an association between swimmers with shoulder pain and years of swimming experience in ages 15–19-year-old swimmers, but not in the other age groups studied. However, two prospective studies (48,418) found that years of swimming experience was not significantly associated with shoulder injury. In addition, one retrospective (423), four case-control studies (444–446,450) and two cross sectional study (427,438) also failed to find an association between shoulder pain in swimmers and year of previous swimming experience. In summary, since the available evidence to assess years of swimming experience as a risk factor for shoulder pain in swimmer is insufficient, the level of certainty that this factor is associated with shoulder pain is low.

3.3.2.4 Breathing Side

Only three studies examined whether breathing side was associated with pain. One study found an association between breathing side and shoulder pain (424) and two found no association (424). In a cross-sectional study by Tate et al (424), bilateral breathing in the 8–11-year-old age group was found to be associated with shoulder pain, but no associations were found in breathing side preference in the other age groups (12-14, 15-18 and Masters). Further, a cross sectional study by Stocker et al. (428) failed to find a significant association between breathing side and shoulder pain in competitive collegiate and masters level swimmer. There is inconsistent and insufficient evidence that breathing side is a risk factor for shoulder pain in swimmers and therefore the level of certainty is low.

3.3.2.5 Swimming Training Equipment

Four studies examined whether Swimming Training Equipment was associated with shoulder pain. Only one study found an association between swimming training equipment and shoulder pain (53) and three studies found no association (424,428,432). In a cross-sectional study by McMaster et al (53), the use of hand paddles and kickboards were found to be significantly associated with shoulder pain in both male and female swimmers with shoulder pain. In three subsequent cross-sectional studies (424,428,432) the use of hand paddles was not significantly associated with shoulder pain or injury. In summary, there is no evidence that training equipment increases the risk of shoulder pain in swimmers. Therefore, the level of certainty that this risk factor is associated with shoulder pain is low.

3.3.2.6 Cross Training and Stretching

Four studies examined whether Cross Training and Stretching was associated with shoulder pain. Two studies found an association between Cross Training and Stretching and shoulder pain (53,426) and two studies found no association (54,429). In a prospective study (426), increased weight training was associated with increased shoulder pain in male swimmers. Furthermore, increased stretching in both males and females was associated with increased incidence of shoulder pain in a cross-sectional study (53). Additionally, shoulder pain was associated with strength training and weight training for both groups. However, in a prospective study (54), a retrospective (423) and two cross sectional studies (429,438) reported that the number of dry-land strength training sessions per week was not significantly associated with shoulder pain. In summary, there is insufficient evidence and inconsistent findings amongst the individual studies to suggest that cross training and stretching is a risk factor for shoulder pain in swimmers. Therefore, the level of certainty that this risk factor is associated with shoulder pain is low.

3.3.3 Demographics

3.3.3.1 Previous history of pain and injury

Four studies examined whether history of previous pain and injury was associated with shoulder pain. Three studies found an association between Previous history of pain and injury and shoulder pain (54,424,431) and one study found no association (429). In a prospective study by Walker et al (54), swimmers with a history of shoulder pain were 4.1 and 11.3 times more likely to sustain a subsequent injury for significant interfering pain in the shoulder and a significant shoulder injury,

respectively. A cross sectional study by Tate et al (424) found that previous history of pain or injury was significantly associated with shoulder pain in high school swimmers and masters' level swimmers. Moreover, Bansal et al (431) found a strong association between past history of shoulder pain and shoulder impingement syndrome. In contrast, four studies including three cross-sectionals (429,437,438) and a case control (450) found no association between injury history and risk. In summary, there is sufficient evidence from both Level I and two Level III studies to determine that previous history of pain and injury is a risk factor for shoulder pain in swimmers. However, due to the inconsistent findings, and lack of repeated level I studies, there is only a moderate certainty that this risk factor is associated with shoulder pain.

3.3.3.2 Sex

Five studies examined whether sex was associated with shoulder pain. Only two studies found an association between sex and shoulder pain (62,423) and four found no association (48,426,429,444). A retrospective study by Sallis et al (62) reported that female swimmers had significantly more shoulder injuries than the male swimmers. In a Level II retrospective study by Tessaro et al (423), female athletes reported significantly more incidents of shoulder pain compared to their male counterparts. Three prospective Level I studies (48,418,428) failed to find a significant difference in injury rates between male and female swimmers over the course of a competitive season. Furthermore, six cross sectional studies (51,435–439) and three case control (444,450,451) reported no significant differences between sexes and injury. In summary, there is insufficient evidence and inconsistencies amongst the individual studies, to suggest that sex is a risk factor for shoulder pain in swimmers. Therefore, the level of certainty that gender is a risk factor for shoulder pain is low.

3.3.3.3 Age

Sixteen studies examined whether age was associated with shoulder pain. Three cross sectional studies (433,436,437) that reported significant associations between age and injury. However, thirteen studies found no association between age and shoulder pain (48,54,423,424,426,429,439,444,450–452). In summary, there is currently no evidence that age is a risk factor for shoulder pain in swimmers and its certainty is therefore low.

3.3.3.4 Competitive Level

Four studies examined whether competitive level was associated with shoulder pain. Three studies found an association between shoulder pain and competitive level (429,441) and two studies found no association (418,444). In a description of elite and recreational swimmers (441), the elite group was found to have significantly higher incidence of shoulder overuse dysfunction and pain than the recreational group. In addition, the cross-sectional MRI study by Sein et al (59)found that competitive level was significantly associated with supraspinatus thickness and impingement pain and that athletes at higher levels of competition were more likely to have supraspinatus thickness and impingement pain than those at lower levels. Furthermore, the cross-sectional study by Kruger et al (429) reported that competitive swimmers were found to have significantly higher incidence of shoulder pain compared to recreational swimmers. In contrast, a large prospective study (418)and a case control (444) did not report a significant association. Therefore, there is insufficient and inconsistent evidence that competitive level is associated with shoulder pain and injury. As a result, our confidence that competitive level is associated with risk is therefore, low.

3.3.4 Musculoskeletal Factors

3.3.4.1 Scapular Kinematics, strength and dyskinesis

Three studies examined whether Scapular Kinematics, strength and dyskinesis was associated with shoulder pain. Only one study found an association between Scapular Kinematics, strength and dyskinesis and shoulder pain (444) and four studies found no association (424,434,436,444). In a study of 20 swimmers with shoulder impingement and 20 healthy swimmers without history of shoulder pain (444); shoulder impingement was significantly associated with reduced strength during the scapular strength test after swimming. In addition, swimmers with shoulder impingement also demonstrated less upward rotation at 45, 90 and 135 degrees of humeral elevation during arm abduction after swim training. Conversely, in the cross-sectional study by Harrington et al (51), scapular depression, abduction and adduction strength were not significantly associated with shoulder pain. Further, the frequency of obvious scapular dyskinesis was not found to be different between the swimmers who had shoulder pain and those that had no shoulder pain in any age group (424). In addition, two cross sectional studies (434,436) that found no significant difference in scapular kinematics between injured and uninjured swimmers. In summary, the evidence supporting scapular kinematics, strength and dyskinesis as a risk factor for shoulder pain in swimmers is insufficient and therefore the level of certainty that this risk factor is associated with shoulder pain is low.

3.3.4.2 Core Stability

Only two studies examined whether core stability was associated with shoulder pain. One study found an association between core stability and shoulder pain (424) and one study found no association (51). In a cross-sectional study by Tate et al (453) it was reported that youth swimmers (12 to 14 years old) with shoulder pain maintained a side bridge for significantly less time than uninjured swimmers. There were however no significant differences found in other age groups and between other tests of core stability (prone bridge and close kinetic chain bridge). A further study by Harrington et al (51) found no association between side bridge and prone bridge endurance and shoulder pain. In summary, the evidence available on trunk muscle endurance as a risk factor for shoulder pain in swimmers is insufficient. Therefore, the level of certainty that this risk factor is associated with shoulder pain is low.

3.3.4.3 Pectoral Length

Only two studies examined whether pectoral length was associated with shoulder pain. Both studies found an association between pectoral length and shoulder pain (51,424). It has been demonstrated by Borstad & Ludewig (454) that a short pectoralis minor demonstrated scapular kinematics like the kinematics exhibited in earlier studies by subjects with shoulder impingement. These results support the theory that an adaptively short pectoralis minor may influence scapular kinematics and is therefore a potential mechanism for subacromial impingement. In a cross-sectional study by Tate et al (424), a significant association was found in the resting normalized pectoralis minor length (normalized to clavicle length) at rest in the 15-year-old age group of girls with shoulder pain but not in any of the other age groups. Additionally, no associations were found

during pectoralis minor stretch and shoulder pain in any of the age categories. Furthermore, the cross-sectional study by Harrington et al (51) found that pectoralis length at rest and at stretch was significantly associated with shoulder pain in Division I female swimmers. In summary, there is currently inconsistent evidence that pectoralis length is a risk factor for shoulder pain in swimmers; therefore, the level of certainty that this risk factor is associated with shoulder pain is low.

3.3.4.4 Sub-acromial bursa thickness

Newly identified literature from 2015-2021, found two studies investigating sub-acromial bursa thickness and shoulder injury risk. In a prospective study by Couanis et al (421) investigating the relationship between bursa thickness and shoulder pain in open water swimmers found that sub-acromial bursa thickness increased in response to swimming training, but was not correlated to pain, suggestive of positive adaptive process. However, it was observed that sub-acromial bursa thickness increased significantly and was correlated with pain following an acute exacerbation event such as an open water race. In contrast, a case control study by Suzuki et al (451) found that while sub-acromial bursa thickness was significantly increased in swimmers compared to non-swimmers, there was no correlation between those swimmers with and without shoulder pain. In summary, due to insufficient evidence from the available literature, our certainty that sub-acromial bursa thickness is associated with shoulder pain is low.

3.4 CONCLUSION

The investigation of risk factors is an important initial step towards the understanding of the aetiology of shoulder injuries in swimmers. Following the inclusion of more recent studies (2015-2021), only a single new risk factor was identified, Sub-Acromial Bursa Thickness (Section

3.3.4.4). Secondly, all previously identified risk factors maintained their level of certainty apart from Competitive Level (Section 3.3.3.4) which decreased to a low level of certainty. Interestingly, despite the inclusion of 16 new articles in the last 6 years, none of the previously identified risk factors achieved a high level of certainty. Additionally, several risk factors were identified that comprised only of a single study *Swimming Scholarship Status* (452), *Inadequate Treatment* (431), *Triceps Length* (424), *Latissimus Length* (424) and *Internal/External Rotation Endurance* (427), *Serratus Anterior Thickness* (450), *Arm Span* (418), *Acute:Chronic Work Load Ratio* (418) and *Lower Trapezius Thickness* (450). Furthermore, some studies have suggested that changes in EMG activation of shoulder stability muscles are associated with shoulder pain during swimming (443,445,446,455–457). However, further investigation is required in order to determine whether the changes in EMG activation shoulder stability muscles are a possible cause of shoulder pain in swimmers or that shoulder pain modifies EMG activation. Finally, some studies have suggested that a genetic predisposition to shoulder injury exists (29) however; no studies have investigated this in swimmers.

Although several risk factors have been previously identified (21), it is important to note that shoulder injuries in swimmers are of complex aetiological origins. It has been well established that swimming training places excessive load on the shoulder (59,458), but in recent years, our understanding and management of shoulder injuries has evolved. It seems that greater measures are taken to reduce that burden including but not limited to changes in traditional paradigm of swimming training (459), prehabilitation (49,460), appropriate strength (461) and cross-training ((462), individual training load monitoring (461), and health surveillance programs (463,464). Therefore, it is not an absence of new studies but instead the publication of higher quality studies that has affected our level of certainty. As more literature becomes available, it is possible that our

confidence in previously identified risk factors will change, including the potential identification of new risk factors. This includes the potential role of genetic risk factors in the aetiology of shoulder injuries in swimmers.

There were several limitations to this systematic review. Although, a number of risk factors were identified in the included studies, poor study methodology or limited available data caused constrained estimation of injury risk to the shoulder for almost all of the investigated risk factors. None of the identified risk factors were found to have a high level of certainty. Prospective cohort design and consistency of measurements of risk should be employed in future studies. There was a clear lack of consistency of the definition of shoulder pain or injury. Several studies utilized self-reported pain without an objective diagnosis of pathology. Additionally, several aspects of one shoulder pathology were investigated which may weaken the significance of factors with only a single supporting study. Therefore, it is possible that a single risk factor for shoulder pain such as scapular rotation may possibly be nullified as a risk factor as it doesn't show a correlation with scapular dyskinesis. Secondly, some studies were excluded because they were unavailable or in a foreign language. Finally, the systematic review was limited by the quality of studies available, as the majority of studies were case control or cross sectional in study design.

This evidence-based systematic review provides descriptive analysis of several non-genetic risk factors for shoulder injury in swimmers. Although the identified risk factors did not achieve a high level of certainty, several of the risk factors were moderately associated with shoulder pain and injury in swimming (Table 3.6). In order to improve the definition and aetiology of shoulder injuries in swimmers, more high-quality studies are required to improve the inconsistency of results. Finally, a shoulder injury can potentially be a career ending happenstance, thus, an

understanding of the risk factors, including genetic risk factors, that predispose a swimmer to shoulder injury can improve the quality of training, reduce injury rates, and prolong a career. The genetic risk factors for shoulder injuries in swimmers will be investigated in the next chapter (Chapter 4).

CHAPTER 4: GENETIC RISK FACTORS FOR ROTATOR CUFF TENDINOPATHY IN SWIMMERS

4.1 INTRODUCTION

As previously outlined in Chapter 2, section 2.2.7, the shoulder is a common site of injury (56,59,442,465). This has been proposed to be due to the joint being optimised for mobility and range of motion, at the expense of mechanical stability (122,170,203,466). It has been reported that approximately 91% of swimmers have reported experiencing significant pain in their shoulder (59,467,468) with a reported incidence rate of 5.3 and 6.5 injuries per 1000 exposure hours in male and female swimmers, respectively (48).

As reviewed in the previous chapter, several non-genetic intrinsic and extrinsic factors have been reported to modulate the risk of shoulder injuries in swimmers. In addition to the non-genetic risk factors, recent literature has suggested that genetic factors contribute, at least in part, to shoulder injuries. However, studies investigating the genetic susceptibility to rotator cuff tendon injuries are limited (55) and were reviewed in Chapter 2, section 2.2.7.

Only a handful of studies have investigated the role of collagen gene variants in shoulder injuries. Khoschnau et al. (27), using a case-control candidate gene approach, published the first study which examined the association of collagen gene polymorphisms, specifically the *COL1A1* rs1800012 (G/T) Sp1 binding site polymorphism, and shoulder pathology. They reported that the GG and the heterozygous GT genotypes displayed a similar risk (OR, 1.06; 94% CI, 0.76-1.49),

whereas the TT genotype was under-represented in the injured population with shoulder dislocation compared with the controls (OR, 0.15; 94% CI, 0.03-0.68). However, a recent study by Vidal Rodríguez & del Castillo. (30) failed to find an association between the *COL1A1* rs1800012 polymorphism and risk of rotator cuff disease.

The *COL5A1* 3'-UTR rs12722 (T/C) has also been shown to modulate the risk of musculoskeletal injuries. Specifically, the CC genotype decreased the risk of ACL rupture in females (73), TEN (82) and CTS (94). The TT genotype was associated with increased risk of Tennis elbow (24). Although, two recent case control studies were unable to find an association between *COL5A1* rs12722 polymorphism and RCT (187) and RCD (30). More recently, Alakhdar et al (469) investigated RCT in 137 youth athletes and found the CC genotype of the *COL5A1* rs12722 polymorphism was significantly over-represented in the tendinopathy group. The authors also investigated the role of *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (A/T), however no significant difference was found between cases and controls.

As outlined in Chapter 2, section 2.7., *COL1A1* rs1800012 (G/T), *COL5A1* rs10628678 (-/AGGG), *COL6A1* rs35796750 (C/T), *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) and *COL12A1* rs970547 (G/A) haven been previously investigated in several injury phenotypes. However, as summarised above, few studies have investigated the possible association of a subset of these polymorphisms with rotator cuff pathology. These polymorphisms were selected for investigation because of their previous association with several muscular skeletal soft tissue injuries (Chapter 2, section 2.7). We hypothesized that, at least a subset of these collagen gene polymorphisms independently or interact to modulator the risk of RCT.

The objective the study in this chapter of the thesis was to investigate the association of collagen gene polymorphisms with risk of RCT, specifically supraspinatus tendinopathy, in a South African swimming cohort of self-reported European ancestry. The aims of this study were to determine whether the *COL1A1* rs1800012 (G/T), *COL5A1* rs12722 (T/C), *COL5A1* rs10628678 (AGGG/-), *COL6A1* rs35796750 (T/C), *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T), *COL11A2* rs1799907 (A/T) and *COL12A1* rs970547 (G/A) polymorphisms either independently or via gene-gene interactions modulated risk. Specifically, based on previous studies, we hypothesised that the 1) TT and CC genotypes of *COL1A1* rs1800012 (G/T) and *COL5A1* rs12722 (T/C), respectively, were associated with reduced risk; 2) the AGGG/AGGG, TT and AA genotypes of *COL5A1* rs10628678 (AGGG/-), *COL6A1* rs35796750 (T/C) and *COL12A1* rs970547 (G/A), respectively, were associated with increased risk; 3) the T allele of the *COL6A1* rs35796750 (T/C) was associated with decreased risk and 4) the inferred T-C-T pseudo-haplotype constructed from *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) was associated with increased injury risk.

4.2 METHODS

4.2.1 Study Design

This case-control genetic association study followed a candidate gene approach using the recommendations made by the Strengthening the Reporting of Genetic Association Studies (STREGA) initiative (470), an extension of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement checklist (471). The recommendations were applied to all subsequent chapters (Chapters 4 to 6) in this thesis. The study in this chapter was

approved by the Faculty of Health Sciences Human Research Ethics Committee within the University of Cape Town, South Africa (HREC 421/2013) (Appendix A1.1).

4.2.2 Participants

A total of 103 (49 females, 54 males) swimmers (RCT) with clinically diagnosed RCT between 2009-2016 were recruited. The recruitment was conducted between 2013-2016. Of the 103 participants in the RCT group, 84.5% (n=87) were diagnosed with a Supraspinatus Tendinopathy (SST) and were also analysed separately as a sub-group. Further, 101 (55 females, 46 males) apparently healthy swimmers with no previous history of shoulder pathology (including RCT, trauma, bursitis, or adhesive capsulitis) (CON) were recruited.

All participants in the RCT and CON groups were current swimmers from the Western Cape Region of South Africa at the time of recruitment and were of self-reported European descent. All participants were required to complete a written informed consent form (Appendix A1.3), in accordance with the Declaration of Helsinki (64th WMA General Assembly, Fortaleza, Brazil, October 2013) (472) and a questionnaire specifying personal details, injury details and medical and sporting history (Appendix A1.4). Participants were excluded from the study if they had been previously diagnosed with any other connective tissue disorders or other systemic diseases believed to be associated with skeletal muscle pathology as previously outlined in (82) e.g. Ehlers-Danlos syndrome, benign hypermobility joint syndrome, rheumatoid arthritis, rhabdomyolysis, myopathy and muscular dystrophy.

4.2.3 Diagnosis of Rotator Cuff Tendinopathy

Participants in the RCT group were identified through patient databases at two Orthopaedic Surgery Clinics in Cape Town, South Africa. All RCT participants were clinically diagnosed by at least one experienced orthopaedic surgeon using standard clinical tests for diagnosis of rotator cuff pathology including Jobe's test, Patte's test, Neer's sign and Hawkins' test. For this study, all rotator cuff tendinopathies were included (including supraspinatus, infraspinatus, subscapularis, teres minor and the long head of the biceps tendon) in order to maintain sample size. Although the biceps tendon is not typically characterised as a rotator cuff tendon, it functions as a secondary shoulder stabiliser (473,474) and is often affected in other rotator cuff disorders (167,475) such as such as supraspinatus tendinopathy (476). Furthermore, the long head of the biceps is closely linked to the glenohumeral complex and is a possible contributor to the onset of anterior shoulder pain (18,107,477).

4.2.4 Blood Collection and Extraction

Approximately 4.5ml of venous blood was collected from each participant by venepuncture of the antecubital vein into an EDTA vacutainer tube by a qualified phlebotomist. The samples were stored at 4˚C until DNA extraction as previously described Lahiri & Numberger (478) with minor modifications (385). The reactions were performed on an Applied Biosystems StepOnePlusTM Real-Time PCR system (Applied Biosystems, Foster City, CA, USA), using the Applied Biosystems StepOnePlusTM Real-Time PCR software Version 2.1 (Applied Biosystems, Foster City, CA, USA. Genotypes were determined by endpoint fluorescence. For PCR and genotype quality control purposes, a number of positive (known genotypes) and DNA-free controls were

randomly included on every 96-welled PCR plate. All control samples were successfully repeated on every plate.

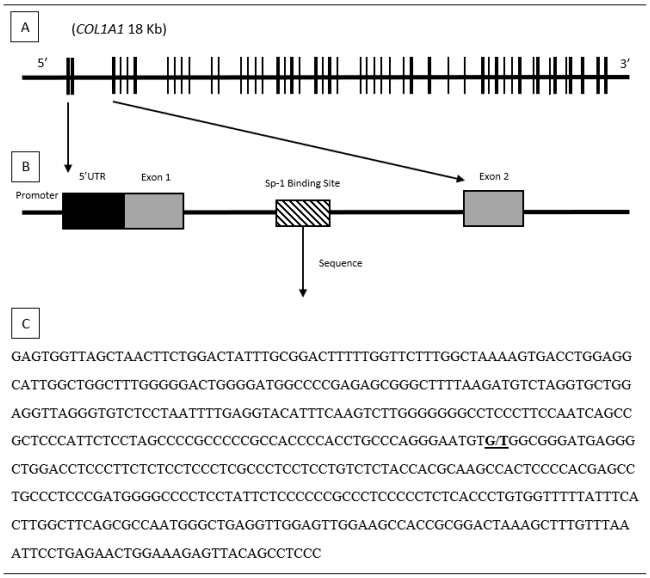
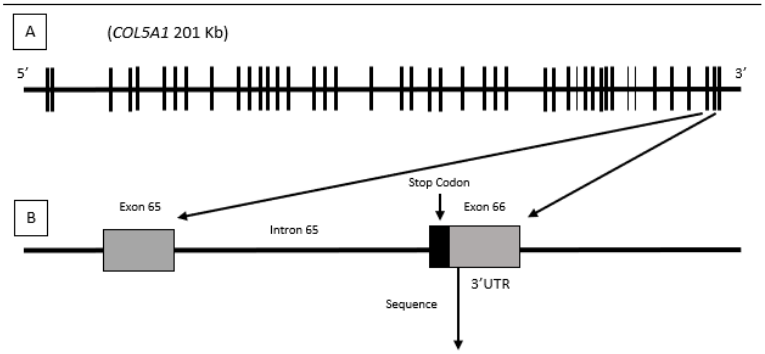


Figure 4.1. A) Schematic representation of the 18 Kb *COL1A1* gene on chromosome 17q21.33. The exon (vertical lines) and intron (horizontal spaces) boundaries are shown. **B)** The 5'-end of the gene spanning exons 1 and 2 has been enlarged with the 5'-UTR of exon 1 represented as a black box and the translated regions of exons 1 and 2 as grey boxes. The Sp1-binding site polymorphism (rs1800012) is represented as a hatched box. **C)** The rs1800012 single nucleotide polymorphism, which is boxed, and flanking sequences. The Sp1-binding site is bold and underlined. The schematic representation is not necessarily drawn to scale.



C

ACCAAGAAAGGCTACCAGAAGACGGTTCTGGAGATCGACACCCCCAAAGTGGAGCAGGTGCCCATCGTGGACATCATGTTCA
ATGACTTCGGTGAAGCGTCACAGAAATTTGGATTTGAAGTGGGGCCGGCTTGCTTCATGGGCTAGGAGCCGCCGAGCCCGGGC
TCCCGAGAGCAACCTCGTGACCTCAGCATGCCATTCGTTCGTGAGTGTCCCGTGCACGTCCTGACCCTGGACAGTGAAGGCTTC
TCCCTCCCCTCCCACCTGACTTCATCTACGCCTCGGCACCACGGGGTGTGGGACCCCAGCCCGGAGAGAACAGAGGGAAGGAG
CCGCGCCCCCACCTGGAGCTGAATCACATGACCTAGCTGCACCCCAGCGCCTGGGCCCGCCCCACGCTCTGTCCACACCCAC
T/CGCGCCCCGGGAGCGGGGCCATGCCTCCAGCCCCCCAGCTCGCCCGACCCATCCTGTTCGTGAATAGGTCTCAGGGGTTGGG
GGAGGGACTGCCAGATTTGGACACTATATTTTTTTCTAAATTCAACTTGAAGATGTGTATTTCCCCTGACCTTCAAAAAATGTT
CCAAGGTAAGCCTCGTAAAGGTCATCCCACCATCACCAAAGCCTCCGTTTTTAACAACCTCCAACACGATCCATTTAGAGGCC
AAATGTCATTCTGCAGGTGCCTTCCCGATGGATTAAAGGTGCTTATGTTTTTGTGAGTTTTAAGTAAATATTTGTATTGTATTGT
TATAAATGTTAAGTGTGCCTGGCTTTCAATCATGCACGGAAACCCAGTCTCAGTCCCACGGACAGAATGGGCGAGGCATGGAT
TCTGGGTTGCAGTACCGTTCTGATTAGAAATAGGAAGTCTCCCCACCCCCGCCCTGGCCAAGAACGTGCAATAAATTGGAAGT
TTGCCCCGGGGCAGCAAGAATTTATGCTGCCATTGAAAAGCAGGTACCAGTGCCCCTTTTCAGACAGTTTTTGATTCGCTCTAG
ACTTTTTTTTTTTTTAATAGGGAD/ACGGAAAAATTTGATAATTTTCTTTTTTCTACATGCACTTAAGACTAAAACACAGGTTTG
GATTAATTTTATTTGCTTCCTTTTTCCGCTTTTCTTCCCGCAGAGCCTGATGGGAGAATGTCCAGGGCAGGGAAACCACATTTTT
TGTAGGTGATAACTCAATGAAAATTGGTGCTTATTTTTTACACTTCTCTCTTGTGGCTCTCTTGTGGTGCTATCTATCTGTTTTAA
GGTCTCCTTGAAGGCGCACTGGGGACCCTGGCCATGCCTCGTTCTCCCTGCTTTCTTTATCCTGTTATTGCCTCCACAGTCTGTT
GCCAAGGACTCTAAGATCAATGCACGTCACTTTCCTTTCCACTGGGCAGGATAGCCAAGCACACTCCCTCCTGCGCTCTCCCGC
CCCGGTGCGTCCACTCCCGAGGGCTGTTATGAGGACTGGGTTGTGCCTACTTGATTTGAAAACACACACAAGCAATAAAAAGC
CTCTTCCTGCATTGTCTGTGG

Figure 4.2. A) Schematic representation of the 201 Kb *COL5A1* gene *1* gene on chromosome 17q21.33. The exon (vertical lines) and intron (horizontal spaces) boundaries are shown. **B)** The 3'-end of the gene spanning exons 65 and 66 has been enlarged with the 3'-UTR of exon 66, which contain the rs12722 (T/C) and rs10628678 (-/AGGG) polymorphisms, is represented as a hatched box and the translated regions of exons 65 and 66 as grey boxes. **C)** The rs12722 (T/C) and rs10628678 (-/AGGG) polymorphisms, which are bold and underlined, and flanking sequences. The schematic representation is not necessarily drawn to scale.

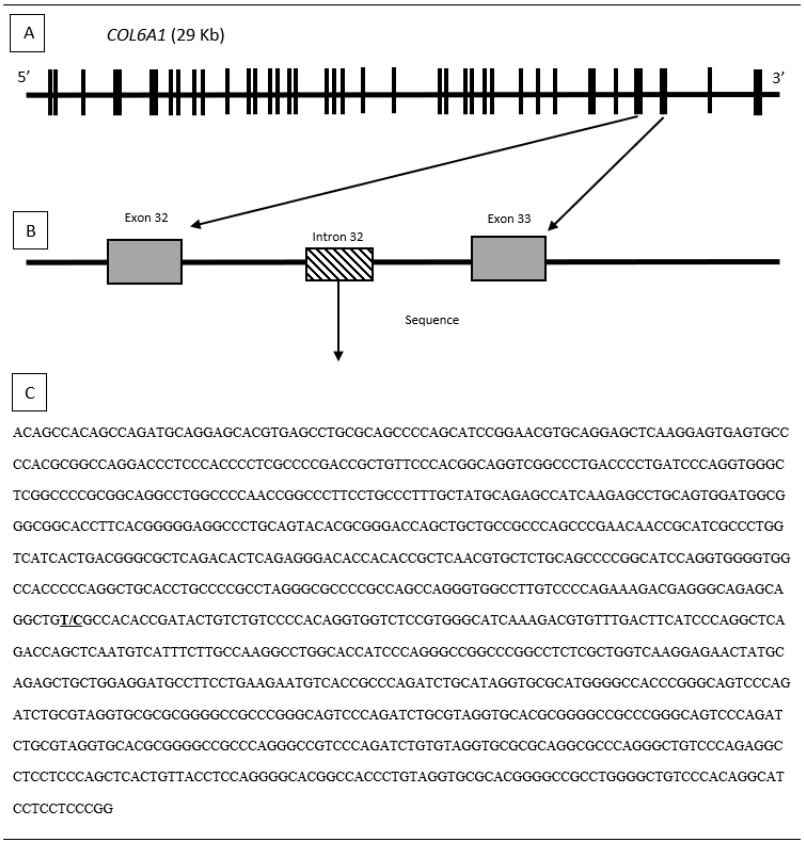


Figure 4.3. A) Schematic representation of the 29 Kb *COL6A1* gene on chromosome 21q22.3. The exon (vertical lines) and intron (horizontal spaces) boundaries are shown. The 3' – end of the gene spanning exons 32 and 33 has been enlarged **B)**. The specific polymorphism location is represented by a hatched box (Intron 32). The flanking sequence for polymorphism rs35796750 is shown **C)**. The rs35796750 polymorphism is in bold and underlined. The schematic representation is not necessarily drawn to scale.

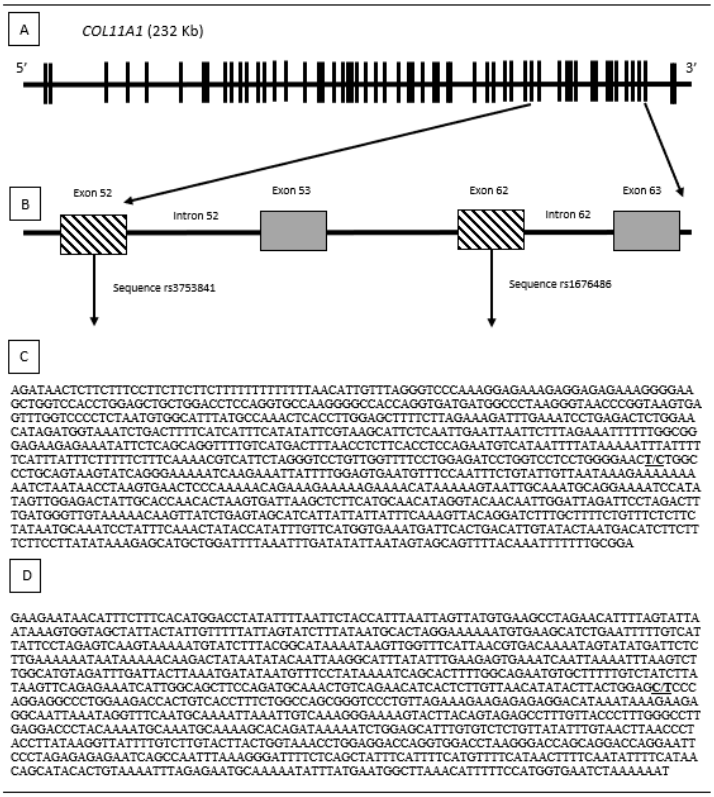


Figure 4.4. A) Schematic representation of the 232 Kb *COL11A1* gene on chromosome 1p21.1. The exon (vertical lines) and intron (horizontal spaces) boundaries are shown. The 3' – end of the gene spanning exons 52 and 53 has been enlarged **B)**. The specific polymorphism location is represented by a hatched box (Exon 52). The flanking sequence for polymorphism rs3753841 is shown **C)**. The rs3753841 polymorphism is in bold and underlined. The 3' – end of the gene spanning exons 62 and 63 has been enlarged **B)**. The specific polymorphism location is represented by a hatched box (Exon 62). The flanking sequence for polymorphism rs1676486 is shown **D)**. The rs1676486 polymorphism is in bold and underlined. The schematic representation is not necessarily drawn to scale.

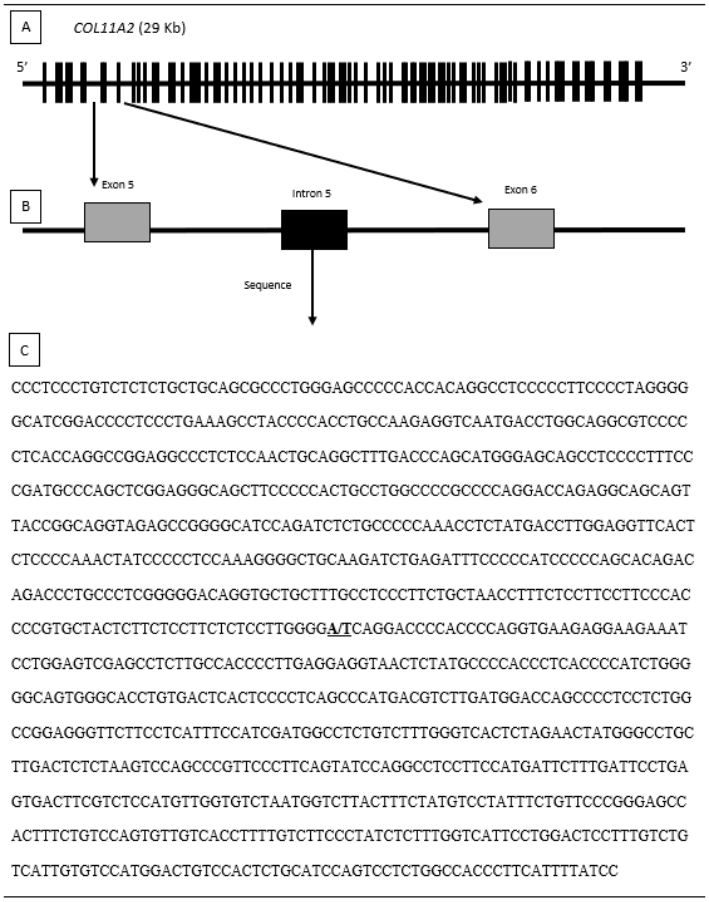


Figure 4.5. Schematic representation of the 29 Kb *COL11A2* gene **A)** on chromosome 6p21.3. The exon (vertical lines) and intron (horizontal spaces) boundaries are shown. The 3' – end of the gene spanning exons 5 and 6 has been enlarged **B)**. The specific polymorphism location is represented by a red box. The flanking sequence for polymorphism rs1799907is shown **C)**. The rs1799907 polymorphism is highlighted in red and is located within a box. The schematic representation is not necessarily drawn to scale.

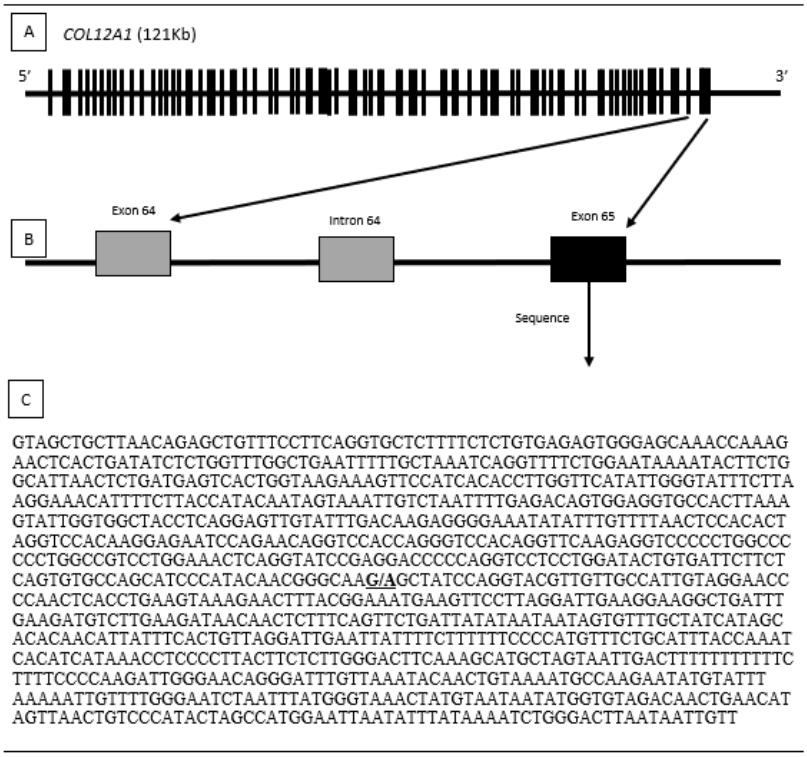


Figure 4.6. Schematic representation of the 121 Kb *COL12A1* gene **A)** on chromosome 6p12. The exon (vertical lines) and intron (horizontal spaces) boundaries are shown. The 3' – end of the gene spanning exons 5 and 6 has been enlarged **B)**. The specific polymorphism location is represented by a red box. The flanking sequence for polymorphism rs970547 is shown **C)**. The rs970547 polymorphism is highlighted in red and is located within a box. The schematic representation is not necessarily drawn to scale.

Genotyping of the *COL5A1* rs12722 (BstUI RFLP) polymorphism (Table 4.1) was performed using PCR and restriction fragment length polymorphism analysis as previously described (73,479)The 667 bp PCR products of the *COL5A1* gene were digested with the restriction endonuclease BstUI (5'- CG'CG -3') to produce 351 and 316 bp for the T allele and 316, 271 and 80 bp for the C allele and the resultant amplicons were resolved together with a 100bp molecular

weight marker and SYBER® Gold nucleic acid gel stain (Invitrogen Molecular Probes^{TM}, Oregon, USA) on 6% non-denaturing polyacrylamide gels (Figure 4.7) . The gels were photographed under UV light using an Uvitec photo-documentation system (Uvitec Limited, Cambridge, UK) and genotypes were determined based on the resultant DNA fragment sizes.

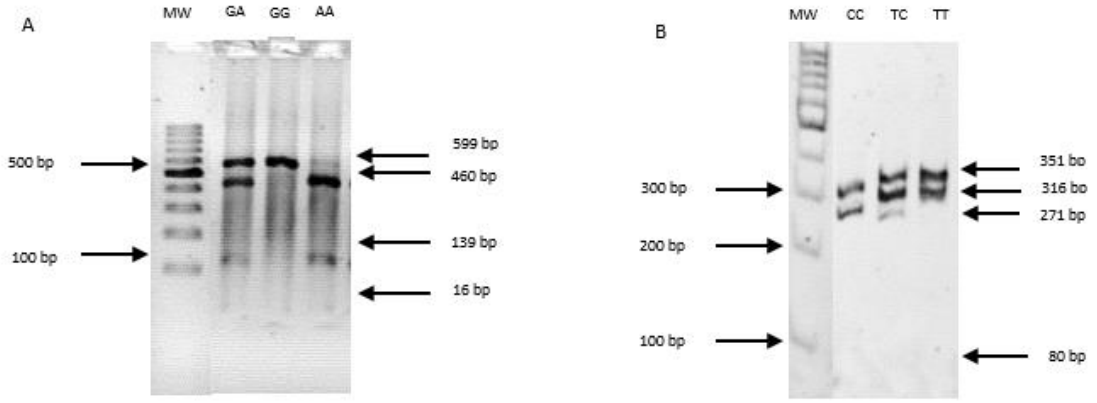


Figure 4.7. Typical non-denaturing visualising gels showing the genotypes of **A)** *COL12A1* rs970547 and **B)** *COL5A1* rs12722 restriction fragment length polymorphisms (RFLPs). In **A)**, digestion of the 615 bp PCR product with *AluI* enzyme produces 599 bp and bp fragments for the G allele and 460 bp, 139 bp and 16 bp fragments for the A allele. The 16 bp fragement is not fully visable on the figure. In **B)**, digestion of the 667 bp PCR product with the *BstUI* enzyme produces 316 bp, 271 bp and 80 bp in the C allele and 351 bp, 316 bp and 271 bp in the T allele. The 80 bp allele in fully visable on the figure. The left lane on each figure shows the 100 bp molecular weight (MW) ladder marker with the appropriate fragment sizes denoted in base pairs (bp).

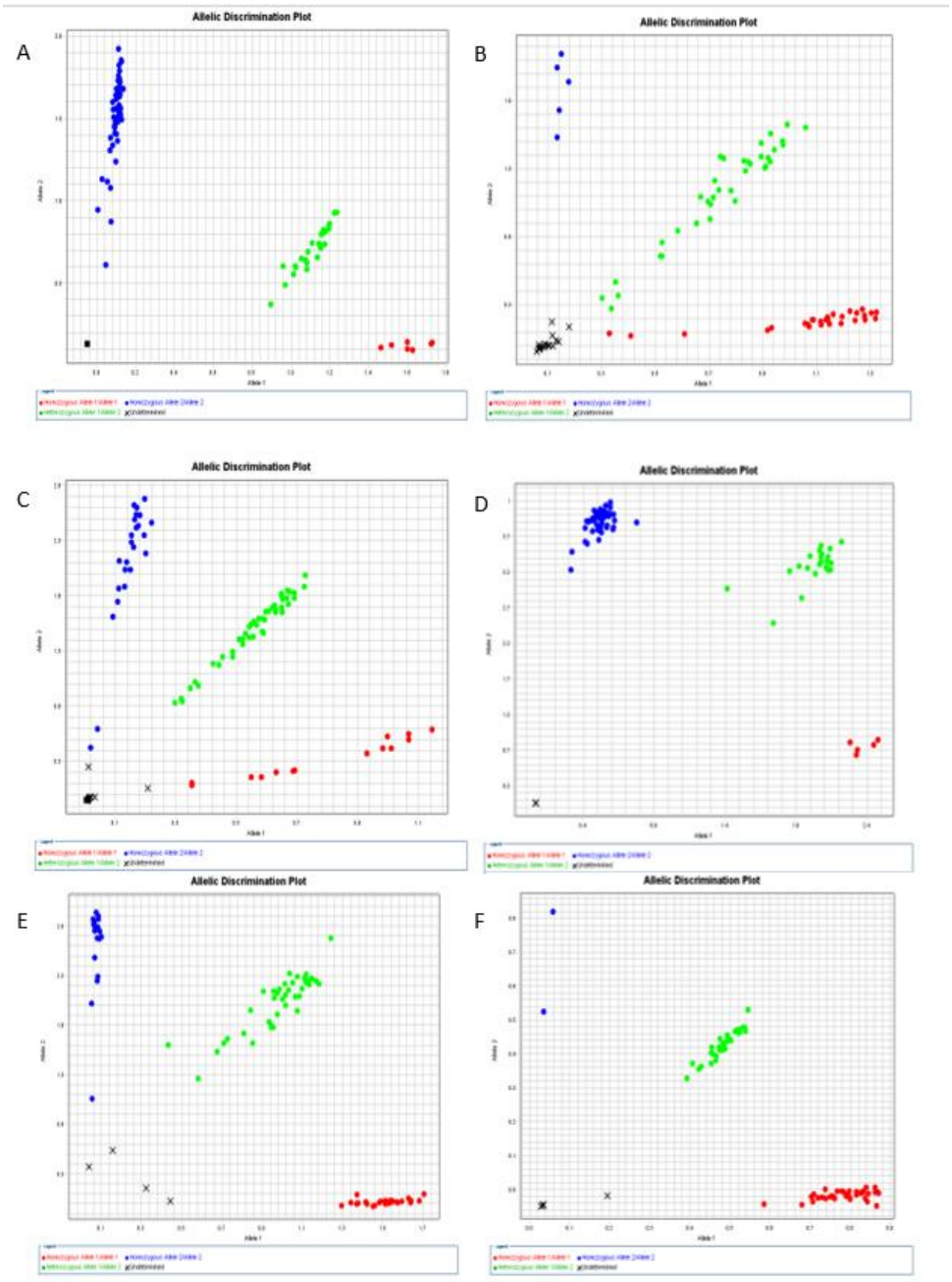


Figure 4.8. Typical allelic discrimination plots using the Taqman® Genotyping assay for **A)** *COL1A1* rs1800012 (G/T), **B)** *COL5A1* rs10628678 (AGGG/-), **C)** *COL6A1* rs3579750 (T/C), **D)** *COL11A1* rs3753841 (C/T), **E)** *COL11A1* rs1676486 (C/T), and **F)** *COL11A2* rs1799907 (T/A) on the StepOnePlus™ Real-time PCR System.

Table 4.1. Details of custom designed Fluorescence-based Taqman® polymerase chain reaction (PCR) assays (Applied Biosystems, Foster City, CA, USA) for genotyping including, Assay ID, polymorphism location and context sequence for *COL1A1* rs1800012, *COL5A1* rs10628678, *COL6A1* rs35796750, *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL11A2* rs1799907.

COL1A1	
rs1800012	
Assay ID	C___7477170_30
Location	Chromosome 17: 50200388
Context Sequence	GGGAGGTCCAGCCCTCATCCCGCCC **[A/C]**
[VIC/FAM]	CATTCCCTGGGCAGGTGGGGTGGCG
COL5A1	
rs10628678	
Assay ID	AH20W1K
Location	Chromosome 9: 134843172
Context Sequence	AGACTTTTTTTTTTTTTAATAGGGA **[-/AGGG]**
[VIC/FAM]	AAAAATTTGATAATTTTCTTTTTTC
COL6A1	
rs35796750	
Assay ID	C__25761725_10
Location	Chromosome 21: 46002498
Context Sequence	AGAAAGACGAGGGCAGAGCAGGCTG **[C/T]**
[VIC/FAM]	GCCACACCGATACTGTCTGTCCCCA
COL11A1	
rs3753841	
Assay ID	C___2947954_10
Location	Chromosome 1: 102914362
Context Sequence	TTCCCTGATACTTACTGCAGGGCCA **[G/A]**
[VIC/FAM]	GTTCCCCAGGAGGACCAGGATCTCC
COL11A1	
rs1676486	
Assay ID	C___8400671_10
Location	Chromosome 1: 102888582
Context Sequence	TCTTGTTAACATATACTTACTGGAG **[A/G]**
[VIC/FAM]	CCCAGGAGGCCCTGGAAGACCACTG
COL11A2	
rs1799907	
Assay ID	C__25474257_10
Location	Chromosome 6: 33185058
Context Sequence	CCTCTTCACCTGGGGTGGGGTCCTG **[A/T]**
[VIC/FAM]	CCCCAAGGAGAGAAGGAGAAGAGTA

The *COL12A1* rs970547 polymorphism was also genotyped using RFLP as previous described (87). PCR reaction was performed in a final volume of 60µl containing at least 100ng DNA, 20pmol of the forward (5'-GAG AAT CCA GAA CAG CTC CAC CAG-3') and reverse (5'-CAT GGC TAG TAT GGG ACA G-3') primers, 2.0mM $MgCl_2$, 50mM KCl, 10mM Tris-HCl (pH 8.3), 200µmol of dNTPs (dATP, dTTP, dCTP and dGTP) and 1 unit of DNA Taq polymerase (New England Biolabs, Ipswich, Massachusetts, USA). Amplifications were conducted by denaturing for 1 cycle at 94ºC for 3 minutes, followed by 30 cycles at 94ºC for 30 seconds, 55ºC for 30 seconds, and 72ºC for 1 minute, followed by a final extension of 1 cycle at 72ºC for 5 minutes on a thermal cycler (Hybaid; PCR Express, Middlesex, UK). The 615 bp PCR products of the *COL12A1* gene was digested with restriction endonuclease AluI (5'-AG'CT-3') to produce 16 bp, 599 bp sized fragments for the G allele and 16 bp, 139 bp, 460 bp sized fragments for the A allele (Figure 4.7 A). The resultant fragments were separated, together with a 100 bp DNA ladder of known size markers (Promega Corporation, Madison, Wisconsin, USA) and SYBER® Gold nucleic acid gel stain (Invitrogen Molecular Probes™, Oregon, USA), on 2% agarose gel. The gels were photographed under UV light using a Uvitec photo-documentation system (Uvitec Limited, Cambridge, UK) and genotypes were determined based on the sizes of the DNA fragments. All sample extraction and genotyping took place at the Division of Exercise Science and Sports Medicine Laboratory within the Faculty of Health Sciences at the University of Cape Town. Investigators were blinded to the phenotypes of the participant samples while genotyping.

Furthermore, positive and negative controls were used to ensure genotyping accuracy and as a preventative measure to detect contamination. A total of 198/204 (97.1%), 201/204 (99.4%), 200/204 (99.4), 187/204 (91.7%), 197/204 (96.6%), 199/204 (97.4%), 197/204 (96.6%) and 200/204 (98.0%) genotypes were identified for *COL1A1* rs1800012, *COL5A1* rs12722 and

rs10628678, *COL6A1* rs35796750, *COL11A1* rs3753841 and rs1676486, *COL11A2* rs1799907 and *COL12A1* rs970547 polymorphisms respectively.

4.2.5 Statistics

QUANTO v1.2.4 (http://hydra.usc.edu/gxe) was used to calculate the sample size required for the study to be statistically powered. Assuming allele frequencies between 0.1 and 0.9 for the 'risk' allele of each polymorphism investigated, our sample size of 204 participants would be adequate to detect an allelic OR of 1.8 and greater at a power of 80% and a significance level of 4%. Genotype and allele frequencies and Hardy-Weinberg equilibrium (HWE) were analysed using the genetics package of the programming language and environment R (http://www.r-project.org), and GraphPad V6 (GraphPad software, San Diego, California, USA). Descriptive statistics were compared using one-way analysis of variance to determine any significant differences between the characteristics of the CON versus RCT groups. The R packages "genetics" (480) and "SNPassoc" (481) were used to analyse any differences in genotype and allele frequencies and to calculate HWE probabilities and linkage disequilibrium (LD).

Inferred pseudo-haplotypes were constructed as per previously literature using specific Polymorphisms within each gene (76,89,94,98). Analysis of the inferred haplotypes was done with the "haplo.stat" package (482,483) of the programming language and environment R (http://www.r-project.org) in order to compare allele frequencies of the polymorphisms within each haplotype between cases and controls. Significance was accepted at $p<0.05$. There were no adjustments performed for multiple testing, as there is currently no obvious appropriate method (484,485). For this study, Bonferroni correction was considered too conservative, as the tests were

conducted on the same group of participants (484). Furthermore, correction for multiple testing was considered inappropriate, as there was an a priori hypothesis that the gene polymorphisms investigated are associated with the RCT phenotype (86,486).

4.3 RESULTS

4.3.1 General Characteristics

The general characteristics of the participants are summarised in Table 4.2. There were no significant differences found between the CON and RCT groups, as well as between the CON group and SST sub-group, with regards to general characteristics (age, sex, height, weight and body mass index), current training loads and years of swimming experience at the time of recruitment.

Table 4.2: General participant characteristics for the control (CON), Rotator Cuff Tendinopathy (RCT) groups and the Supraspinatus Tendinopathy (SST) sub-group.

	CON (n=101)	**RCT** (n=103)	p-value$^{\alpha}$	**SST** (n=87)	p-value$^{\beta}$
Age (years) *	35.1 ± 14.5 (97)	37.7 ± 14.1 (102)	0.256	37.3 ± 14.2 (86)	0.241
Female Sex (%) (n)	55.4 (55)	54.9 (49)	0.291	52.9 (43)	0.187
Height (cm)	173.9 ± 9.3 (99)	175.9 ± 8.6 (101)	0.259	175.9 ± 15.3 (87)	0.257
Weight (kg) *	74.0 ± 17.0 (99)	77.6 ±13.9 (99)	0.454	78.3 ± 15.3 (87)	0.314
BMI (kg.m^{-2})	24.6 ± 6.2 (96)	25.6 ± 5.7 (98)	0.781	24.9 ± 5.9 (87)	0.811
Training load (hrs. wk^{-1})	4.7 ± 3.0 (98)	5.2 ± 4.9 (99)	0.463	5.2 ± 5.5 (87)	0.693
Swim Experience (yrs.)	19.3 ± 13.6 (98)	21.5 ± 13.7 (99)	0.275	21.8 ± 13.9 (87)	0.237

Values are expressed as mean ± standard deviation or as a frequency (%).
The number of participants (n) is indicated in parentheses.
BMI: Body Mass Index, cm: centimeters, kg: kilograms, m: meters, hrs. wk^{-1}: hours per week
Current Training load: Current hours per week of swimming training at time of recruitment.
$^{\alpha}$ = CON vs. RCT.
$^{\beta}$ = CON vs. SST.
* = age and weight at injury for the RCT group and SST sub-group; and at recruitment for CON group

4.3.2 Characteristics of Clinical Examination

Supraspinatus tendinopathy was the most diagnosed (84.2%, n=85) cause of pathology followed by tendinopathy of the long head of the bicep tendon (9.9%, n=10), then the infraspinatus (6.9%, n=7), subscapularis (6.9%, n=7) and teres minor being the least common diagnosis (3.0%, n=3). Some participants underwent additional examination which included ultrasound (27.8%, n=28), X-ray (10.9%, n=11), Magnetic Resonance Imagining (MRI) (8.9%, n=9), and surgery (7.9%, n=8).

4.3.3 Participant and Familial History of Soft Tissue Injuries

Participants in the RCT group and SST sub-group (Table 4.3) reported significantly more multiple joint injuries (>2 different tendon and ligament injuries) than those in the CON group. Similarly, participants in the RCT group, but not the SST sub-group, also reported significantly more tendon injuries (>2 different tendons injured) than those in the CON group. However, there were no significant differences in participant ligament injury, participant injuries to multiple ligaments. Similarly, there was no other significant difference in history of tendon and ligament injuries between the groups. Furthermore, the groups were similarly matched for family history of tendon, ligament, and joint injuries.

Table 4.3: Participant and family history of soft tissue injuries for the participants in the control group (CON), Rotator Cuff Tendinopathy (RCT) group as well as the Supraspinatus Tendinopathy (SST) sub-group. Sex specific comparisons between the two groups are not reported.

	CON (n=100)	**RCT (n=101)**	**p-value$^{\alpha}$**	**SST (n=87)**	**p-value$^{\beta}$**
Participant multiple joint injury	21.0 (21)	37.6 (38)	**0.009***	35.6 (31)	**0.026***
Participant tendon injury	38.0 (38)	45.4 (46)	0.296	46.9 (38)	0.458
Participant injuries to multiple tendons	7.0 (7)	17.8 (18)	**0.020***	13.8 (12)	0.125
Participant ligament injury	45.2 (42)	57.7 (56)	0.083	58.3 (49)	0.055
Participant injuries to multiple ligaments	19.0 (19)	27.7 (28)	0.144	26.4 (23)	0.224
Family history of joint injury	48.0 (48)	50.5 (51)	0.724	51.7 (45)	0.611
Parent joint injury	30.0 (30)	30.7 (31)	0.915	32.2 (28)	0.747
Sibling joint injury	17.0 (17)	22.8 (23)	0.305	21.8 (19)	0.403
Child joint injury	7.0 (7)	5.0 (5)	0.540	5.8 (5)	0.727
Family history of tendon injury	19.0 (19)	27.7 (28)	0.098	27.6 (24)	0.164
Parent tendon injury	11.0 (11)	15.8 (16)	0.314	16.1 (14)	0.307
Sibling tendon injury	8.0 (8)	10.9 (11)	0.484	10.3 (9)	0.578
Child tendon injury	4.0 (4)	2.0 (2)	0.400	2.3 (2)	0.510
Family history of ligament injury	41.0 (41)	44.8 (43)	0.920	43.5 (37)	0.960
Parent ligament injury	25 (25)	24.8 (25)	0.967	25.3 (22)	0.963
Sibling ligament injury	13.0 (13)	19.8 (20)	0.193	19.5 (17)	0.154
Child ligament injury	7.0 (7)	4.0 (4)	0.343	4.6 (4)	0.486

Values are expressed as percentages with the number of participants (n) indicated in parentheses. Values in bold typeset are significant ($p<0.05$).
CON, apparently healthy controls.
RCT, rotator cuff tendinopathy.
SST, supraspinatus tendinopathy.
Joint injuries include both ligament and tendon injuries (fingers, wrist, elbow, shoulder, pelvis, knee and ankle).
$^{\alpha}$ = CON vs. RCT.
$^{\beta}$ = CON vs. SST.

4.3.4 Swimming Training Variables

The results of swimming training variables are summarised in Table 4.4. The RCT group and SST sub-group reported significantly more summer and winter training hours per week than the CON

group. Participants were found to use swimming training equipment equally across the groups with no significant differences in usage detected between the groups.

Table 4.4: Summary of Swimming Training Variables control group (CON), Rotator Cuff Tendinopathy (RCT) group Supraspinatus Tendinopathy (SST) sub-group. Sex specific comparisons between the two groups are not reported.

	CON (100)	**RCT (101)**	**p-value$^{\alpha}$**	**SST (n=87)**	**p-value$^{\beta}$**
Summer Hrs/wk	5.4 ± 3.4 (97)	7.1 ± 5.3 (94)	**0.006***	7.3 ± 5.5 (86)	**0.005***
Winter Hrs/wk	4.2 ± 3.1 (97)	5.7 ± 5.3 (94)	**0.011***	5.9 ± 5.6 (86)	**0.009***
Paddles Usage (Use/wk)	1.3 ± 1.7 (95)	1.7 ± 2.3 (97)	0.126	1.5 ± 2.0 (87)	0.699
Kickboard Usage (Use/wk)	2.1 ± 2.1 (95)	2.1 ± 2.2 (97)	0.977	2.4 ± 2.0 (87)	0.371
Pool buoy (Use/wk)	1.5 ± 1.8 (96)	2.0 ± 2.2 (97)	0.096	1.9 ± 2.1 (87)	0.122

Swimming training variables are expressed as mean ± standard deviation and the number of participants shown in parenthesis (n).
(Hrs/wk) = Hours of training per week.
(Use/wk) = Usage of item per week.
$^{\alpha}$ = CON vs. RCT.
$^{\beta}$ = CON vs. SST.
Global P-values are given with bold typeset indicating significant differences (P<0.05).

4.3.5 Independent Genotype Analysis

There were no significant differences in the genotype or allele frequency distributions of any of the polymorphisms investigated (Table 4.5) in this chapter between the CON and RCT groups (*COL1A1* rs1800012, p=0.567; *COL5A1* rs12722, p=0.531; COL5A1 rs10628678, p=0.519; *COL6A1* rs35796750 p=0.480; *COL11A1* rs3753841, p=0.741; *COL11A1* rs1676486, p=0.902; *COL11A2* rs1799909, p=0.241, and *COL12A1* rs970547, p=0.967). Similarly, there were no significant differences found between the CON group and SST sub-group (*COL1A1* rs1800012, p=0.852; *COL5A1* rs12722, p=0.596; *COL5A1* rs10628678, p=0.236; *COL6A1* rs35796750

p=0.592; *COL11A1* rs3753841, p=0.858; *COL11A1* rs1676486, p=0.858; *COL11A2* rs1799909, p=0.208, and *COL12A1* rs970547, p=0.954) (Table 4.5). The *COL1A1*, *COL5A1*, *COL6A1*, *COL11A1*, *COL11A2*, *COL11A2* and *COL12A1* polymorphisms were in Hardy-Weinberg Equilibrium (HWE).

Table 4.5: Genotype frequency distributions for *COL1A1* rs1800012, *COL5A1* rs12722, *COL5A1* rs10628678, *COL6A1* rs35796750, *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 and *COL12A1* rs970547 in the control (CON), Rotator Cuff Tendinopathy (RCT) group and Supraspinatus Tendinopathy (SST) sub-group.

		CON	RCT	p-value [a]	SST	p-value [b]
***COL1A1* rs1800012 (G/T)**	*N*	94	101		77	
	GG	58.9 (56)	63.7 (64)	0.567	62.3 (48)	0.852
	GT	32.6 (30)	31.6 (32)		31.2 (24)	
	TT	8.4 (8)	4.9 (5)		6.5 (5)	
	T Minor allele	24.5 (46)	20.8 (42)	0.492	22.1 (34)	0.651
	HWE	0.258	0.761		0.504	
***COL5A1* rs12722 (T/C)**	*N*	94	102		77	
	TT	36.2 (34)	30.4 (31)	0.531	29.9 (23)	0.596
	TC	51.1 (48)	52.0 (53)		53.2 (41)	
	CC	12.8 (12)	17.6 (18)		16.9 (13)	
	C Minor allele	38.3 (72)	43.6 (89)	0.452	43.5 (67)	0.492
	HWE	0.379	0.689		0.642	
***COL5A1* rs10628678 (AGGG/-)**	*N*	95	101		76	
	AGGG/AGGG	43.2 (41)	49.5 (50)	0.519	55.3 (42)	0.236
	AGGG/(-)	47.4 (45)	44.6 (45)		39.5 (30)	
	(-)/(-)	9.5 (9)	5.9 (6)		5.3 (4)	
	(-) Minor allele	33.2 (63)	28.2 (57)	0.373	25.0 (38)	0.115
	HWE	0.477	0.460		0.768	
***COL6A1* rs35796750 (T/C)**	*N*	89	97		73	
	TT	34.8 (31)	27.8 (27)	0.480	27.4 (20)	0.592
	TC	46.1 (41)	54.6 (53)		52.1 (38)	
	CC	19.1 (17)	17.5 (17)		20.5 (15)	
	C Minor allele	42.1 (75)	44.8 (87)	0.304	46.6 (68)	0.309
	HWE	0.411	0.498		0.815	
***COL11A1* rs3753841 (T/C)**	*N*	94	102		77	
	CC	40.4 (38)	37.3 (38)	0.741	41.6 (32)	0.858
	TC	43.6 (41)	49.0 (50)		45.5 (35)	
	TT	16.0 (15)	13.7 (14)		13.0 (10)	
	T Minor allele	37.8 (71)	38.2 (78)	0.976	35.7 (55)	0.635
	HWE	0.343	0.834		1.000	
***COL11A1* rs1676486 (C/T)**	*N*	93	99		75	
	TT	64.5 (60)	64.6 (64)	0.902	68.0 (51)	0.858
	CT	30.1 (28)	31.3 (31)		28.0 (21)	
	CC	5.4 (5)	4.0 (4)		4.0 (3)	
	C Minor allele	19.4 (38)	19.7 (39)	0.985	18.0 (27)	0.832
	HWE	0.495	0.834		1.000	
***COL11A2* rs1799907 (A/T)**	*N*	94	102		77	
	AA	45.7 (43)	55.9 (57)	0.241	58.4 (45)	0.208
	AT	44.7 (42)	39.2 (40)		36.4 (28)	
	TT	9.6 (9)	4.9 (5)		5.2 (4)	
	T Minor allele	31.9 (60)	24.5 (50)	0.156	25.0 (36)	0.098
	HWE	1.000	0.789		1.000	
***COL12A1* rs970547 (A/G)**	*N*	91	102		77	
	AA	59.3 (54)	60.8 (62)	0.967	59.7 (46)	0.954
	GA	33.0 (30)	32.4 (33)		33.8 (26)	
	GG	7.7 (7)	6.9 (7)		6.5 (5)	
	G Minor allele	24.2 (44)	23.0 (47)	0.838	25.0 (36)	0.958
	HWE	0.255	0.402		0.541	

From previous page:
Genotype and allele frequencies are expressed as percentages with the number of participants shown in parenthesis (n). Global *P*-values are given with bold typeset indicating significant differences ($P<0.05$). HWE, Hardy-Weinberg equilibrium.
AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 Polymorphism.
[a] = RCT vs. CON, [b] = SST vs. CON

4.3.6 Collagen Gene-Gene Interactions

4.3.6.1 Type V and XI Collagen Polymorphism and Gene-Gene Interactions

Inferred haplotypes constructed from *COL5A1* 3'-UTR have previously been reported to be associated with TEN (76) and CTS (94). Of the four possible allele combinations, only three were inferred at a haplotype frequency above 4% (Table B1.1). However, none of the inferred haplotypes constructed from *COL5A1* rs12722 (T/C) and *COL5A1* rs10628678 (AGGG/-) were found to be associated with modulating risk (Figure 4.9).

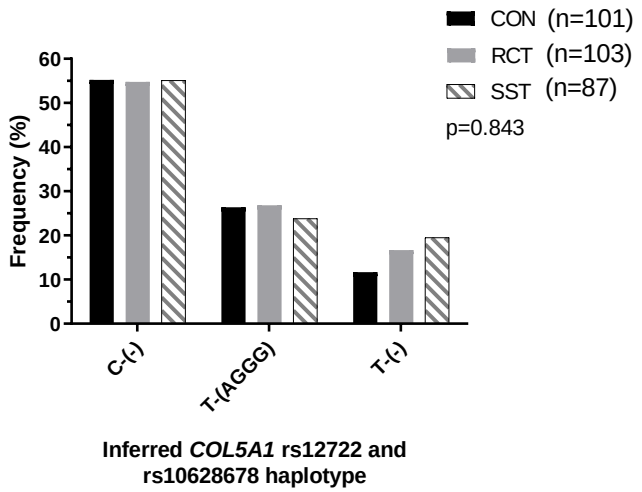


Figure 4.9. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) and *COL5A1* rs10628678 (AGGG/-) polymorphisms for the control (CON, black bars), the rotator cuff tendinopathy (RCT, grey bars) groups and SST sub-group (SST, hatched bars). Global p-value is indicated under the legend key. The number (n) of subjects in each group is in parenthesis. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

Since type XI collagen gene-gene interactions had previously been reported to be associated with chronic TEN (76) and CTS (98), inferred haplotypes were constructed between *COL11A1* rs3753841 (T/C) and *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (A/T). Of the six possible allele combinations of the three SNPs (Table B1.2), all were inferred at a haplotype frequency great than 4%, however none were found to be significant (Figure 4.11A).

Further, of the 4 possible allele combinations that were constructed using the *COL11A1* rs3753841 (T/C) and *COL11A1* rs1676486 (C/T) (Table B1.3), only three were inferred at a haplotype frequency greater than 4%. However, none of the allele combinations constructed were found to

be associated with risk between the groups (Figure 4.10 B). Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and *COL11A2* rs1799907 (A/T) polymorphisms (Figure 4.10 C) found all four combinations were inferred at a haplotype frequency greater than 4% (Table B1.4), however no of the combinations were found to be significant. However, a trend towards the C-T inferred haplotype being significantly over-represented in the CON group vs. the RCT (p=0.052) and the SST (p=0.068) groups was noted. Finally, inferred haplotype frequency distributions constructed from *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) polymorphisms (Figure 4.10 D). All combinations were inferred at a haplotype frequency greater than 4%. However, none were found to be significant, although a trend towards significance was noted. The T-A haplotype was shown to be trending towards a significant over-representation in the CON group vs. the RCT (p=0.090) and STT (p=0.074) groups.

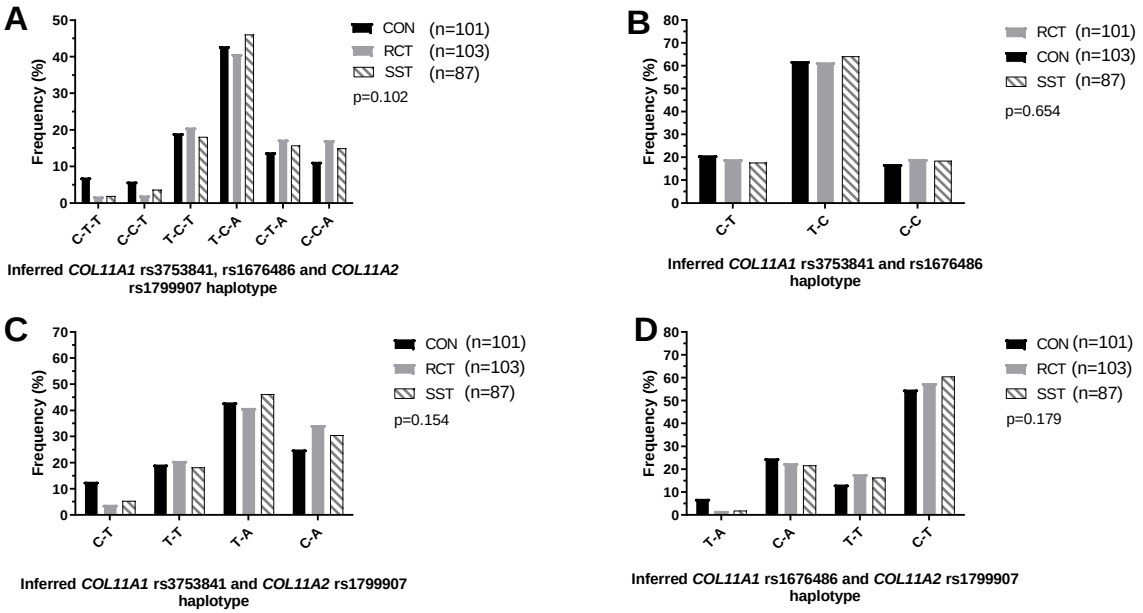


Figure 4.10 A) Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and rs1676486 (C/T), *COL11A2* rs1799907 (A/T) polymorphisms. **B)** Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and rs1676486 (C/T). **C)** Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and *COL11A2* rs1799907 (A/T) polymorphisms. **D)** Inferred haplotype frequency distributions constructed from *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) polymorphisms for the control (CON, black bars) and the rotator cuff tendinopathy (RCT, grey bars) groups and supraspinatus tendinopathy (SST, hatched bars). Trends are indicated with an # and the p-value. The number (n) of subjects in each group is in parenthesis.

Previous research has hypothesised that type XI collagen may interact with type V collagen during fibrillogenesis in developing tendons (252) and gene-gene interactions between the genes encoding the two collagen gene types have been reported to modulate the risk of chronic TEN (76) and CTS (98). Therefore, inferred pseudo-haplotypes were constructed between *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T), *COL11A2* rs1799907 (A/T) *COL5A1* rs10628678 (AGGG/-).

Only seven of 12 possible allele combinations were inferred at a frequency greater than 4% (Table B1.6). None of the possible allele combinations were found to be significantly associated with risk (Figure 4.11 A). Although trends toward significance were noted. The T-C-A-AGGG inferred haplotype demonstrated a trend towards significant over-representation in the SST group vs. the CON (p=0.055), however the same trend was not observed when compared to the RCT group (p=0.0654). Moreover, the C-C-A-AGGG inferred haplotype demonstrated a trend towards significant over-representation in the RCT group (p=0.057) and the SST group (p=0.051) when compared to the CON group.

Further, none of the inferred haplotype constructed from the two *COL11A1* and the *COL5A1* rs10628678 polymorphisms (Table B1.7) were significantly different between the groups (Figure 4.11 B). Further, inferred pseudo-haplotypes were constructed between *COL11A1* rs3753841 (T/C), *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) (Figure 4.11 C). All eight of the haplotypes constructed were inferred at a frequency greater than 4% (Table B1.8). It was found that the C-A-(-) inferred haplotype was significantly over-represented in the CON group compared to the SST group (6.0% vs. 0.4%, hap.score: -2.12, p=0.034, CON vs. SST, respectively). Although a significant difference was noted between the CON and RCT groups, the haplotype frequency was inferred at 3.2% (Table B1.8). Trends toward significance were noted in the T-A-AGGG haplotype whereby the SST trended towards over-representation vs the CON group (p=0.059), however a trend was not observed in the RCT group (p=0.589). Further, the C-T-AGGG inferred haplotype trended towards significance in the RCT (p=0.060) and SST (p=0.061) groups vs. the CON group. When haplotypes were constructed using the *COL11A1*

rs1676486 (C/T), *COL11A2* rs1799907 (A/T) *COL5A1* rs10628678 (AGGG/-), (Figure 4.11 D) only five of eight possible combinations were inferred at a frequency above 4% (Table B1.9). However, none of the possible haplotype combinations were found to be significant but a trend towards significance was observed. It was noted that the C-A-AGGG haplotype trended towards significant over-representation in the RCT (p=0.091) and SST (p=0.071) groups vs. the CON group.

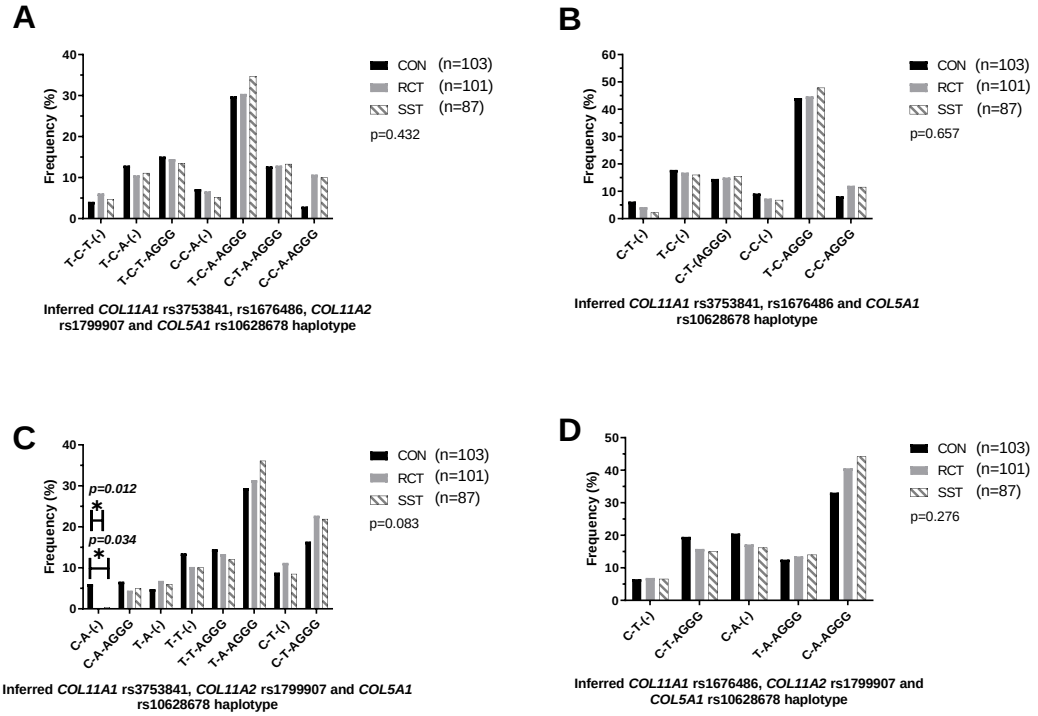


Figure 4.11 A) Inferred haplotype frequency distributions constructed from *COL5A1* rs10628678 (AGGG/-), *COL11A1* rs3753841 (T/C) and rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) polymorphisms. **B)** Inferred haplotype frequency distributions constructed from *COL5A1* rs10628678 (AGGG/-), *COL11A1* rs3753841 (T/C) and rs1676486 (C/T) polymorphisms. **C)** Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C), *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-). **D)** Inferred haplotype frequencies constructed from *COL11A1* rs1676486 (C/T), *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) polymorphisms for the control (CON, black bars) and the rotator cuff tendinopathy (RCT, grey bars) groups and supraspinatus tendinopathy (SST, hatched bars). Significant differences between the groups ($p<0.05$) are indicated with a solid line and the p-value. Trends are indicated with an # and the p-value. Global p-value is indicated below the legend. The number (n) of subjects in each group is in parenthesis. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

Finally, inferred haplotypes were constructed polymorphism from each of the type XI collagen gene polymorphisms and *COL5A1* rs10628678 (AGGG/-). All 4 of the possible allele combinations of the haplotypes constructed from *COL11A1* rs3753841 (T/C) and *COL5A1*

rs10628678 (AGGG/-) (Figure B1.9 A), *COL11A1* rs1676486 (C/T) (Figure B1.9 B) and *COL5A1* rs10628678 (AGGG/-) and *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) (Figure B1.9 C) were inferred at a frequency greater than 4% (Table B1.10). However, none of the combinations were found to be significant between the groups.

4.3.6.2 Types V and VI Collagen Gene-Gene Interactions

Previously, O'Connell et al (89) investigated the gene-gene interaction between *COL5A1* rs12722 (T/C) and *COL6A1* rs35796750 (T/C) and risk of Exercise Associated Muscle Cramps in Ironman Triathletes as well as triathlon performance (410). All four of the four possible allele combinations were inferred at a frequency greater than 4% (Table B1.13). However, none of the possible inferred allele combinations were found to be significant between any of the groups (Figure 4.12). Although, a trend towards significance was noted in the T-T inferred haplotype where the RCT (p=0.098) and SST (p=0.062) trended towards over-representation when compared to the CON group.

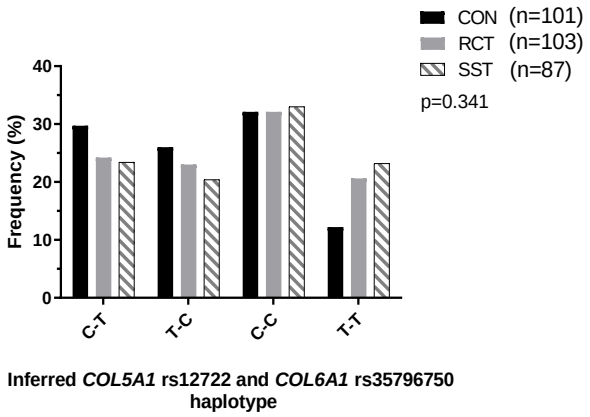


Figure 4.12. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) *COL6A1* rs35796750 (T/C) polymorphisms for the control (CON, black bars) and the rotator cuff tendinopathy (RCT, grey bars) groups and the SST sub-group (SST, hatched bars). Trends are indicated with an # and the p-value. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

4.3.6.3 Types V and XII Collagen Gene-Gene Interactions

In a study by O'Connell et al (89) investigating the interaction effect of collagen gene polymorphisms and risk of ACL rupture where inferred haplotypes were constructed using *COL5A1* rs12722 and *COL12A1* rs970547. All four possible allele combinations were inferred at a frequency greater than 4% (Table B1.14). However, none of the four possible inferred pseudo-haplotype combinations in *COL5A1* rs12722 (T/C) and *COL12A1* rs970547 (G/A) were found to be significant between the CON, RCT and SST sub-groups (Figure 4.13).

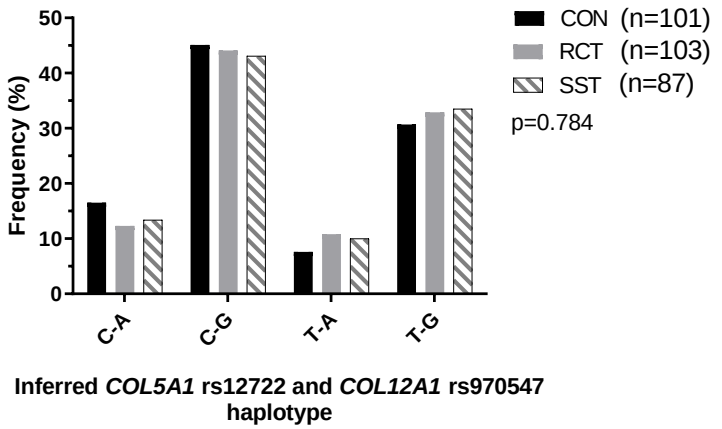


Figure 4.13 Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) and *COL12A1* rs970547 (G/A) polymorphisms for the control (CON, black bars) and the rotator cuff tendinopathy (RCT, grey bars) groups and supraspinatus tendinopathy (SST, hatched bars). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

4.3.6.4 Types VI and XII Collagen Gene-Gene Interactions

Previously, O'Connell et al (487) investigating genetic markers and range of motion in addition to endurance running performance (398) where inferred pseudo-haplotypes were constructed using *COL6A1* rs35796750 and *COL12A1* rs970547. All four of the possible allele combinations were inferred at a frequency greater than 4% (Table B1.15). None of the four possible inferred pseudo-haplotypes combinations in *COL6A1* rs35796750 (T/C) and *COL12A1* rs970547 (G/A) were found to be significant (Figure 4.14) between any of the groups.

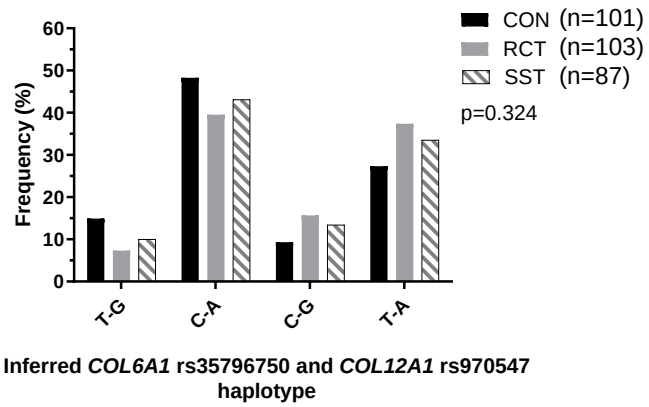


Figure 4.14. Inferred haplotype frequency distributions constructed from *COL6A1* rs35796750 (T/C) and *COL12A1* rs970547 (G/A) polymorphisms for the control (CON, black bars) and the rotator cuff tendinopathy (RCT, grey bars) groups and supraspinatus tendinopathy (SST, hatched bars). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

4.3.6.5 Types V, VI and XII Collagen Gene-Gene Interactions

Six of the possible allele combinations were inferred at a frequency greater than 4% (Table B3.15). None of the possible inferred pseudo-haplotypes combinations in *COL5A1* rs12722 (T/C), *COL6A1* rs35796750 (T/C) and *COL12A1* rs970547 (G/A) were found to be significant (Figure 4.15) between any of the groups.

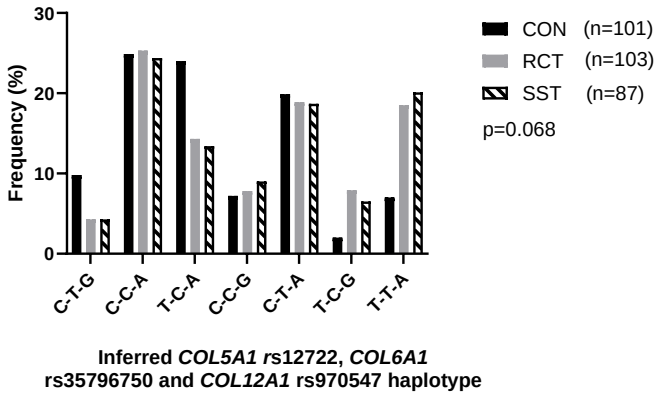


Figure 4.15. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (C/T) *COL6A1* rs35796750 (T/C), *COL12A1* rs970547 (G/A) polymorphisms for the control (CON, black bars) and the rotator cuff tendinopathy (RCT, grey bars) groups and supraspinatus tendinopathy (SST, hatched bars). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

4.4 DISCUSSION

The main finding of this chapter was that, although previously associated with other MSK injuries (21,37,99), none of the investigated collagen gene polymorphisms were independently associated with RCT risk. Of the several previously investigated collagen polymorphism interactions (76,94,98), only the C-A-(-) inferred haplotype constructed from *COL11A1* rs3753841(T/C), *COL11A2* rs1799907 (C/T) and *COL5A1* rs10628678 (AGGG/-) was found to be significantly over-represented in the CON group compared to the SST sub-group. Although the RCT group was also found to be significantly different to the CON group, the haplotype frequency was inferred below 4%. This is in agreement with previous studies where the complementary T-C-T-(AGGG) inferred haplotype constructed from *COL11A1* rs3753841(T/C), *COL11A1* rs1676486 (C/T),

COL11A2 rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) and the T-C-(AGGG) inferred haplotype constructed from *COL11A1* rs3753841(T/C), *COL11A1* rs1676486 (C/T) and *COL5A1* rs10628678 (AGGG/-) was associated with increased risk chronic TEN (76)and CTS respectively (98). Although not always in total alignment with the previously reported T-C-T-(AGGG) inferred risk haplotype, as summarised in Table 4.6 there were trends for some of the other haplotype combinations constructed from *COL11A1* rs3753841(T/C), *COL11A1* rs1676486 (C/T), *COL11A2* rs1799907 (A/T) and/or *COL5A1* rs10628678 (AGGG/-) to modulate risk of RCT and/or SST in this study.

It has been previously shown that type V collagen interacts with type XI collagen (103,252) during fibrillogenesis during tendon development. Type V and type XI collagen are minor fibrillar collagens that share many similar structures and functions (253) and have been shown to initiate fibril nucleation of collagens during early fibrillogenesis and have a large role in regulating fibril diameter (103,356). Previously is has been suggested that the similarities between the α1 chains encoded by the *COL11A1* and *COL5A1* genes, allow the two polypeptide chains to interchange which results in the formation of hybrid collagen proteins (98,216,488). Therefore, it has been hypothesized that these hybrid collagens are able to supplement or replace the existing type V collagen located in tendons (103). In so doing, it is possible that individuals who possess these functional polymorphisms could produce type V and XI collagens, in conjunction to altered functionality, could be responsible for changes in the biomechanical properties of the tendon which could result in pathology.

As previously described in Chapter 2, section 2.7.4, the rs1676486 (C>T) polymorphism of the α1(XI) chain is located within exon 62. This results in proline being substituted to serine at the

amino acid position of 1535 (400). Mio et al. (400) reported that T allele of the *COL11A1* rs1676486 was associated with an increased in mRNA degradation of the type XI collagen, resulting in a decrease in α1(XI) chain production. Therefore, changes in the combination of alleles within these polymorphisms could potentially result in changes to the structural and functional properties of the collagen fibril (98,409).

Furthermore, we hypothesized that haplotype constructed from the *COL5A1* rs10628678 (AGGG/-) and rs12722 (T/C), which has been previously reported to be associated with AT (385) and CTS (94), would be associated with RCT in the current study. Although, no significant difference was found between the CON, RCT or SST groups, the *COL5A1* rs10628678 (AGGG/-) and rs12722 (T/C) have been theorised to be involved in altering the stability of *COL5A1* messenger RNA (mRNA) (386). More specifically, the 3′-UTR containing rs10628678 region has been shown to be functional and implicated in the aetiology of AT (489). The T allele of rs12722 and the AGGG allele of rs10628678 were shown to associated with increased mRNA stability, leading to the hypothesis that an increased type V collagen production, resulting in a reduced fibril diameter (389,489,490). As a result, this could potentially contribute to affect the mechanical properties of MSK tissues (491). Considering the limited space within a shoulder capsule, changes in the diameter of tendons could possibly contribute to pathogenesis including increased mechanical impingement.

As previously discussed in Chapter 2, section 2.3, and in agreement with Posthumus et al. (87), participants who reported having sustained a previous joint injury or a history of multiple tendon injuries were 2.3x greater risk of RCT and 2.1x great risk of SST. Although a family history of MSK injuries has been shown to increase the risk of injury (87,141,166), the current study did not

find any significant differences between CON, RCT and SST. However, a trend towards significance was noted in the RCT and SST group and a family history of tendon injuries.

To the best of our knowledge, this is the first study to investigate multiple collagen gene polymorphisms in a RCT cohort. As is discussed in Chapter 3, swimming is a sport that places an incredibly high demand on the rotator cuff complex (48,55,492). One of the main risk factors for injury is training load. Within this cohort, both RCT and SST groups trained significantly more hours per week than the CON group. Both the RCT and SST groups reported training significantly more hours per week in both summer and winter compared to the CON group. During training, a significant amount of repetitive load is placed on the shoulder, resulting in approximately 2500-9600 overhead rotations completed per day (493). As a result, several possible physiological adaptations have been noted in the literature, which include tendon thickness (59,493,494), tendon degeneration (433), reduced joint stability (204,205), increased range of motion (51,441,442), muscle fatigue (495), altered scapular kinematics (434,444) and increased mechanical impingement (206,433,466) (Figure 4.15). These adaptations have previously associated with tendinopathy (146,421,496).

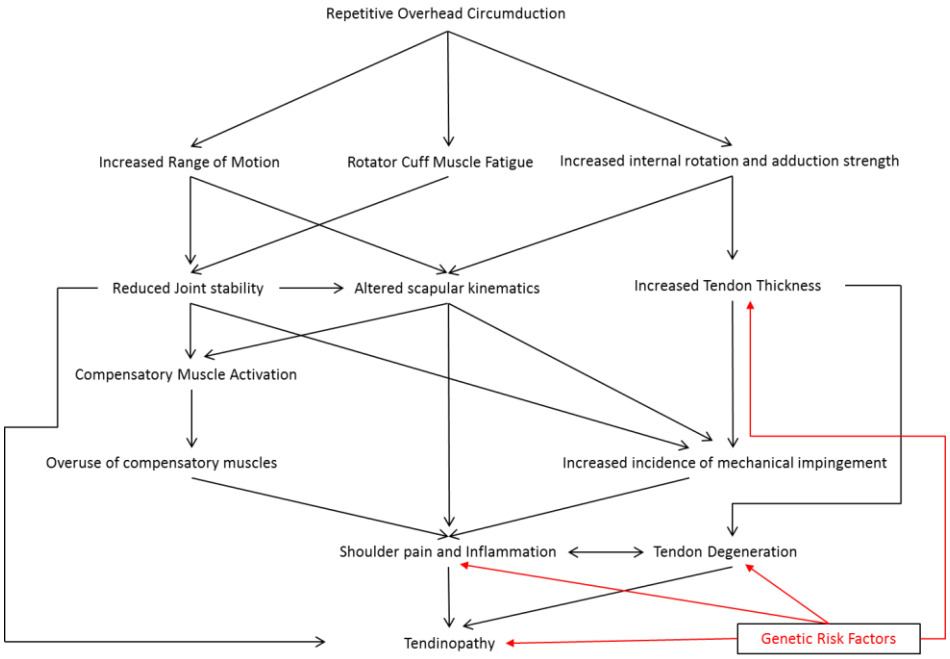


Figure 4.15. A schematic diagram of the proposed mechanisms that predispose a swimmer to RCT.

Accordingly, Sein et al. (59) was demonstrated that intensive and repetitive swimming training lead to increased tendon thickness of rotator cuff tendons in elite swimmers. Changes in tendon thickness have been shown to alter the mechanical properties of the tendon and the increase the incidence of mechanical impingement (59,433,496,497). Further, chronic loading of tendons without adequate recovery has been shown to contribute to overuse tendinopathy (45,146). If managed correctly, reactive tendinopathy as suggested by Cook and Purdam (40) may result in short term adaptive changes to the molecular components of the tendon. However, if the load continues without modification, the result is tendon disrepair including further extracellular matrix breakdown (40). This stage has been reported in chronically overloaded tendons but may appear across a spectrum of ages and loading environments (498). Therefore, the combination of intrinsic

and extrinsic factors may contribute to the onset of injury risk. Although in this chapter, none of the investigated polymorphisms were associated with risk, it does not exclude the potential that an associated exists. Genes that modulate structure and function of these collagens within the rotator cuff tendons are vital in determining risk of risk of injury.

Although none of the investigate polymorphisms were independently associated with RCT risk, on-broader investigation one inferred pseudo-haplotype was to be associated with reduced risk. Interestingly, a trend towards significance was found in several of the inferred haplotypes. It possible that the small sample of this cohort limited our ability to detect meaningful associations.

In summary, this study offers evidence indicating that the different collagens composing the tendon fibre of the RC interact with each other, both individually and collectively, influencing the risk outcomes of RCT. However, this does not rule out the possibility of other genetic variants potentially playing a role in the aetiology of RCT. Further research is needed to comprehend the functional mechanisms underlying this intricate phenotype and the genetic associations, as well as the gene-gene interactions, that may influence the risk of this pathology. A continuous summary of the results of this thesis are presented in Table 4.5 and Table 4.6.

Table 4.6: A continuing summary of the genotype results from each chapter in this thesis. The results of Chapter 4 are shown.

Investigated Gene Variant	Investigated Phenotype			
	Chapter 4		Chapter 5	Chapter 6
	RCT	SST	EA ACL	MA ACL
COL1A1 rs1800012 (G/T)	n.s	n.s / Prev pub CC (30)	Prev pub TT (84)	
		n.s		
			Prev pub GG (90)	
COL5A1 rs12722 (T/C)	n.s	n.s	Prev pub CC (73)	
COL5A1 rs10628678 (AGGG/-)	n.s	n.s		
COL6A1 rs35796540 (C/T)	n.s	n.s	Prev pub n.s (89)	
COL11A1 rs3754841(T/C)	n.s	n.s		
COL11A1 rs16746744 (C/T)	n.s	n.s		
COL11A2 rs1799907 (A/T)	n.s	n.s		
COL12A1 rs970547 (G/A)	n.s	n.s	Prev pub AA (87,405)	

Green shading indicates a reduction in injury risk. Red shading indicates an increased risk of injury. Grey shading indicates no association with risk. Prev pub; Previously published. RCT; rotator cuff tendinopathy. SST; Supraspinatus tendinopathy. EA; European Ancestry. ACL; Anterior cruciate ligament rupture. MA; Mixed Ancestry.

Table 4.7: A continuing summary of the haplotype results from each chapter in this thesis. The results of Chapter 4 are shown.

Phenotype		Investigated Gene variants						
		COL11A1 rs3753841 (T/C)	*COL11A1* rs1676486 (C/T)	*COL11A2* rs1799907 (A/T)	*COL5A1* rs12722 (T/C)	*COL5A1* rs10628678 (AGGG/-)	*COL6A1* rs35796750 (C/T)	*COL12A1* rs970547 (A/G)
Chapter 4	RCT SST	C		A	(-) (Thesis)			
Chapter 5	EA ACL	Prev pub C	T	A (409)				
		Prev pub C	T (409)					
		Prev pub T	C	T	AGGG (409)			
		Prev pub T	C	T	(-) (409)			
		Prev pub C	T	A	AGGG (409)			
		Prev pub T	C	A	AGGG (409)			
					Prev pub T			A (89)
Chapter 6	MA ACL							

Green shading indicates a reduction in injury risk. Red shading indicates an increased risk of injury. Grey shading indicates no association with risk. Prev pub; Previously published. RCT; rotator cuff tendinopathy. SST; Supraspinatus tendinopathy. EA; European Ancestry. ACL; Anterior cruciate ligament rupture. MA; Mixed Ancestry.

CHAPTER 5: COLLAGEN GENE RISK FACTORS FOR ACL RUPTURE IN A SWEDISH AND COMBINED POPULATIONS

5.1 INTRODUCTION

ACL ruptures are the most common knee ligament injury with an incidence of 1 in every 3500 people (33). It is a multifactorial injury, for which various extrinsic and intrinsic risk factors, including genetic factors, have been identified (17,409). As reviewed in Chapter 2 section 2.7, a genetic component, including polymorphisms within collagen encoding genes, has been found to associate with several chronic tendon injuries including TEN (76,81,82), rotator cuff (27,29,30,180,183,190), CTS (94,98), and tennis elbow (24,25). Further, several studies have also found association between collagen encoding genes and the risk of ACL rupture (27,77,84,85,87,499), suggesting some similarities between genetic risk for acute ligament and chronic tendon injuries (Chapter 2, section 2.7.1).

Most studies investigating populations with European ancestry have reported an independent association of the rare *COL1A1* rs1800012 TT genotype with decreased ACL rupture risk (84,380,500,501). In a combined analysis containing previously published data (500) consisting of 1427 asymptomatic controls and 407 participants with non-contact ACL rupture, the TT genotype was significantly associated with reduced risk of ACL rupture. Only a single cohort has reported an association between the rs12722 polymorphism within the functional *COL5A1* 3'-untranslated region (UTR) and ACL rupture in a population with European ancestry. Specifically, the CC genotype was associated with reduced risk of ACL rupture in South African females (73,502). This independent association has however has not been repeated in other cohorts,

including the cohort investigated in this chapter of the thesis (502,503). Although no independent associations were observed for *COL5A1* rs12722 or rs10628678 (AGGG/-) in a combined analysis, the C-(-) haplotype was under-represented and T-(-) haplotype over-represented in the female ACL combined group (502). These and a previous study investigating the association of *VEGFA* polymorphisms in a larger combined cohorts (240) highlights the limitations of testing associations in multiple populations of limited sample populations.

Although the *COL12A1* rs970547 AA genotype has been reported to be independently associated with increased risk of ACL rupture (87), this *COL12A1* polymorphism was also not independently associated with ACL rupture in male Polish football players (88) nor Scandinavian female elite team athletes (503). The inferred T-A haplotype constructed from *COL5A1* rs12722 and *COL12* rs970547 was associated with increased risk of ACL rupture in Polish and South African Cohorts (89). Finally, the *COL6A1* rs3579650 polymorphism was not associated with ACL rupture in a single study (89).

As summarised above and extensively reviewed in Chapter 2, section 2.7, the association of several collagen gene polymorphisms with ACL rupture have been investigated in several populations of European ancestry. Except for *COL1A1* rs1800012, the association of other collagen variants is less clear. Therefore, the aim of this chapter was to investigate the independent association of the previously investigated *COL1A1* rs1800012, *COL5A1* rs12722, rs10628678, *COL6A1* rs3579650, *COL11A1* rs3753841, rs1676841, *COL11A2* rs1799907 and *COL12A1* rs970547 polymorphisms with risk of ACL rupture in a Swedish population and more importantly a combined cohort containing previously published populations of European ancestry to increase the sample size.

Finally, the association of collagen gene-gene and haplotype interactions with risk of ACL rupture was also investigated in the Swedish and the combined population.

5.2 METHODS

This chapter followed the methodology as previously outlined in Chapter 4, section 4.2.1. This study was approved by the University of the Faculty of Health Sciences Human Research Ethics Committee within the University of Cape Town, South Africa (HREC 269/2014) (Appendix A1.9) and the and the Regional Ethical Review Board in Umeå, Sweden (dnr. 2011-200-31M) (Appendix A1.10).

5.2.1 Swedish population Participants

One-hundred and ninety-five (n=195) physically active and unrelated participants of self-reported European ancestry between the ages of 19 and 65 years were recruited between 2011 and 2013 from the University Hospital in Umeå and orthopaedic clinics in Luleå, Sweden (642,645–647) using previously described inclusion and exclusion criteria (504). These participants within this cohort comprised of 79 individuals who had clinically diagnosed ACL injuries with a non-contact mechanism of rupture (NON group) and 116 apparently healthy, asymptomatic individuals with no history of ACL injuries (CON group). The NON participants were recruited from a long-term follow ACL rupture study (504–506). ACL ruptures were diagnosed based on physical examination, magnetic resonance imaging and arthroscopically confirmed at the University Hospital in Umeå. All surgeries were done at the University Hospital in Umeå, Sweden.

All participants were given full written and verbal explanation of the study (Appendix A1.11) and were required to complete a written informed consent form (Appendices A1.12, A1.14 and A1.15),

in accordance with the Declaration of Helsinki (472), and were requested to complete a questionnaire regarding personal details, medical history, personal and family ligament and tendon rupture history, as well as sports participation (Appendix A1.13).

5.2.2 DNA Extraction and genotyping

DNA samples from participants were extracted from venous blood at the University of Umeå (Sweden) and an aliquot was sent to the University of Cape Town (South Africa) for genotyping and subsequent analysis. All participants were genotyped for the *COL1A1* rs1800012 (G/T), *COL5A1* rs10628678 (-/AGGG), *COL6A1* rs35796750 (T/C), *COL11A1* rs3753841 (T/C) and rs1676486 (C/T), *COL11A2* rs1799907 (A/T) and *COL12A1* rs970547 (G/A) polymorphisms as described in in section 4.2.4 of the previous chapter. The *COL5A1* rs12722 (T/C) polymorphism was previously genotyped by Suijkerbuijk et al. (507) and included in this study to construct inferred haplotypes.

To ensure genotyping accuracy and as a preventative measure to detect contamination, positive and negative controls were used. No discrepancies were observed. A total of 209/226 (92.4%), 203/226 (89.8%), 206/226 (91.2), 209/226 (92.4%), 209/226 (92.4%), 209/226 (92.4%), 210/226 (92.9%) and 205/226 (90.7%) genotypes were identified for *COL1A1* rs1800012, *COL5A1* rs12722 and rs10628678, *COL6A1* rs35796750, *COL11A1* rs3753841 and rs1676486, *COL11A2* rs1799907 and *COL12A1* rs970547 polymorphisms, respectively.

5.2.3 Combined Analysis

The published genotype data from cohorts consisting of individuals with self-reported European ancestry were combined with the data included in this current study, to further analyse the

independent association of the for 1) *COL1A1* rs1800012, 2) *COL5A1* rs12722, 3) *COL5A1* rs10628678 4) *COL6A1* rs35796750 5), *COL11A1* rs3753841, 6) *COL11A1* rs1674684, 7) *COL11A2* rs1799907 and 8) *COL12A1* rs970547 polymorphisms with risk of ACL injuries in a larger sample size (Table 5.1). Although skiing mechanisms of rupture were excluded, this analysis nevertheless included a mixed mechanism of ACL rupture (ACL Group). Males and females were also analysed separately.

Table 5.1: Summary of previously investigated European Ancestry Anterior Cruciate Ligament (ACL) rupture and related healthy Control (CON) populations.

COHORT	*COL1A1* rs1800012	*COL5A1* rs12722	*COL5A1* rs10628678	*COL6A1* rs35796750	*COL11A1* rs3753841	*COL11A1* rs1676486	*COL11A2* rs1799907	*COL12A1* rs970547
Polish (88,508)	CON: 143 ACL: 91	CON: 134 ACL: 211						CON: 143 ACL: 91
RSA (76,89)		CON: 231 ACL: 224	CON: 216 ACL: 214	CON: 199 ACL: 119	CON: 206 ACL: 203	CON: 203 ACL: 209	CON: 192 ACL: 190	CON: 226 ACL: 220
Norwegian & Finnish (503)	CON: 732 ACL: 119	CON: 732 ACL: 119						CON: 732 ACL: 119
Swedish (507)		CON: 109 ACL: 77						
RSA, Polish, Swedish, Norwegian & Finnish (500)	CON: 407 ACL: 1425							
Australian (502)		CON: 79 ACL: 354	CON: 80 ACL: 356					

RSA; South African, CON; Healthy Controls, ACL; Anterior cruciate ligament rupture

5.2.4 Statistics

The same statistical tests and haplotype analysis methods were performed as described in Chapter 4, section 4.2.5. Significance was accepted when $p<0.05$. No adjustments were made for multiple testing in this study for the same reasons as outlined in Chapter 4, section 4.2.6.

5.3 RESULTS

5.3.1 Participant Characteristics

Participants in the Swedish cohort NON and CON groups were similarly matched for weight, height, BMI and Country of birth (Table 5.2). Participants in the CON group were significantly older (44.7 ± 11.9, n=114, p<0.001) than the NON group (36.5 ± 13.7, n=78). Furthermore, males were significantly under-represented in the CON group (34.4%, n=40) compared to the NON group (54.4%, n=42, p=0.014).

Table 5.2: General participant characteristics for the Swedish CON and NON groups*.

	CON (n=116)	**NON** (n=79)	p value
Age (years)$^{\alpha}$	44.7 ± 11.9 (114)	36.5 ± 13.7 (78)	**p<0.001**
Sex (% male)	34.5 (40)	54.4 (42)	**0.014**
Height (cm)	172.3 ± 10.1 (108)	173.4 ± 8.6 (71)	0.438
Weight (kg)$^{\alpha}$	72.1 ± 13.6 (107)	75.0 ± 12.8 (71)	0.149/0.688[a]
BMI (kg.cm^{-2})	24.4 ± 2.9 (108)	24.7 ± 2.9 (70)	0.466/0.427[a]
COB (% Swedish)	94.9 (110)	89.9 (71)	0.154

* Results previously reported by Suijkerbuijk and Ponzetti et al. (507)
Data reported as mean ± SD,
Sex and county of birth (COB) are reported as frequency (%).
The number of participants with available data for each variable is reported in parentheses.
Age, weight and BMI are self-reported values at the time of ACL rupture.
[a] indicates adjusted for age and sex. Significant p values are noted in bold.
α age and weight at rupture for the NON group; and at recruitment for CON group

5.3.2 Participant and Family History of Musculoskeletal Soft Tissue Injuries

Participants who reported a family history (first degree relatives) of at least one joint injury at the time of recruitment was significantly higher in the NON group (p=0.014) when compared to CON

group (Table 5.3). Joint injuries included injuries to the knee, ankle, toe, elbow, wrist, finger, spine or shoulder and were self-reported by the participants on behalf of a relative. Additionally, there was a significantly higher reported history of parents with a history of at least one joint injuries in the NON group (p=0.018) when compared to the CON group. Similarly, reported history of sibling with history of at least one joint injury was significantly higher in the NON group (p=0.003) when compared to the CON group. Familial history of NON injuries (including parent, sibling and child) were not significantly different between the CON and NON groups. Additionally, previous injuries to multiple joints were not significantly different between groups.

Table 5.3. Medical history and family injury for CON and NON groups of the Swedish cohort*.

	Male			**Female**			**All**
	CON (n= 40)	NON (n= 42)	p-value[a]	CON (n= 76)	NON (n= 37)	p-value[a]	p-value[b]
Previous ligament injury	88.6 (35)	100.0 (35)	0.114	73.5 (68)	81.8 (33)	0.458	**0.035**
Previous joint injury	35.1 (37)	51.3 (39)	0.235	37.5 (72)	41.7 (36)	0.834	0.230
Family history of ACL injury	15.2 (33)	24.3 (37)	0.384	17.4 (69)	33.3 (33)	0.121	0.093
• Grandparent	0.0 (33)	0.0 (37)	-	0.0 (69)	3.0 (33)	0.323	0.407
• Parent	0.0 (33)	2.7 (37)	1.000	7.2 (69)	12.1 (33)	0.466	0.531
• Sibling	12.1 (33)	10.8 (37)	1.000	2.9 (69)	6.1 (33)	0.593	0.551
• Child	3.0 (33)	10.8 (37)	0.361	7.2 (69)	6.1 (33)	1.000	0.551
	0.0 (33)	0.0 (37)	-	0.0 (69)	6.1 (33)	0.103	0.164
Family history of joint injury*	59.5 (37)	69.2 (39)	0.516	47.2 (72)	72.2 (36)	**0.024**	**0.014**
• Parent	21.6 (37)	46.2 (39)	**0.031**	31.9 (72)	47.2 (36)	0.181	**0.018**
• Sibling	43.2 (37)	48.7 (39)	0.804	15.3 (72)	44.4 (36)	**0.002**	**0.003**
• Child	27.0 (37)	17.9 (39)	0.415	20.8 (72)	19.4 (36)	1.000	0.608

* results previously reported by Suijkerbuijk and Ponzetti et al (507).

*Joint injuries included injuries to the knee, ankle, toe, elbow, wrist, finger, spine or shoulder and were self-reported by the participants on behalf of a relative.

Values are expressed as percentages with the number of participants (n) with available data in parentheses.

[a] CON vs. NON, p-values in bold typeset indicate significance ($p < 0.050$).

[b] CON (male + female) vs. NON (male + female), p-values in bold typeset indicate significance ($p < 0.050$).

5.3.3 Independent Genotype Analysis

None of the seven investigated polymorphisms (*COL1A1* rs1800012, *COL5A1* rs10628678, *COL6A1* rs35796750, *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 and *COL12A1* rs970547) were independently associated with ACL rupture in the Swedish population (Table 5.4). All polymorphisms were found to be in HWE. Additionally, none of the alleles' frequencies were found to be significantly different between the groups. It was considered inappropriate to stratify genotype results by sex as it would render the sample size too small to reflect any meaningful associations.

Table 5.4. Genotype frequency distributions for *COL1A1* rs1800012, *COL5A1* rs10628678, *COL6A1* rs35796750, *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 and *COL12A1* rs970547 in the Swedish control (CON) and Swedish non-contact anterior cruciate ligament injuries (NON).

SNP	Allele	CON (117)	NON (79)	*P*-value
	N	106	73	
	G/G	61.3 (65)	71.2 (52)	0.556
***COL1A1* rs1800012**	G/T	33.0 (35)	26.0 (19)	
(G/T)	T/T	5.7 (6)	2.7 (2)	
	T Allele	22.2 (47)	15.8 (23)	0.160
	HWE	0.585	1.000	
	N	105	73	
	AGGG/AGGG	48.6 (51)	41.1 (30)	0.501
***COL5A1* rs10628678**	AGGG/-	35.2 (37)	43.8 (32)	
(AGGG/-)	-/-	16.2 (17)	15.1 (11)	
	(-) Allele	33.8 (71)	36.9 (54)	0.510
	HWE	0.130	0.803	
	N	59	59	
	T/T	42.4 (25)	40.7 (24)	0.981
***COL6A1* rs35796750**	T/C	47.5 (28)	49.2 (29)	
(T/C)	C/C	10.1 (6)	10.1 (6)	
	C Allele	33.9 (40)	34.7 (41)	0.857
	HWE	0.434	0.575	
	N	106	73	
	T/T	40.6 (43)	37.0 (27)	0.685
***COL11A1* rs3753841**	T/C	46.2 (49)	45.2 (33)	
(T/C)	C/C	13.2 (14)	17.8 (13)	
	C Allele	36.3 (77)	40.4 (59)	0.305
	HWE	1.000	0.809	
	N	107	73	
	C/C	60.7 (65)	54.8 (40)	0.651
***COL11A1* rs1676486**	C/T	32.7 (35)	38.4 (28)	
(C/T)	T/T	6.5 (7)	6.8 (5)	
	T Allele	22.9 (49)	26.0 (38)	0.472
	HWE	0.419	1.000	
	N	107	73	
	A/A	52.3 (56)	58.9 (43)	0.082
***COL11A2* rs1799907**	A/T	41.1 (44)	27.4 (20)	
(A/T)	T/T	6.5 (7)	13.7 (10)	
	T Allele	27.1 (58)	27.8 (40)	0.884
	HWE	0.808	0.080	
	N	101	66	
	A/A	62.4 (63)	59.1 (39)	0.832
***COL12A1* rs970547**	A/G	33.7 (34)	37.9 (25)	
(A/G)	G/G	4.0 (4)	3.0 (2)	
	G Allele	20.8 (42)	22.0 (29)	0.834
	HWE	0.767	1.000	

Genotype and allele frequencies are expressed as percentages with the number of participants shown in parenthesis (n). Global *P*-values are given with bold typeset indicating significant differences ($P<0.05$).
HWE, Hardy-Weinberg equilibrium.
AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 Polymorphism.

5.3.4 Combined Genotype Analysis

However, when the *COL1A1* rs1800012 genotype data from this study was combined with previously reported cohorts of European ancestry (Table 5.1), the TT genotype was significantly under-represented (p = 0.027, OR = 0.45, 95% CI 0.20 – 1.02) in the combined ACL group (GG 68.0%, GT 30.9%, TT 1.1%) compared to the combined CON group (GG 67.0%, GT 29.0%, TT 4.0 %) when both male and female participants were included in the analysis (Figure 5.1 A). Interestingly, when analysed separately the TT genotype was significantly under-represented in the combined female (p = 0.045, OR = 0.00, CI 0.00 – 0.71; ACL: GG 68.1%, GT 31.9%, TT 0.0%; CON: GG 68.3%, GT 27.0%, TT 4.8%) but not male (p =0.299; ACL: GG 68.0%, GT 39.5%, TT 1.5%; CON: GG 66.5%, GT 31.0%, TT 3.6%) ACL groups (Figure 5.1 B and C).

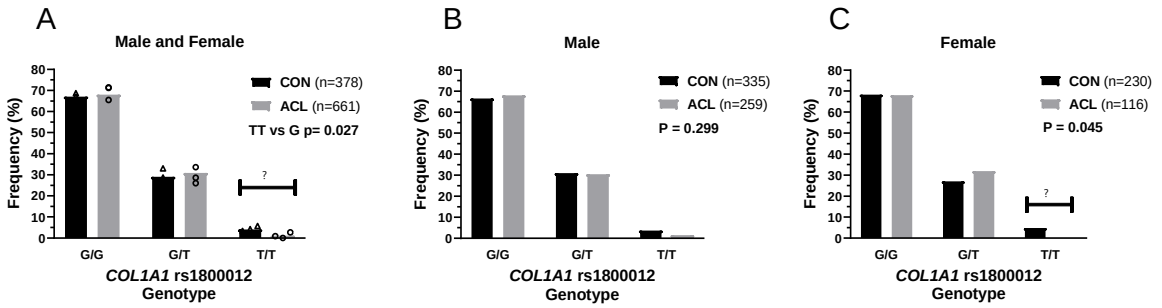


Figure 5.1 The *COL1A1* rs1800012 (G/T) genotype distributions of the **(A)** male and female, **(B)** male only and **(C)** female only combined CON (white bars) and ACL (grey bars) groups of South African, Polish, Swedish, Norwegian, Finish and Australian participants with self-reported European ancestry. The genotype distributions of each induvial cohort are shown as symbol in only the combined male and female panel. Sample sizes were too small for most cohorts for individual cohort sex sub-analysis. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

Interestingly, the AGGG/AGGG genotype was significantly over-represented (p=0.035) in the CON group when all participants were analysed, but not when males and females were analysed separately. There were no significant differences (p = 0.060) between the *COL5A1* rs12722 genotype distribution of a combined analysis with previously reported cohorts of European ancestry consisting of Polish, South African, Norwegian, Finish, Australian and Swedish populations between the ACL group (TT 29.1% + TC 49.9% vs CC 21.0%) compared to CON group (TT 32.2% + TC 48.4% vs CC 19.5%) (Figure 5.2A). Similarly, there were also no significant differences (p = 0.074) between the *COL5A1* rs10628678 genotype distribution of the ACL (AGGG/AGGG 45.9% + AGGG/- 44.0% vs -/- 10.2%) and CON (AGGG/AGGG 52.2% + AGGG/- 37.7% vs -/- 10.1%) groups consisting of South African, Australian, and Swedish populations (Figure 5.2D). There were also no significant differences in the genotype distributions

of either *COL5A1* polymorphisms when the male (Figure 5.2B & E) and female (Figure 5.2C & F) participants when analysed separately.

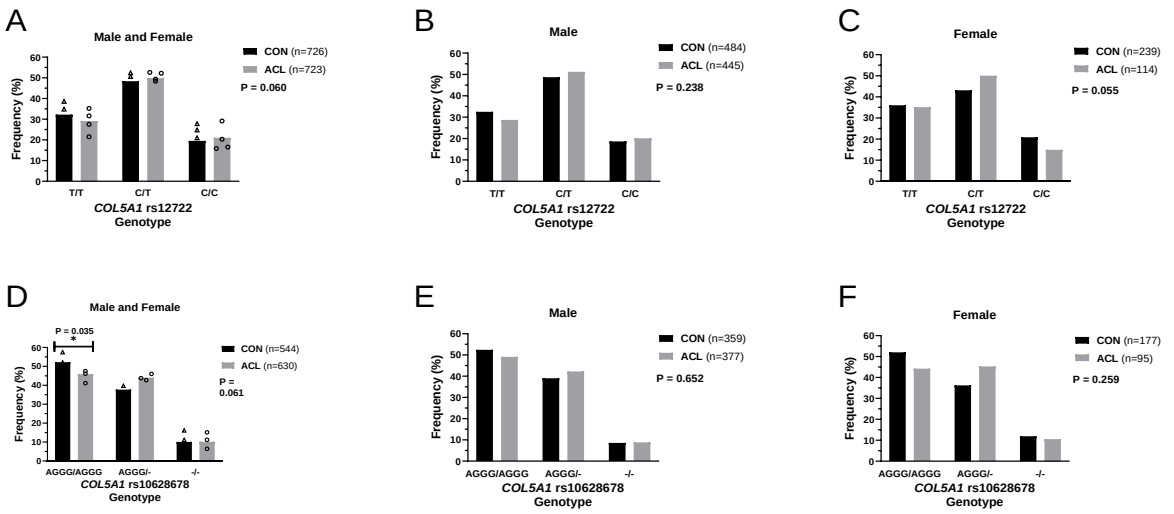


Figure 5.2 The *COL5A1* rs12722 genotype distributions of the **(A)** male and female, **(B)** male only and **(C)** female only combined CON (white bars) and ACL (grey bars) groups of South African, Polish, Swedish, Norwegian, Finish and Australian participants with self-reported European ancestry. The *COL5A1* rs10628678 genotype distributions of the **(D)** male and female, **(E)** male only and **(F)** female only combined CON (white bars) and ACL (grey bars) groups of South African, Swedish and Australian participants with self-reported European ancestry. The genotype distributions of each induvial cohort are shown as symbol in only the combined male and female panel. Sample sizes were too small for most cohorts for individual cohort sex sub-analysis. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

There were also no significant differences (p=0.394) in the *COL6A1* rs35796750 genotype distribution of a combined analysis with previously reported cohorts of European ancestry consisting of South African and Swedish populations between the ACL group (TT 33.5%, TC 52.5%, CC 12.5%) compared to the combined CON group (TT 36.8%, TC 46.6%, CC 16.7%)

(Figure 5.3 A). Further, male, and female participants were not found to be significantly different between the ACL and CON group (Figure 5.3 B and C).

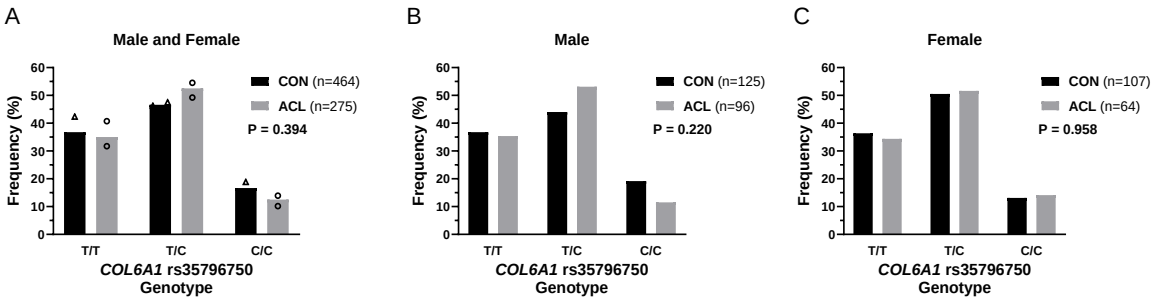


Figure 5.3 The *COL6A1* rs35796750 genotype distributions of the **(A)** male and female, **(B)** male only and **(C)** female only combined CON (white bars) and ACL (grey bars) groups of South African and Swedish participants with self-reported European ancestry. The genotype distributions of each induvial cohort are shown as symbol in only the combined male and female panel. Sample sizes were too small for most cohorts for individual cohort sex sub-analysis. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

No significant differences were found *COL11A1* rs3753841 genotype distribution (p = 0.131) of a combined analysis with previously reported cohorts of European ancestry consisting of South African and Swedish populations between the ACL group (Figure 5.4 A, TT 37.1%, TC 45.1%, CC 17.8%) and the CON group (TT 34.3%, TC 52.6%, CC 13.1%). Similarly, male and female participants were not found to be significantly different between ACL and CON groups (Figure 5.4 B and C).

The *COL11A1* rs1676486 genotype distribution (p = 0.440) of a combined analysis with previously reported cohorts of European ancestry also consisting of the same South African and Swedish

populations was not found to be significantly different between ACL (Figure 5.4 D, CC 64.9%, CT 30.1%, TT 5.1%) and CON (CC 59.8%, CT 34.8%, TT 5.4%) groups. Genotype distributions within the male-only and female-only sub-groups were also not found to be significantly different (Figure 5.4 E and F).

Similarly, no significant differences were found *COL11A2* rs1799907 genotype distribution ($p = 0.764$) of a combined analysis with same previously reported cohorts of European ancestry between the ACL group (Figure 5.4 G, AA 43.5%, AT 45.0, TT 11.5%) and CON group (AA 46.5%, AT 43.1%, TT 10.4%). In addition, male-only and female-only analysis found no difference in genotype distributions between ACL and CON groups (Figure 5.4 H and I).

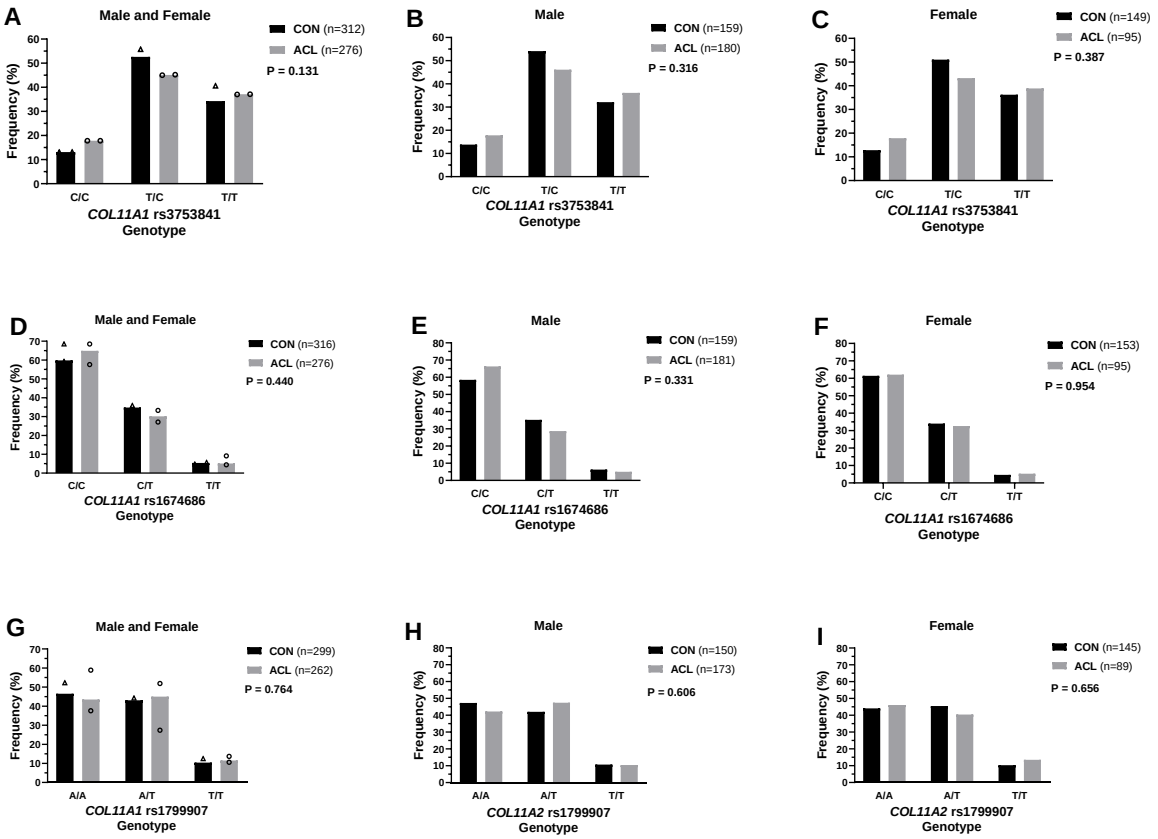


Figure 5.4 The *COL11A1* rs3753841 genotype distributions of the **(A)** male and female, **(B)** male only and **(C)** female only combined CON (white bars) and ACL (grey bars) groups of South African and Swedish participants with self-reported European ancestry. The *COL11A1* rs1676486 genotype distributions of the **(D)** male and female, **(E)** male only and **(F)** female only combined CON (white bars) and ACL (grey bars) groups of European ancestry. The *COL11A2* rs1799907 genotype distributions of the **(G)** male and female, **(H)** male only and **(I)** female only combined CON (white bars) and ACL (grey bars) groups of European ancestry. The genotype distributions of each induvial cohort are shown as symbol in only the combined male and female panel. Sample sizes were too small for most cohorts for individual cohort sex sub-analysis. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

Finally, there were also no significant differences (p = 0.423) in the *COL12A1* rs970547 genotype distribution of a combined analysis with previously reported cohorts of European ancestry consisting of South African, Swedish, and Polish populations populations between the ACL group

(AA 63.7%, GA 31.3%, GG 5.1%) compared to the combined CON group (AA 60.4%, GA 35.3%, GG 4.3%) (Figure **5.5** A). Genotype distributions within the male-only and female-only sub-groups were also not found to be significantly different (Figure 5.5 A, B and C).

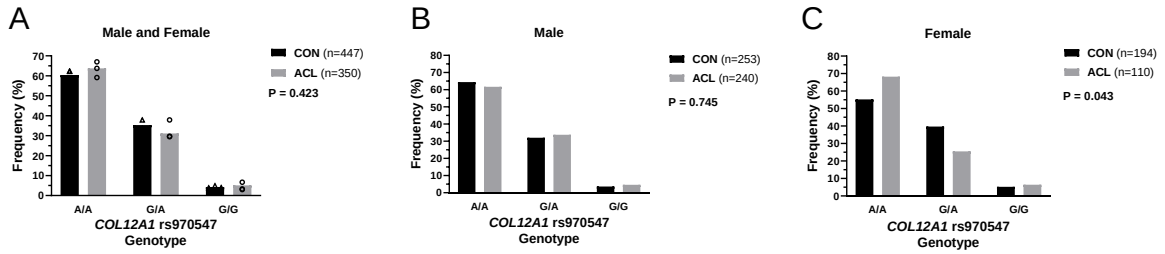


Figure 5.5 *COL12A1* rs970547 genotypes of combined ACL (A), male only (B) and female only (C) for CON (white bar) and combined ACL groups (grey bar) groups of South African, Swedish and Polish participants with self-reported European ancestry. Each induvial cohort is shown as symbol in only the male and female panel. Sample size too small for most cohorts for sex sub-analysis so symbols not shown. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

5.3.5 Inferred Genotype-Genotype Interaction Analysis

Inferred genotype-genotype interactions were constructed for the Swedish cohort, as well as the larger combined Swedish, South African (409) and/or Polish (89) cohorts. Only allele combinations of each haplotype that were inferred at a frequency above 4% are shown in the figures in the subsequent sections. Detailed results of each the constructed haplotypes for each population are shown in the Appendices (Tables B3 – Swedish Cohort, Tables B4 – Combined

ACL Cohort, Tables B5 – Male-only Combined ACL Cohort and, Tables B6 – Female-only Combined ACL Cohort).

5.3.5.1 Type V and XI Collagen Gene-Gene Interactions

None of the inferred haplotypes constructed from *COL5A1* rs12722 (T/C) and *COL5A1* rs10628678 (AGGG/-) were found to be associated with risk of ACL rupture in the Swedish cohort (Figure 5.6 A, Table B3.1). However, a combined group comprising of Swedish, South African, Polish and Australian groups were analysed (Figure 5.6 B, Table B4.1), it was found that the C-AGGG (31.2% vs. 20.6%, p=0.001, ACL vs. CON, respectively) and T-(-) (14.4% vs. 5.7%, p=0.010, ACL vs. CON, respectively) inferred haplotypes were significantly over-represented in the ACL group. Whereas the T-AGGG inferred haplotype was found to be significantly over-represented in the CON group (50.5% vs. 36.6%, p>0.001, CON vs. ACL, respectively). Similarly, the Female only analysis (Figure 5.6 D, Table B6.1) found the C-AGGG (33.8% vs. 16.3%, p=0.004, ACL vs. CON, respectively) and T-(-) (21.4% vs. 4.3%, p=0.006, ACL vs. CON, respectively) inferred haplotypes were significantly over-represented in the ACL group. Whereas the T-AGGG inferred haplotype was found to be significantly over-represented in the CON group (54.0% vs. 30.0%, p>0.001, CON vs. ACL, respectively). However, in the Male only analysis (Figure 5.6 C Table B5.1), none of the inferred haplotypes were found to be significant when the Swedish cohort was combined the South African, Australian and Polish cohorts.

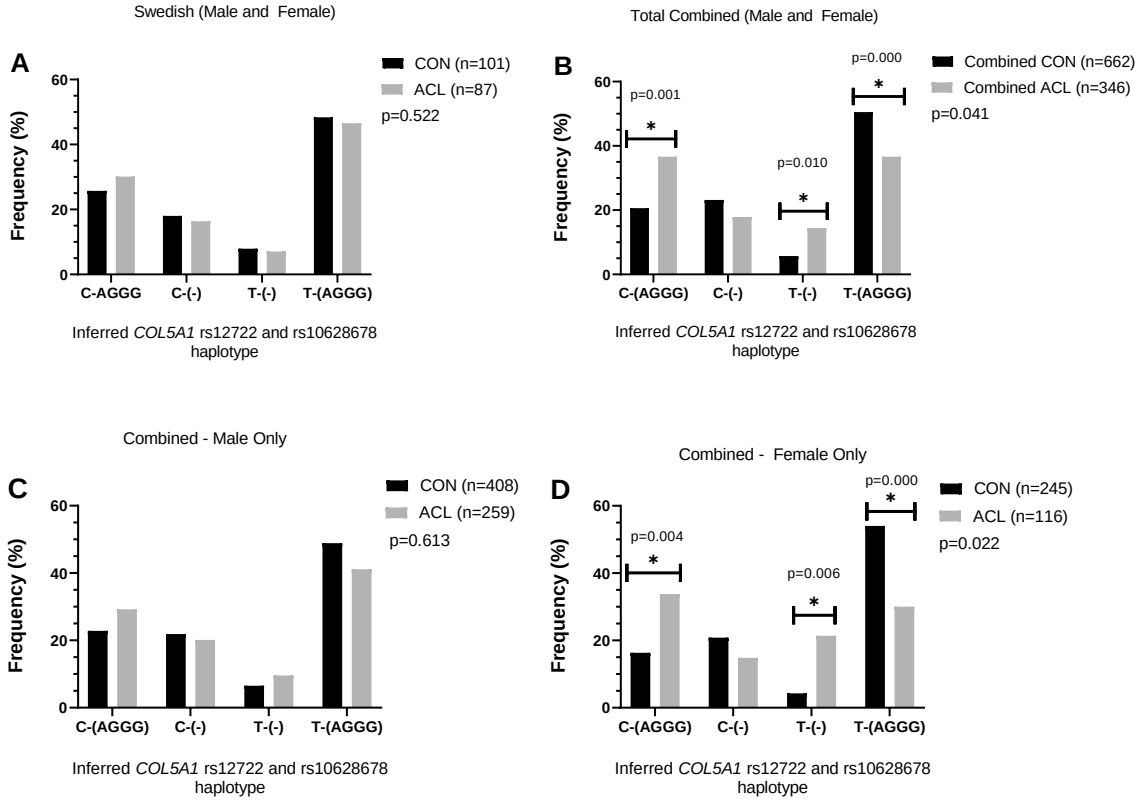


Figure 5.6. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) and *COL5A1* rs10628678 (AGGG/-) of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D) for CON (black bar) and combined ACL groups (grey bar). The combined groups consisted of South African, Swedish, and Australian participants with self-reported European ancestry. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

Pseudo-haplotypes were constructed between *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (A/T). None of the possible inferred haplotypes were found to be associated with risk of ACL rupture in the Swedish (Figure 5.7 A, Table B3.2) or the combined cohorts (Figure 5.7 B to D, Table 4.2, Table B5.2, Table B6.2).

Similarly, none of the possible inferred haplotypes constructed from any two of the type XI collagen gene polymorphisms (*COL11A1* rs3753841, *COL11A1* rs1676486 and *COL11A2* rs1799907) were found to be associated with risk of ACL rupture in the Swedish and combined populations, or when the combined populations were stratified by sex (Figure 5.7 A to D, Table B4.9).

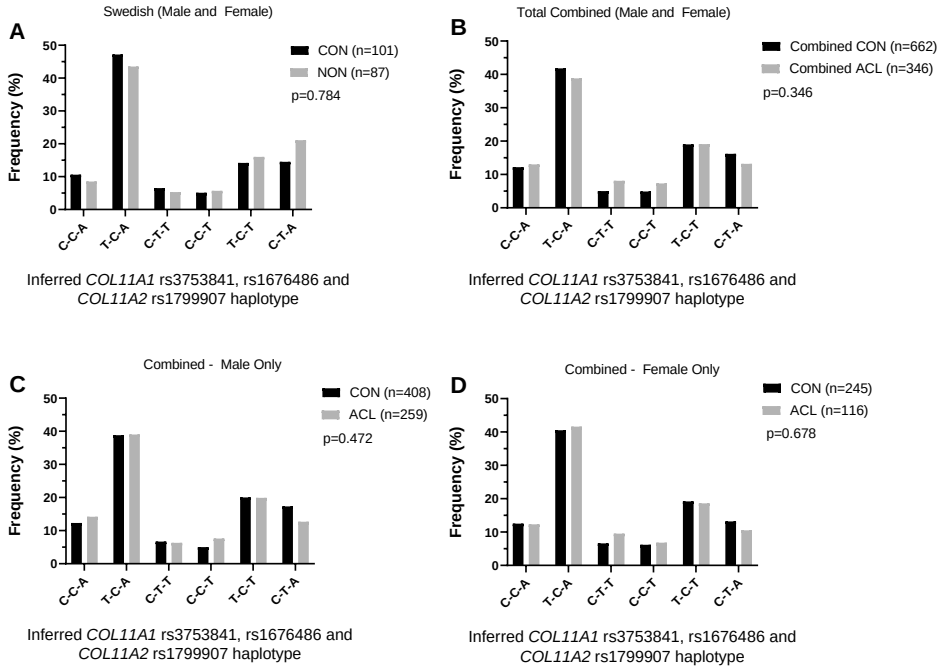


Figure 5.7. Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and rs1676486 (C/T), *COL11A2* rs1799907 (A/T) polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D) for CON (black bar) and combined ACL groups (grey bar). The combined groups consisted of South African and Swedish participants with self-reported European ancestry. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

Inferred haplotypes constructed from the *COL11A1* rs3753841 (T/C) and rs1676486 (C/T) polymorphisms (Figure B4.2 A to B4.2 D). None of the possible inferred haplotypes were found to be associated with risk of ACL rupture in the Swedish (Figure B4.2 A, Table B3.3) combined analysis or when stratified by sex (B4.2 B to D, Table B4.3, Table B5.3, Table B6.3).

Inferred haplotypes constructed from the *COL11A1* rs3753841 (T/C) and *COL11A2* rs1799907 (A/T) polymorphisms (Figure B4.3). There were no significant differences between the ACL group and CON group in the Swedish (Figure B4.3 A, Table B3.4) combined analysis or when stratified by sex (Figure B4.3 B to D, Table B4.4, Table B5.4, Table B6.4).

Inferred haplotypes constructed from the *COL11A1* rs1676484 (T/C) and *COL11A2* rs1799907 (A/T) polymorphisms (Figure B4.4). There were no significant differences between the ACL group and CON group in the Swedish (Figure B4.4 A, Table B3.5) combined analysis or when stratified by sex (Figure B4.4 B to D, Table B4.5, Table B5.5, Table B6.5).

Inferred haplotypes were constructed from, *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T), *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-). None of the possible allele combinations were found to be significantly associated with risk ACL rupture risk (Figure 5.8) in Swedish (Figure 5.8 A, Table B3.6), the combined ACL (Figure 5.8 B, Table B4.6), the male (Figure 5.8 C, Table B5.6) or female (Figure 5.8 D, Table B6.6) cohorts.

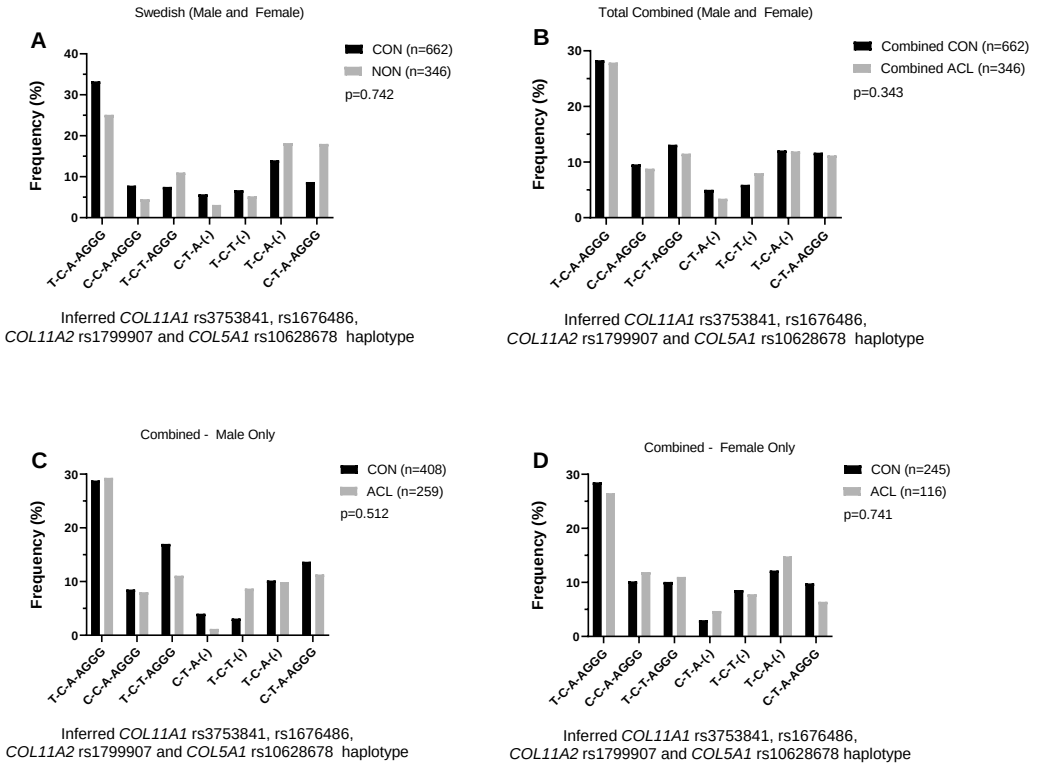


Figure 5.8. Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The combined groups consisted of South African and Swedish participants with self-reported European ancestry. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

Similarly, none of the possible inferred haplotypes constructed from any two of the type XI collagen gene polymorphisms (*COL11A1* rs3753841, *COL11A1* rs1676486 and *COL11A2* rs1799907) and COL5A1 rs10628678, or any individual type XI collagen gene polymorphisms with the *COL5A1* polymorphism were found to be associated with risk of ACL rupture in the Swedish and combined populations, or when the combined populations were stratified by sex (Figures B4.5 to B4.10).

Inferred haplotypes were constructed from *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T), and *COL5A1* rs10628678 (AGGG/-). However, none of the possible allele combinations were found to be significantly associated with risk ACL rupture risk (Figure 5.12) in Swedish (Figure B4.5 A, Table B3.7), Combined ACL (Figure B4.5 B, Table B4.7), Male-only (Figure B4.5 C, Table B5.7) or Female-only (Figure B4.5 D, Table B6.7) cohorts.

Inferred haplotypes were constructed from *COL11A1* rs3753841 (T/C), *COL11A2* rs1799907 (A/T), and *COL5A1* rs10628678 (AGGG/-). However, none of the possible allele combinations were found to be significantly associated with risk ACL rupture risk (Figure B4.6) in Swedish (Figure B4.6 A, Table B3.8), Combined ACL (Figure B4.6 B, Table B4.8), Male-only (Figure B4.6 C, Table B5.8) or Female-only (Figure B4.6 D, Table B6.8) cohorts.

Inferred haplotypes were constructed from *COL11A1* rs1676486 (C/T), *COL11A2* rs1799907 (A/T), and *COL5A1* rs10628678 (AGGG/-). However, none of the possible allele combinations were found to be significantly associated with risk ACL rupture risk (Figure B4.7) in Swedish (Figure B4.7 A, Table B3.9), Combined ACL (Figure B4.7 B, Table B4.9), Male-only (Figure B4.7 C, Table B5.9) or Female-only (Figure B4.7 D, Table B6.9) cohorts.

Inferred pseudo-haplotypes were constructed using the *COL5A1* rs10628678 (AGGG/-) polymorphism with each of the three type XI collagen gene polymorphisms (Figure B4.8 to B4.10). None of the possible allele combinations in the *COL11A1* rs3753841 (T/C) and *COL5A1* rs10628678 (AGGG/-) were found to be significantly associated with risk ACL rupture risk (Figure B4.8) in Swedish (Figure B4.8 A, Table B3.10), Combined ACL (Figure B4.8 B, Table B4.10), Male-only (Figure B4.8 C, Table B5.10) or Female-only (Figure B4.8 D, Table B6.10) cohorts.

In the *COL11A1* rs1676486 (C/T) and *COL5A1* rs10628678 (AGGG/-) inferred haplotypes, none of the possible allele combinations were found to be significantly associated with rupture risk (Figure B4.9) in Swedish (Figure B4.9 A, Table B3.11), Combined ACL (Figure B4.9 B, Table B4.11), Male-only (Figure B4.9 C, Table B5.11) or Female-only (Figure B4.9 D, Table B6.11) cohorts.

In the *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) inferred haplotypes, none of the possible allele combinations were found to be significantly associated with rupture risk (Figure B4.10) in Swedish (Figure B4.10 A, Table B3.12), Combined ACL (Figure B4.10 B, Table B4.12), Male-only (Figure B4.10 C, Table B5.12) or Female-only (Figure B4.10 D, Table B6.12) cohorts.

5.3.5.2 Types V, VI and XII Collagen Gene-Gene Interactions

Pseudo-haplotypes were constructed with *COL12A1* rs970547 (G/A) and *COL6A1* rs35796750 (T/C) and *COL5A1* rs12722 (T/C) (Figure 5.8). However, none of the possible allele combinations were found to be significantly associated with rupture risk (Figure 5.9) in Swedish (Figure 5.9 A, Table B3.14), the combined ACL (Figure 5.9 B, Table B4.14), male (Figure 5.9 C, Table B5.14) or female (Figure 5.9 D, Table B6.14) cohorts.

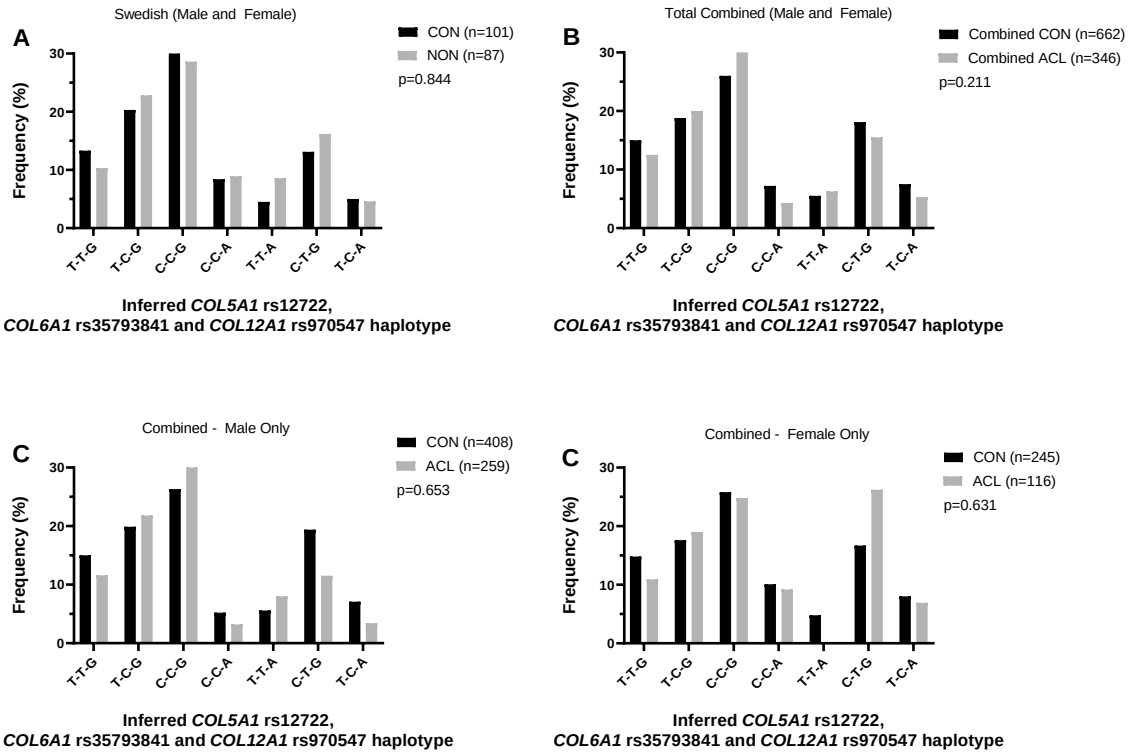


Figure 5.9. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C), *COL6A1* rs35796751 (T/C) and *COL12A1* rs970547 polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The combined groups consisted of South African and Swedish participants with self-reported European ancestry. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

5.3.5.3 Types V and XII Collagen Gene-Gene Interactions

Pseudo-haplotypes were constructed using *COL5A1* rs12722 and *COL12A1* rs970547 (Figure 5.9). However, none of the possible allele combinations were found to be significantly associated with rupture risk in Swedish (Figure 5.10 A, Table B3.15), Combined ACL (Figure 5.10 B, Table

B4.15), Male-only (Figure 5.10 C, Table B5.15) or Female-only (Figure 5.10 D, Table B6.15) cohorts.

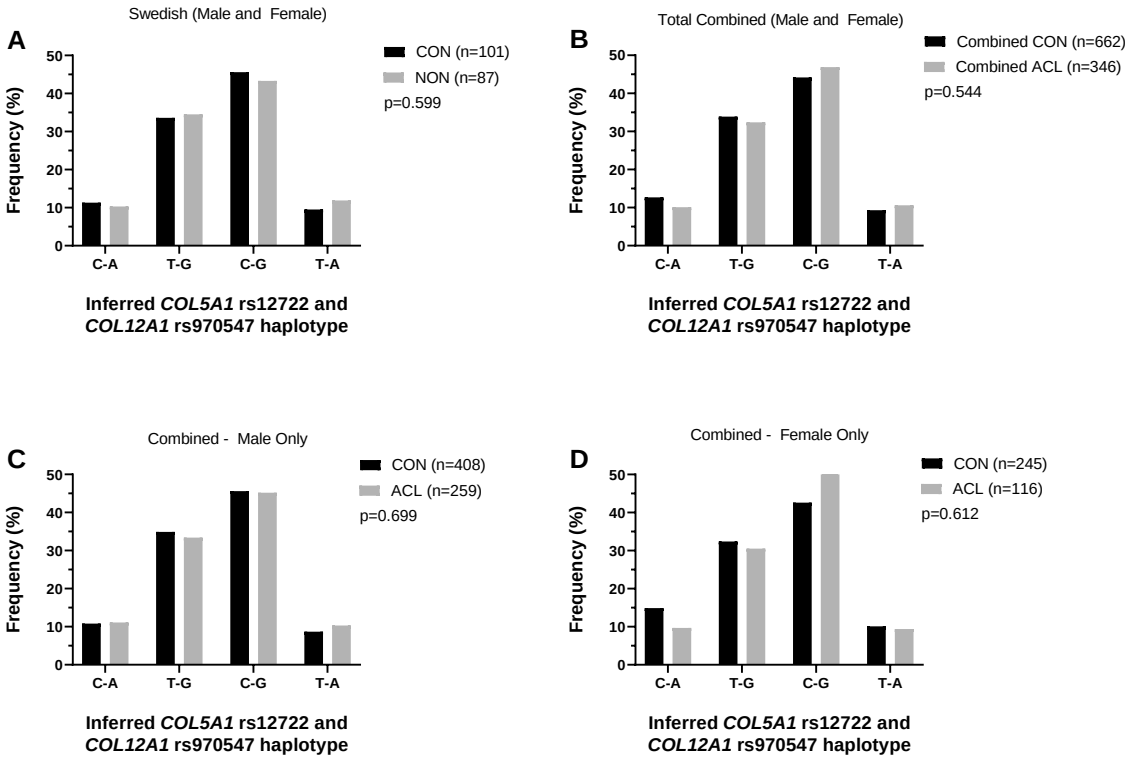


Figure 5.10. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C), and *COL12A1* rs970547 polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The combined groups consisted of South African, Polish and Swedish participants with self-reported European ancestry. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

5.3.5.4 Types V and VI Collagen Gene-Gene Interactions

Pseudo-haplotype was constructed using the *COL5A1* rs12722 (T/C) and *COL6A1* rs35796750 (T/C) (Figure 5.11). However, none of the possible allele combinations were found to be

significantly associated with rupture risk in Swedish (Figure 5.10 A, Table B3.13), Combined ACL (Figure 5.10 B, Table B4.13), Male-only (Figure 5.10 C, Table B5.13) or Female-only (Figure 5.10 D, Table B6.13) cohorts.

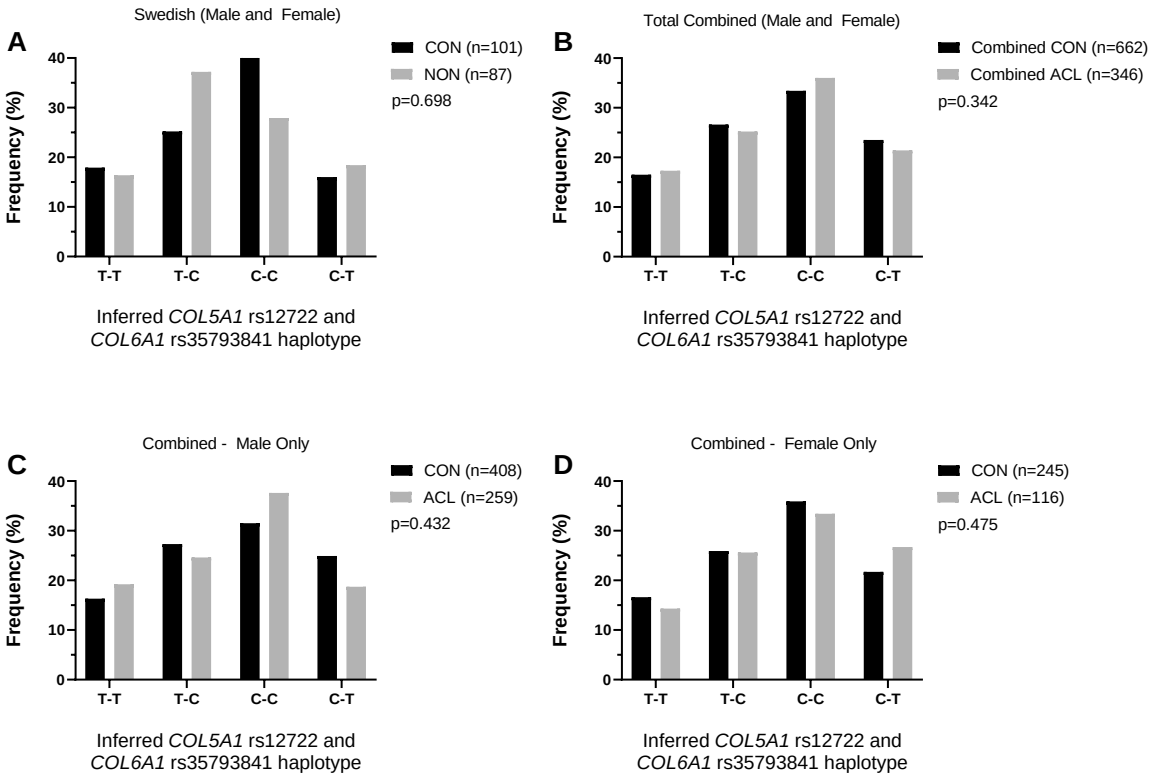


Figure 5.11. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) and *COL6A1* rs35796751 (T/C) polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The combined groups consisted of South African and Swedish participants with self-reported European ancestry. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

5.3.5.5 Types VI and XII Collagen Gene-Gene Interactions

Pseudo-haplotypes were constructed using *COL6A1* rs35796750 and *COL12A1* rs970547 (Figure 5.12). None of the four possible inferred pseudo-haplotypes combinations in *COL6A1* rs35796750 (T/C) and *COL12A1* rs970547 (G/A) were found to be significant in the Swedish population (Figure 5.12, Table B3.16). However, in the Combined ACL group, the T-A haplotype was found to be significantly over-represented in the ACL group compared to the CON group (p=0.030, 11.0% vs. 7.1%, ACL vs CON, respectively). Moreover, in the Male-only combined ACL cohort, the T-G haplotype was found to be significantly over-represented in the CON group compared to the ACL group (p=0.010, 33.7% vs. 23.0%, CON vs. ACL, respectively). Finally, in the Female-only combined ACL cohort, none of the possible allele combinations were found to be significant.

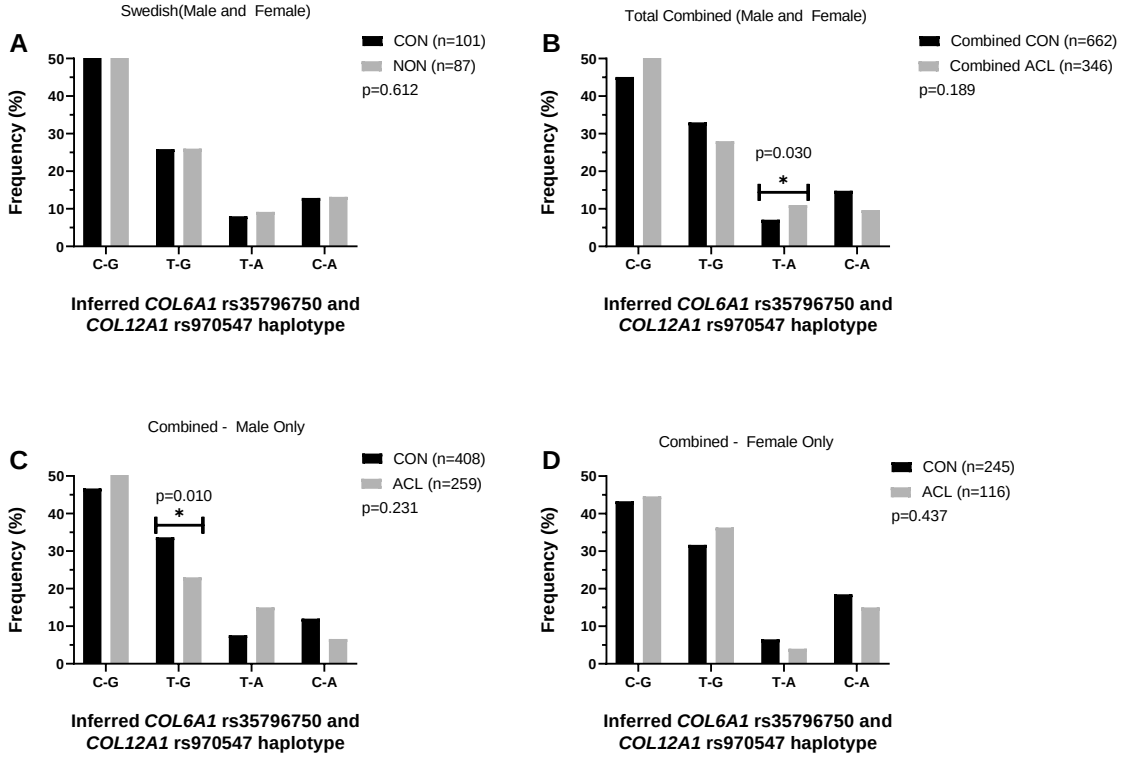


Figure 5.12. Inferred haplotype frequency distributions constructed from *COL6A1* rs35796751 (T/C) and *COL12A1* rs970547 polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The combined groups consisted of South African and Swedish participants with self-reported European ancestry. The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

5.4 DISCUSSION

The first main finding of this study was that there were no individual associations between *COL1A1* rs1800012, *COL5A1* rs10628678, *COL6A1* rs35796750, *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 and *COL12A1* rs970547 and ACL rupture risk in a Swedish population. *COL5A1* rs12722 was previously also reported not to be associated with ACL

rupture in the same Swedish cohort (507). Similarly, none of the investigated inferred haplotypes constructed from the biologically relevant collagen gene variants were associated with ACL rupture risk in the Swedish cohort. Since there were no independent associations identified within the Swedish cohort, a larger combined analysis was conducted using self-reported European ancestry ACL rupture cohorts consisting of previously published South African (409), Polish (89), Swedish (507), Norwegian and Finish, (503) and/or Australian populations (502).

The second main finding the combined cohort analysis was that the *COL1A1* rs1800012 TT genotype was associated with reduced risk of ACL rupture within a combined cohort consisting of South African, Polish, Swedish, Norwegian, Finish and Australian participants with self-reported European ancestry. A novel finding of this thesis was that this *COL1A1* TT genotype was associated with reduced risk in the female, but not male participants. Although the exact mechanism is unknown, several other intrinsic risk factors have been proposed to alter the risk of ACL ruptures in females (38,231,509). One possible hypothesis is that sex-specific gene-hormone interactions may modulate risk of ACL rupture in females. Future work is required to repeat this sex-specific association and determine the molecular mechanism(s). Although previously reported in some smaller cohorts, none of the other investigated collagen variants, namely *COL5A1* (rs12722 and rs10628678), *COL6A1* (rs35796750), *COL11A1* (rs3753841 and rs1676486), *COL11A2* (rs1799907) and *COL12A1* (rs970547) were individual associated with ACL rupture risk in the larger combined cohort.

Previous literature has demonstrated that there is some degree of heritability for ACL injuries (237,510,511). A longitudinal study of approximately 88,000 Swedish twins found that the genetic contribution of ACL injuries was approximately 70% (512). In the current Swedish cohort,

participants in the NON group reported have significantly more previous ligament injuries than the CON group. Although a family history of ACL rupture was not significantly different between groups, we found that participants in the NON group reported significantly more family history of joint injuries compared to the CON group. Participants in the NON group were twice more likely to have sustained an ACL rupture if a parent reported a history of previous joint rupture. In addition, risk of ACL rupture increased to 4-times if a sibling had reported a previous joint rupture. These findings are consistent with previous studies investigating familial risk of ACL injuries. In a case control study by Myer et al. (510) of patients undergoing ACL reconstruction, men with ACL injuries were more likely to have a sibling with an ACL compared to rupture free men. No associations in risk were found between women with ACL injuries and healthy controls. However, a prospective study found that women with had suffered and ACL rupture, were more likely to have a parent with a history of ACL than healthy controls (223).

As previously discussed in Chapter 2, section 2.7.1, type I collage is the predominant fibrillar collagen in non-cartilaginous connective tissues. Several studies have investigated the role of the *COL1A1* gene as a potential predisposing factor of acute and chronic musculoskeletal soft tissue injuries (500). More specifically, the TT genotype of rs1800012 polymorphism has been shown to be associated with reduced risk of acute injuries including shoulder dislocations and ACL injuries (27,84,500) in most of the investigated European Ancestry populations. Recently, Gibbon et al. (500), in a large multivariate modelling approach, demonstrated similar results a significant under-representation of the TT genotype in ACL injuries. The T allele of the rs1800012 polymorphism has been shown to enhance the binding affinity of the Sp1 transcription factor, producing a three-fold increases in the primary level of the messenger RNA (mRNA) in contrast to the G allele (374,377,513). Further, Ficek et al (380) also noted a decreased risk of both ACL and Achilles

tendon ruptures when inferred haplotypes were constructed using the *COL1A1* rs1107946 and rs1800012 variants. The reduction of acute soft tissues injuries, in theory, may be due to an increase in homotrimeric type-I collagen molecules resulting in increased tensile strength and a greater diffusion of force in tendons and ligaments. Further work should focus on the rs1107946, rs1800012 and other *COL1A1* functional variant interactions in the modulation of ACL ruptures.

Although the *COL5A1* rs12722 and rs10628678 genotypes were not independently associated with ACL rupture risk in the combined analysis comprising the Swedish, South African, Polish and Australian cohorts, an additional finding of this study was that the C-AGGG and T-(-) inferred haplotypes constructed from the these *COL5A1* variants were associated with increased risk of ACL rupture, while the T-AGGG haplotype was associated with reduced risk, in the combined male and female group and female-only group. Recently, Alvarez-Romero et al (502) reported also reported that the T-(-) inferred haplotype was associated with increased risk when only the same South African and Australian cohorts were analysed together. The inclusion of the additional Polish and Swedish cohorts in the analysis did not alter this finding. However, the previous finding analysing the smaller combined cohort reported that the opposite C-(-) haplotype, rather than the T-AGGG haplotype reported in this study, was associated with reduced risk (502). The C-AGGG haplotype was not shown to modulate risk in the previous study. There are limited studies that have investigated the association of the *COL5A1* rs10628678 variant with MSK injuries. The AGGG/AGGG genotype has been associated with increased risk of TEN in South African and Australian cohorts (76,490). However, no association was observed in a British TEN (80).

Haplotypes constructed from several variants located within the *COL5A1* 3'-UTR were nevertheless shown to modulate risk of TEN in the British cohort (80). Taken together, the simplest

explanation of these findings suggests that the specific region within the *COL5A1* 3'-UTR, still requires to be defined. Future work should consider sequencing the entire 3'-UTR in larger cohorts and constructed inferred haplotypes from all the identified variants to determine whether this functional region modulates the risk of ACL rupture (386,389).

None of the investigated type XI collagen gene variants were independently associated with ACL rupture in the combined cohort consisting of the South African and Swedish participants with self-reported European ancestry. Similarly, none of the inferred haplotypes constructed from the type XI and/or type V collagen gene polymorphisms were found to be associated with risk modulation. Previously, Hay et al(76) examined *COL11A1* (rs3753841 and rs1676486), *COL11A2* (rs1799907) and *COL5A1* (rs10628678) within an Australian and South African TEN cohort. The T-C-T inferred haplotypes constructed from *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 was found to be significantly associated with increased tendinopathy risk. In addition, the T-C-T-AGGG inferred haplotype constructed from *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 and *COL5A1* rs10628678, was found to be significantly over-represented in the tendinopathy group. More recently, Dada et al (98) found that the TT genotype of the *COL11A1* rs3753841 (T/C) was significantly associated with increased risk of CTS. Additionally, the T-C haplotype constructed from *COL11A1* rs3753841 and *COL11A1* rs167648 was associated with increased risk (98). Type XI collagen is a heterotrimer that is encoded by *COL11A1*, *COL11A2* and *COL11A3* genes, which produces of α1(XI), α2(XI) and α3(XI) chains, respectively. Since it has a similar structure and function to type V collagen, it has been theorised to interact with the type V collagen during tendon development, regulating fibrillogenesis (252). Type XI collagen gene variants need to be investigated in additional cohorts, these findings suggest

that, unlike chronic tendinopathy, the investigated variants within the type V and XI collagen genes do no interact to alter the risk of ACL rupture.

Finally, although the *COL6A1* rs35796750 and the *COL12A1* rs970547 variants were not independently associated with ACL ruptures, the T-A inferred haplotype was found to be significantly associated with increased ACL rupture risk. Interestingly, the T-G inferred haplotype was significantly associated with reduced risk when only the male participants were analysed. None of the other investigated interactions between the type V, VI and/or XII collagen gene variants were associated with ACL rupture in this study. Both collagen type VI and XII are non-fibrillar collagens and are expressed in connective tissues such as bone and skeletal muscle (393,394,514,515). The function of type VI collagen is still being investigated; however, it is suggested to be involved with the development and maintenance of the basement membrane (394). It has also been observed to interact with types I, V and XII collagens to assist with collagen fibrillogenesis (198) in which *COL5A1* and *COL12A1* also plays a role as well. Both collagens have the ability to mediate cell-matrix and matrix-matrix interactions (515), which are important features regulating cell migration, adhesion, apoptosis, and are crucial for matrix bridge formation. Izu et al. (515) demonstrated that deficiencies in collagens VI or XII impair matrix bridge formation particularly in bone forming sites. The authors suggested that matrix bridge formation requires the presence and interaction of both type VI and XII collagens (515).

As outlined in Chapter 2, section 2.2.7, several studies have investigated the role of collagen gene polymorphisms and ACL injury or rupture (31,73,84,87–89,503). Several genes have been previously implicated as potential risk factors including those investigated in this thesis. These previous studies have however generally investigated associations in individual populations with

European ancestry, which are limited by small sample sizes and and may not have sufficient statistical power. In many cases any initially reported independent association has not been repeated in follow-up studies. In addition to investigating the independent association of collagen gene variants, as well as gene-gene interactions, within a Swedish cohort, the association of these variants with a larger combined cohort was therefore investigated in this chapter to address the sample size and subsequent power limitations.

ACL rupture is a multifaceted injury phenotype with several intrinsic and extrinsic factors altering the risk of rupture (219,228,233,412). Even though there were no independent associations of risk in the Swedish cohort, a combined analysis did show that the collagen gene variants investigated in this thesis, at least in part, contributed to the genetic susceptibility of ACL ruptures. In conclusion, the results of this study in addition to previous findings collectively strengthen the role of collagen gene variants and their modulation of ACL rupture risk. Further investigations into these genes will add to the ever-growing body of literature and towards the development of a risk profile for ACL rupture. As additional variants are investigated, the picture surround the aetiology of injury risk will be clearer and more complete, giving physicians and health science practitioners better ability to predict and manage future injuries before they occur. A continuous summary of the results of this thesis are presented in Table 5.5 and Table 5.6.

Table 5.5: A continuing summary of the genotype results from each chapter in this thesis. The results of Chapter 5, and preceding the preceding chapter, are shown.

Investigated Gene Variant	Investigated Phenotype			
	Chapter 4		Chapter 5	Chapter 6
	RCT	SST	EA ACL	MA ACL
COL1A1 rs1800012	n.s	n.s / Prev pub CC (30)	TT[a] (thesis)	
		n.s	Prev pub TT (84)	
			Prev pub GG (90)	
COL5A1 rs12722	n.s	n.s	n.s	
			Prev pub CC (73)	
COL5A1 rs10628678	n.s	n.s	n.s	
COL6A1 rs35796540	n.s	n.s	n.s	
			n.s Prev pub (89)	
COL11A1 rs3754841	n.s	n.s	n.s	
COL11A1 rs16746744	n.s	n.s	n.s	
COL11A2 rs1799907	n.s	n.s	n.s	
COL12A1 rs970547	n.s	n.s	n.s	
			Prev pub AA (87,405)	

Green shading indicates a reduction in injury risk. Red shading indicates an increased risk of injury. Grey shading indicates no association with risk. Prev pub; Previously published. RCT; rotator cuff tendinopathy. SST; Supraspinatus tendinopathy. EA; European Ancestry. ACL; Anterior cruciate ligament rupture. MA; Mixed Ancestry.

[a] significant association in combined EA cohort but not in Swedish only cohort.

Table 5.6: A continuing summary of the haplotype results from each chapter in this thesis. The results of Chapter 5, and preceding chapter, are shown.

Phenotype		Investigated Gene variants						
		COL11A1 rs3753841 (T/C)	*COL11A1* rs1676486 (C/T)	*COL11A2* rs1799907 (A/T)	*COL5A1* rs12722 (T/C)	*COL5A1* rs10628678 (AGGG/-)	*COL6A1* rs35796750 (C/T)	*COL12A1* rs970547 (A/G)
Chapter 4	RCT SST	C		A	(-) (thesis)			
Chapter 5	EA ACL				C	AGGG (thesis)		
					T	(-) (thesis)		
					T	AGGG (thesis)		
					Prev pub T	(-)(502)		
		Prev pub C	T	A (409)				
		Prev pub T	C	T	AGGG (409)			
		Prev pub T	C	T	(-)(409)			
		Prev pub C	T	A	AGGG (409)			
		Prev pub T	C	A	AGGG (409)			
							T	G (thesis)
							T	A (thesis)
Chapter 6	MA ACL							

Green shading indicates a reduction in injury risk. Red shading indicates an increased risk of injury. Grey shading indicates no association with risk. Prev pub; Previously published. RCT; rotator cuff tendinopathy. SST; Supraspinatus tendinopathy. EA; European Ancestry. ACL; Anterior cruciate ligament rupture. MA; Mixed Ancestry.

CHAPTER 6: GENETIC RISK FACTORS FOR ACL RUPTURE IN A MIXED ANCESTRY COHORT

6.1 INTRODUCTION

As discussed in the previous chapter, the literature has primarily focused on investigating the association of the collagen variants included in this thesis with ACL rupture in populations of European ancestry. A limited number of studies have investigated their association in other population groups (31,405,499,516,517). Specifically, the association of *COL1A1*, *COL5A1* and *COL12A1* variants with ACL rupture have previously been investigated in Indian cohorts (657,659,660). Prabhakar et al. (499) did not find a significant association between *COL1A1* rs1800012 and rs1107946 and the risk of ACL rupture. Similarly, in a study by Shukla et al. (516), *COL1A1* rs1800012 was not associated with ACL ruptures or non-contact ACL rupture in a cohort of Indian athletes.

This is not surprising since the T allele is more common in European populations (19%) when compared to South Asian (10%), East Asian (0%) and African (6%) populations (www.ensembl.org) (500) and therefore may not informative in non-European populations. However, the AG and GG genotypes of the *COL12A1* rs970547 variant was significantly associated with ACL risk in an Indian population (31). A more recent study of Korean ACL rupture patients found that the opposite AA genotype was significantly over-represented in the ACL group (517). Similarly, Zhao et al. (405) also reported a significant association of the A allele and AA genotype of the *COL12A1* rs970547 with an increased risk of ACL rupture in the male sub-group of a Chinese Han population. Zhao et al. (405) further reported that *COL1A1* and *COL5A1*varaints were not associated with an ACL rupture in this cohort.

Previous investigations within a South African mixed ancestry population focused on the role of angiogenesis genes and ACL rupture risk (518). The study by Rahim et al. (518) found that the AG genotype of the *KDR* gene was significantly associated with reduced risk, whilst GG genotype was significantly associated with increased risk of ACL rupture. Moreover, the authors observed that within this population, first-degree relatives in the ACL group were reported to have significantly more ACL injuries than the CON group (518).

The limited available literature highlights the importance of investigating genetic associations in genetically diverse populations and the ethnicities, which allows for a deeper understanding of the potential biological mechanisms that may influence the susceptibility to MSK injuries (518,519). Therefore, the aims of this study within the thesis were to i) determine if *COL1A1* rs1800012 G/T, *COL5A1* rs12722 T/C, *COL5A1* rs10628678 AGGG/-, *COL11A1* rs3753841 T/C, *COL11A1* rs1676486 C/T, *COL11A2* rs1799907 A/T and *COL12A1* rs970547 G/A are independently associated with risk of ACL rupture in a South African mixed ancestry cohort, and ii) to investigate whether any of the gene-gene interactions between the collagen variants investigated in the previous chapter modulated risk of ACL rupture.

6.2 METHODS

This chapter followed the methodology as previously outlined in Chapter 4, section 4.2.1. This study in this chapter was approved by the Faculty of Health Sciences Human Research Ethics Committee within the University of Cape Town, South Africa (HREC 164/2006) (Appendix A1.5).

6.2.1 Participants

A total of 209 participants were included in this study. The population has been previously described by Rahim et al. (518), who investigate the role of angiogenesis genes and ACL rupture risk. Participants were recruited between January 2012 and May 2016 from Groote Schuur Hospital, Victoria Hospital, and the Sports Science Orthopaedic Clinic within the Cape Town area of South Africa. Participants were included in the study based on the previously described inclusion and exclusion criteria (518), and all were requested to give informed written consent following written and verbal explanation of the study (Appendices A1.6 and A1.7). Furthermore, each participant was required to complete questionnaires related to personal details, injury details, and medical and sporting history (Appendix A1.8).

Ninety-four participants (77 males and 17 females) were included in the ACL group, of which 51 had a non-contact mechanism of injury (NON sub-group). NON participants were also analysed separately as a sub-group. ACL injuries were diagnosed as previously described in Rahim et al. (518). Furthermore, 100 (81 males and 19 females) apparently healthy, self-reported mixed ancestry individuals with no history of ACL rupture or injury were recruited from gyms and local sports clubs within the Cape Town area of South Africa.

The participants within this cohort are from an ethnic group, specifically from the Western Cape province of South Africa. The group's ancestry is derived from a combination of immigrants and settlers from Europe, slave labourers originating from Malaysia, India, Java, Madagascar, Indonesia, West Africa, and the indigenous African Khoe- and San- speaking or Bantu-speaking populations (520). Therefore, the ethnic classification "mixed ancestry" is an inclusive term that

incorporates a broad array of people who are ethnically and culturally unique to the South African region (520). Populations within South Africa who self-identify as mixed ancestry have diverse genetic profile that resulted from the ethnic intermixing between immigrating and settler populations since 1652.

6.2.2 Blood Collection, DNA Extraction, and collagen gene polymorphism selection

Blood collection and extraction was done, as described in Chapter 4. All participants were genotyped for the following collagen gene polymorphisms: *COL1A1* rs1800012 (G/T), *COL5A1* rs12722 (T/C) and rs10628678 (AGGG/-), *COL6A1* rs35796750 (T/C), *COL11A1* rs3753841 (T/C) and rs1676486 (C/T), *COL11A2* rs1799907 (A/T) and *COL12A1* rs970547 (G/A) polymorphisms as described in Chapter 4.

6.2.3 Statistics

The same statistical tests and haplotype analysis methods were performed as described in Chapter 4. Significance was accepted when $p<0.05$. No adjustments were made for multiple testing in this study for the same reasons as outlined in Chapter 4. Stratification by sex was not performed as the reduction in sample size may render the results unreliable due to small sample sizes.

6.3 RESULTS

6.3.1 Participant Characteristics

The participants were matched for sex, height, weight, body mass index (BMI), years of sport participation, non-contact sport, and non-contact jumping (Table 6.1). However, the participants

in the CON group were significantly older than the ACL (p=0.007) and NON (p=0.019) groups. As reported by Rahim et al. (518), the average self-reported age and weight at the time of recruitment for the ACL group was 2.5 ± 4.4 years (n = 94) older and 0.9 ± 9.6kg (n = 87) heavier than at the time of their rupture. NON subgroup was similarly older (1.9 ± 3.2 years, n = 51) and heavier (1.1 ± 6.9kg, n = 49) at the time of recruitment.

Table 6.1: General participant characteristics and sport participation for the controls (CON), ACL Rupture (ACL) group and non-contact ACL rupture (NON) sub-group.*

	CON (n=100)	**ACL** (n=98)	p value[a]	**NON** (n=51)	p value[b]
Age (years) Ψ	27.4 ± 6.9 (99)	24.5 ± 7.5 (91)	**0.007**	24.5 ± 7.7 (51)	**0.019**
Female Sex (%) (n)	19.0 (100)	17.3 (98)	0.764	12.2 (51)	0.259
Height (cm)	172.2 ± 8.3 (95)	173.7 ± 8.6 (86)	0.221	174.3 ± 8.7 (48)	0.162
Weight (kg) Ψ	77.2 ± 15.6 (94)	77.7 ±15.7 (88)	0.825	74.2 ± 12.9 (50)	0.240
BMI ($kg.m^{-2}$) Ψ	25.3± 5.6 (95)	25.1 ± 6.1 (84)	0.792	24.3 ± 3.9 (47)	0.278
Sport Participation (yrs)	14.8 ± 9.7 (78)	15.3 ± 8.8 (74)	0.730	14.8 ± 8.9 (40)	0.989
Non-Contact Sport (yrs)	13.3 ± 13.3 (76)	13.6 ± 11.2 (49)	0.422	14.2 ± 8.9 (31)	0.140
Non-Contact Jump (yrs)	12.3 ± 9.5 (16)	10.2 ± 8.4 (15)	0.514	10.9 ± 9.5 (10)	0.523

* results previously reported by Rahim et al. (662).
Expect for sex, which is expressed as a frequency (%), values are expressed as mean ± standard deviation. The number of participants (n) is indicated in parentheses.
BMI: Body Mass Index, cm: centimetres, kg: kilograms, m: meters, hrs. wk-1: hours per week
Ψ Age, weight and BMI are self-reported values at the time of first ACL rupture for the ACL group and NON sub-group. They are the self-reported values at recruitment for CON group.
[a] *CON vs. ACL.*, [b] *CON vs. NON.*

6.3.2 Participant and Family History of Soft Tissue Injuries

Participant History of Soft Tissue Injuries and Sport Participation History were previously reported by Rahim et al. (518).

6.3.3 Genotyping

There was no significant difference in the relative genotype or allele distributions of the collagen gene polymorphisms investigated in this thesis between CON and ACL groups (Table 6.2). When stratified by the mechanism of injury, only the *COL12A1* rs970547 GG genotype was significantly over-represented in the NON sub-group (20.9% vs. 5.4%; p=0.021; NON vs. CON, respectively) when compared to the CON group. No other significant differences were observed between the remaining seven polymorphisms. The *COL1A1*, *COL5A1*, *COL6A1*, *COL11A1*, *COL11A2*, *COL11A2* and *COL12A1* polymorphisms were in HWE (Table 6.2).

Table 6.2. Genotype frequency distributions for *COL1A1* rs1800012, *COL5A1* rs12722, *COL5A1* rs10628678, *COL6A1* rs35796750, *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 and *COL12A1* rs970547 in the control (CON), anterior cruciate ligament rupture (ACL) and non-contact anterior cruciate ligament injuries (NON).

SNP	Allele	CON	ACL	*P*-value[a]	NON	*P*-value[b]
	N	96	96		50	
	G/G	82.3 (79)	85.4 (82)	0.552	90.0 (45)	0.204
***COL1A1* rs1800012**	G/T	17.7 (17)	13.5 (13)		10.0 (5)	
(G/T)	T/T	0.0 (0)	1.0 (1)		0.0 (0)	
	T Minor Allele	8.9 (17)	7.3 (14)	0.566	5.0 (5)	0.399
	HWE	1.000	0.447		1.000	
	N	98	97		50	
	C/C	29.9 (29)	42.3 (41)	0.131	34.0 (17)	0.623
***COL5A1* rs12722**	T/C	58.2 (57)	44.3 (43)		50.0 (25)	
(T/C)	T/T	12.2 (12)	13.4 (13)		16.0 (8)	
	T Minor Allele	41.3 (81)	35.6 (69)	0.248	41.0 (41)	0.972
	HWE	0.561	0.579		0.554	
	N	100	97		50	
	AGGG/AGGG	38.0 (38)	35.1 (34)	0.759	30.0 (15)	0.365
***COL5A1* rs10628678**	AGGG/-	50.0 (50)	49.5 (48)		62.0 (31)	
(AGGG/-)	-/-	12.0 (12)	15.5 (15)		8.0 (4)	
	(-) Minor Allele	37.0 (74)	40.2 (78)	0.824	39.0 (39)	0.812
	HWE	1.00	0.679		1.000	
	N	97	92		47	
	T/T	24.2 (24)	30.4 (28)	0.506	31.9 (15)	0.621
***COL6A1* rs35796750**	C/T	50.5 (50)	50.0 (46)		44.7 (21)	
(T/C)	C/C	25.3 (25)	19.6 (18)		23.4 (11)	
	C Minor Allele	51.5 (100)	44.6 (82)	0.180	45.7 (43)	0.515
	HWE	0.806	1.000		1.000	
	N	97	92		45	
	T/T	20.6 (20)	34.8 (32)	0.075	33.3 (15)	0.218
***COL11A1* rs3753841**	T/C	55.7 (54)	42.4 (39)		42.2 (19)	
(T/C)	C/C	23.7 (23)	22.8 (21)		24.4 (11)	
	C Minor Allele	50.5 (100)	42.5 (81)	0.123	45.6 (41)	0.588
	HWE	0.765	0.731		1.000	
	N	99	97		50	
	T/T	55.6 (55)	55.7 (54)	0.951	54.0 (27)	0.940
***COL11A1* rs1676486**	C/T	35.4 (35)	34.0 (33)		38.0 (19)	
(C/T)	C/C	9.1 (9)	10.3 (10)		8.0 (4)	
	C Minor Allele	26.8 (53)	27.3 (53)	0.937	27.0 (27)	0.979
	HWE	0.520	0.832		0.150	
	N	99	97		50	
	A/A	44.4 (44)	45.4 (44)	0.675	54.0 (27)	0.473
***COL11A2* rs1799907**	A/T	48.5 (48)	44.3 (43)		38.0 (19)	
(A/T)	T/T	7.1 (7)	10.3 (10)		8.0 (4)	
	T Minor Allele	31.3 (62)	32.5 (63)	0.857	27.0 (27)	0.589
	HWE	0.599	1.000		0.706	
	N	93	83		43	
	A/A	34.4 (32)	37.3 (31)	0.078	34.9 (15)	**0.021**
COL12A1	G/A	60.2 (56)	48.2 (40)		44.2 (19)	
rs970547 (A/G)	G/G	5.4 (5)	14.5 (12)		20.9 (9)	
	G Minor Allele	35.5 (66)	38.6 (64)	0.672	43.0 (37)	0.237
	HWE	0.348	0.445		0.730	

Genotype and allele frequencies are expressed as percentages with the number of participants shown in parenthesis (n). Global *P*-values are given with bold typeset indicating significant differences ($P<0.05$).
HWE, Hardy-Weinberg equilibrium.
AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 Polymorphism.
[a] CON vs. ACL, [b]CON vs. NON

6.3.4 Collagen Gene-Gene Interactions

6.3.4.1 Type V and XI Collagen Gene-Gene Interactions

All four possible haplotypes constructed from the *COL5A1* rs12722 (T/C) and *COL5A1* rs10628678 (AGGG/-) polymorphisms were inferred at a frequency above 4% (Table B2.1). In addition, there were no significant differences in the inferred *COL5A1* 3'-UTR haplotype distributions between the CON and ACL groups or the CON group and NON sub-group (Figure 6.1).

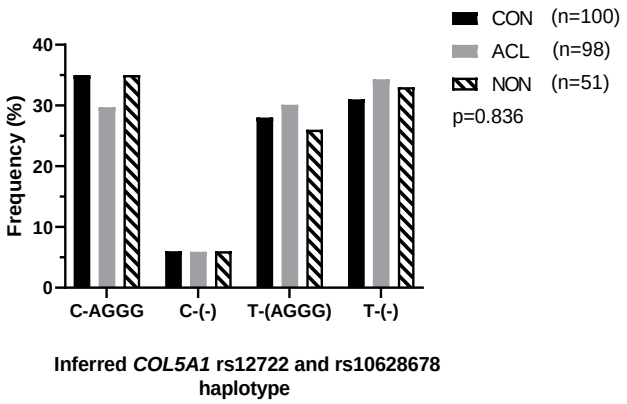


Figure 6.1. The distribution of the inferred haplotypes constructed from the *COL5A1* rs12722 (T/C) and *COL5A1* rs10628678 (AGGG/-) polymorphisms for the CON (black bars) group, ACL (grey bars) group and NON (hatched bars) sub-group. The global p-value is indicated under the legend key. The number (n) of subjects in each group is in parenthesis. The number (n) of subjects in each group is in parenthesis. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

Of the eight possible haplotypes constructed from *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (A/T), six were inferred at a frequency above 4% (Table B2.2). There were no significant differences in the inferred haplotype distributions between the CON and ACL groups, or the CON group and NON sub-group (Figure 6.2 A). Similarly, all four possible haplotypes constructed from the *COL11A1* rs3753841 (T/C) and *COL11A1* rs1676486 (C/T) polymorphisms were inferred at a frequency above 4% (Table B2.3). There were no significant differences in the inferred COL11A1 haplotype distributions between the CON and ACL groups or the CON and NON sub-group (Figure 6.2 B). Further, all four combinations constructed from *COL11A1* rs3753841 (T/C), and *COL11A2* rs1799907 (A/T) were inferred at a frequency greater than 4% (Table B2.4). However, none of the combinations were significant (Figure 6.2 C). Finally, haplotypes constructed from *COL11A1* rs1676486 (C/T) and *COL11A2*

rs1799907 (A/T) were all inferred at a frequency above 4% (Table B2.5), however, none were found to be significant (Figure 6.2 D).

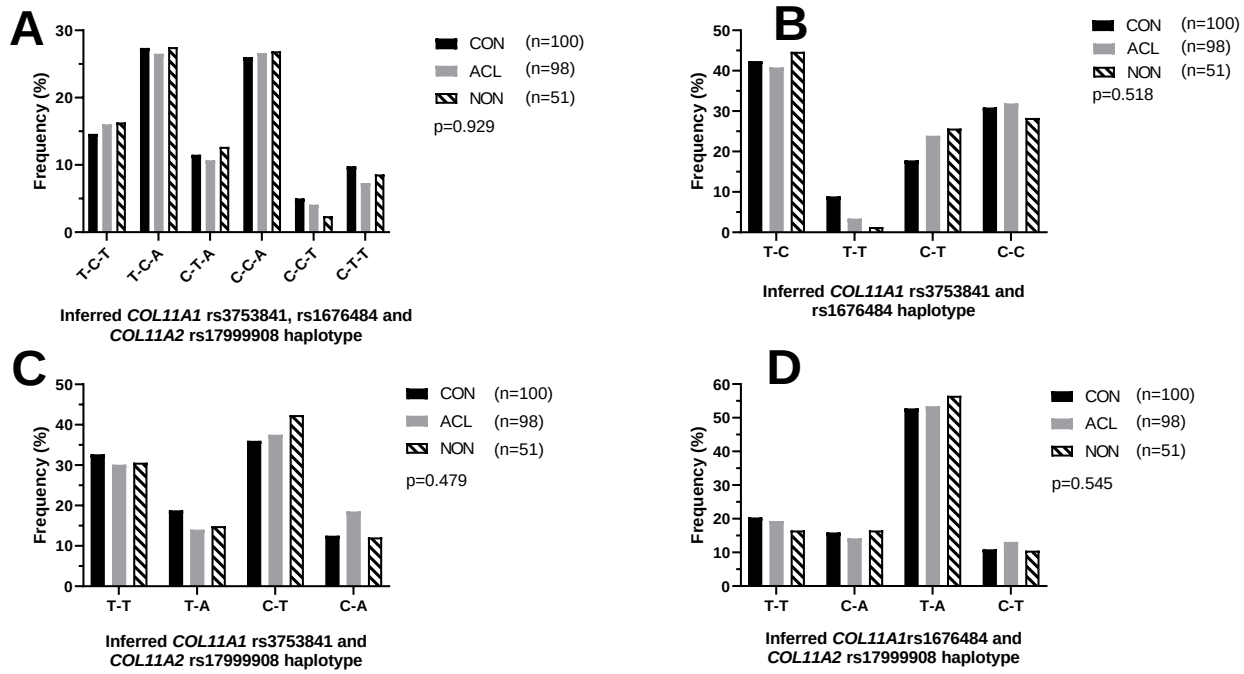


Figure 6.2. Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and rs1676486 (C/T), *COL11A2* rs1799907 (A/T) polymorphisms. **B)** Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and rs1676486 (C/T). **C)** Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and *COL11A2* rs1799907 (A/T) polymorphisms. **D)** Inferred haplotype frequency distributions constructed from *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) polymorphisms for the CON (black bars) group, ACL (grey bars) group and NON (hatched bars) sub-group. Significant differences ($p<0.05$) between the groups are indicated with a solid line and an asterisk (*). Global p-value is indicated under the legend key. The number (n) of subjects in each group is in parenthesis.

Nine of the 14 possible haplotypes constructed from *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T), *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) were inferred at a frequency greater than 4% (Table B2.6). None of the inferred haplotypes were significantly associated with ACL rupture (Figure 6.3 A).

Six of the eight possible haplotypes constructed from *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T), and *COL5A1* rs10628678 (AGGG/-) were inferred at a frequency above 4% (Table B2.7). However, none of the inferred haplotypes were significantly associated with ACL rupture (Figure 6.3 B). All eight possible haplotype constructed from COL*11A1* rs3753841 (T/C), *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) were inferred at a frequency above 4% (Table B2.8). However, none of the inferred haplotypes were significantly associated with ACL rupture irrespective of the mechanism of injury or when only those with the non-contact mechanism of injury were analysed (Figure 6.3 C).

Seven of the eight possible combinations constructed from *COL11A1* rs1676486 (C/T), *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) were inferred at a frequency above 4% (Table B2.9). However, none of the inferred haplotypes were significantly associated with ACL rupture (Figure 6.3 D). Furthermore, inferred haplotypes were constructed from *COL5A1* rs10628678 (AGGG/-) and each of the type XI collagen polymorphisms (rs3753841, rs1676486 and *COL11A2* rs1799907) (Figure B2.9 A-C). Although all possible combinations for each possible haplotype construction were inferred at a frequency greater than 4%, none were inferred haplotypes were significantly associated with ACL rupture (Table B2.10 – B2.12).

Finally, inferred haplotypes were constructed from each of the type XI collagen gene polymorphisms and *COL5A1* rs10628678 (AGGG/-). All 4 of the possible allele combinations of

the haplotypes constructed from *COL11A1* rs3753841 (T/C) and *COL5A1* rs10628678 (AGGG/-) (Figure B2.9 A), *COL11A1* rs1676486 (C/T) (Figure B2.9 B) and *COL5A1* rs10628678 (AGGG/-) and *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) (Figure B2.9 C) were inferred at a frequency greater than 4% (Table B2.10- B2.12). However, none of the inferred haplotypes were significantly associated with ACL rupture.

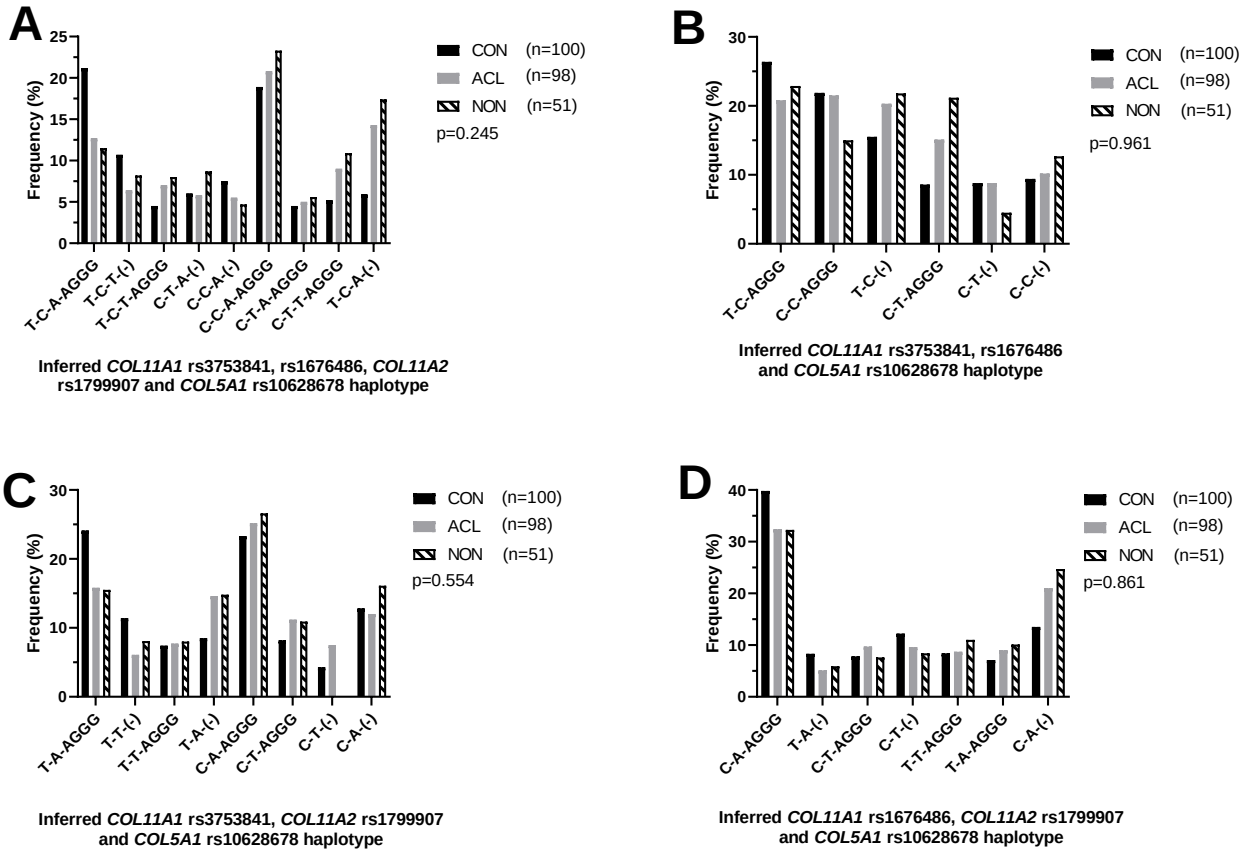


Figure 6.3. A) Inferred haplotype frequency distributions constructed from *COL5A1* rs10628678 (AGGG/-), *COL11A1* rs3753841 (T/C) and rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) polymorphisms. **B)** Inferred haplotype frequency distributions constructed from *COL5A1* rs10628678 (AGGG/-), *COL11A1* rs3753841 (T/C) and rs1676486 (C/T) polymorphisms. **C)** Inferred haplotype frequency distributions constructed from *COL5A1* rs10628678 (AGGG/-), *COL11A1* rs3753841 (T/C) and *COL11A2* rs1799907 (A/T). **D)** Inferred haplotype frequencies constructed from *COL5A1* rs10628678 (AGGG/-), *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (A/T) polymorphisms for the CON (black bars) group, ACL (grey bars) group and NON (hatched bars) sub-group. Global p-value is indicated under the legend key. The number (n) of subjects in each group is in parenthesis.

6.3.4.2 Types V and VI Collagen Gene-Gene Interactions

All four possible haplotypes constructed from *COL5A1* rs12722 (T/C) and *COL6A1* rs35796750 (T/C) were inferred at a frequency greater than 4% (Table B2.13). There were no significant differences in the inferred haplotype distributions constructed from *COL5A1* rs12722 (T/C) and *COL6A1* rs35796750 (T/C) between the CON and ACL groups, or the CON group and NON sub-group (Figure 6.4).

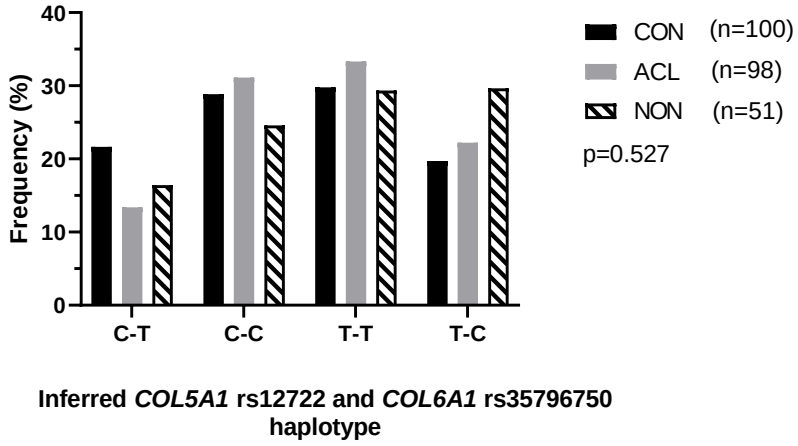


Figure 6.4. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) *COL6A1* rs35796750 (T/C) polymorphisms for the CON (black bars) group, ACL (grey bars) group and NON (hatched bars) sub-group. Global p-value is indicated under the legend key. The number (n) of subjects in each group is in parenthesis.

6.3.4.3 Types V and XII Collagen Gene-Gene Interactions

All four possible haplotypes constructed from *COL5A1* rs12722 (T/C) and *COL12A1* rs970547 (G/A) were inferred at a frequency greater than 4% (Table B2.14). The T-A haplotype was significantly under-represented in the CON group (17.9%) when compared to the NON group (27.3%; p=0.039), but not the ACL group (26.0%; p=0.090) (Figure 6.5).

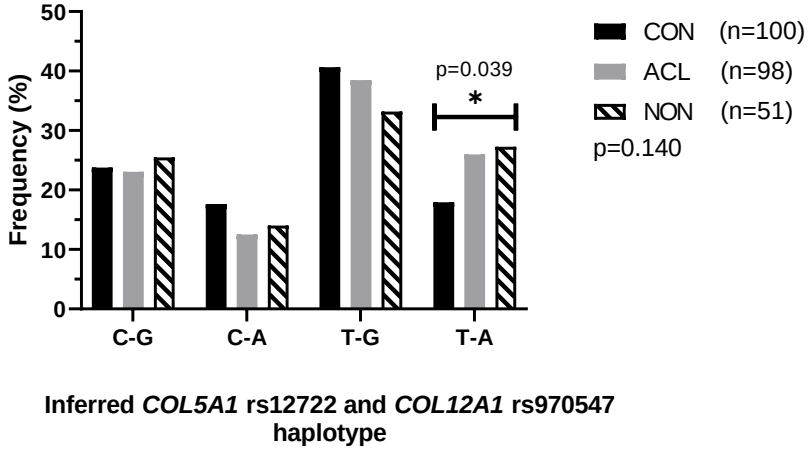


Figure 6.5. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C) and *COL12A1* rs970547 (G/A) polymorphisms for the CON (black bars) group, ACL (grey bars) group and NON (hatched bars) sub-group. Significant differences between the groups are indicated with a solid line and an asterisk for CON vs. ACL (*) and a pound (#) for CON vs. NON. The global p-value is indicated under the legend key. The number (n) of subjects in each group is in parenthesis.

6.3.4.4 Types VI and XII Collagen Gene-Gene Interactions

All four possible haplotypes constructed from *COL6A1* rs12722 (T/C) and *COL12A1* rs970547 (G/A) were inferred at a frequency greater than 4% (Table B2.15). In addition, it was found that the C-G inferred haplotype was significantly over-represented in the MACL and NON groups when compared to the CON group (37.2 vs. 31.1, p=0.029, MACL vs. CON and 34.3 vs. 31.1, p=0.027; NON vs. CON, respectively) (Figure 6.6).

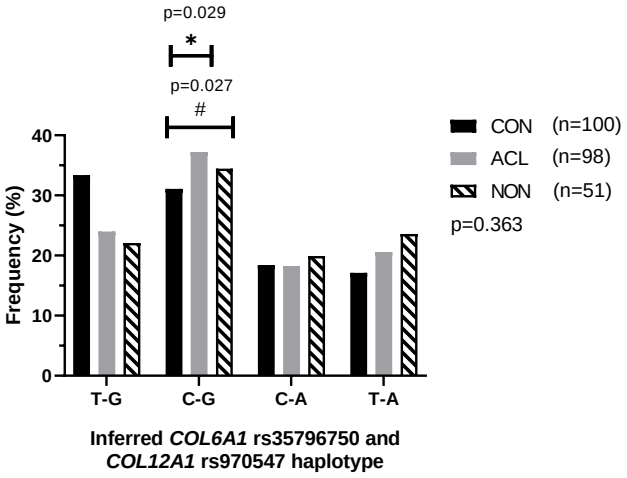


Figure 6.6. Inferred haplotype frequency distributions constructed from *COL6A1* rs35796750 (T/C) and *COL12A1* rs970547 (G/A) polymorphisms for the CON (black bars) group, ACL (grey bars) group and NON (hatched bars) sub-group. Significant differences between the groups are indicated with a solid line and an asterisk for CON vs. ACL (*) and a pound (#) for CON vs. NON. Global p-value is indicated under the legend key. The number (n) of subjects in each group is in parenthesis.

6.3.4.5 Types V, VI and XII Collagen Gene-Gene Interactions

Seven of the eight possible haplotypes constructed *COL5A1* rs12722 (T/C), *COL6A1* rs35796750 (T/C) and *COL12A1* rs970547 (G/A) were inferred at a frequency of greater than 4% (Table B2.16). None of the inferred haplotypes were significantly associated with ACL rupture.

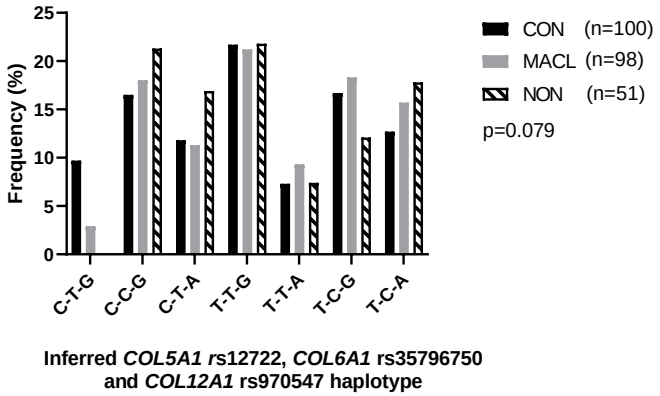


Figure 6.7. Inferred haplotype frequency distributions constructed from *COL5A1* rs12722 (T/C), *COL6A1* rs35796750 (T/C) and *COL12A1* rs970547 (G/A) polymorphisms for CON (black bars) group, ACL (grey bars) group and NON (hatched bars) sub-group. Significant differences between the groups are indicated with a solid line and an asterisk for CON vs. ACL (*) and a pound (#) for CON vs. NON.

6.4 DISCUSSION

Although genetic factors have been previously reported to influence the risk of ACL rupture, only a handful of studies have investigated genetic risk factors in a population of non-European ancestry

(94–98,518). Therefore, this study aimed to determine if previously investigated variants within the collagen encoding genes, specifically the *COL1A1* rs1800012, *COL5A1* rs12722, and rs10628678, *COL6A1* rs35796750, *COL11A1* rs3753841, and rs1676486, *COL11A2* rs1799907 and *COL12A1* rs970547, were associated with an ACL rupture in a mixed ancestry cohort.

The first finding of this study was that the GG genotype of *COL12A1* rs970547 polymorphism was independently associated with an increased risk of NON rupture. The remaining seven polymorphisms were not significantly associated with ACL rupture risk in this population.

Several studies have investigated the association of *COL12A1* rs970547 with ACL rupture risk in European (87,88,91), Asian (405,517,521), and Indian (499,522) populations. Previously, the *COL12A1* (rs970547 A/G) gene has been implicated in ACL ruptures within a European Ancestry South African population (87) but only after stratification by sex. The AA genotype was significantly associated with ACL rupture risk in Females only. No significant association was noted when the Male-only group was analysed. When assessed in a South African and Polish ACL rupture cohort, O'Connell et al. (89) the *COL12A1* rs970547 gene was not independently associated with risk. The authors observed that the T-A inferred pseudo-haplotype constructed from *COL5A1* rs12722 and *COL12A1* rs970547 was significantly associated with ACL rupture in females within a South African and Combined South African and Polish cohort. More recently, in non-European Ancestry populations, Zhao et al. (405) found that the *COL12A1* rs970547 polymorphism was not independently associated with ACL injury risk in a Chinese Han ACL injury population. Although, when stratified by sex, the AA genotype and A allele were significantly over-represented in the Male-only ACL injury group when compared to controls.

Similarly, a study of Korean ACL rupture patients found the opposite AA genotype was significantly associated with rupture risk (517). In contrast, the *COL12A1* rs970547 G allele was significantly associated with ACL risk in an Indian population (31). It is noteworthy that several studies have failed to show a significant association between the rs970547 and various MSK injuries (87,98).

As discussed in Chapter 2, section 2.7.5, type XII collagen is a member of the FACIT family of collagens (98,102) and is facilitates interactions between the cell-matrix and collagen fibres (309,402). Type XII is involved in the up regulation during wounding in response to excess mechanical stress (523). Type XII like type V collagen, have similar biological processes and both are theorised to regulate fibrillogenesis of collagen fibrils (403). The rs970547 SNP within the exon 65 is a non-synonymous coding variant resulting in an amino acid substitution of serine to glycine at position 3058 (524). This change could potentially damage the *COL12A1* peptide and may result in an altered function of type XII collagen. A possible reason for the difference in allele frequency could be due to a mixed ancestry population's heterogeneity. Thus, due to the inconsistency within the literature, further studies are required to elucidate the association of rs970547 and other *COL12A1* polymorphisms in modulating injury risk.

As discussed in Chapter 4, section 4.3.6, 15 gene-gene interactions were constructed using an a priori hypothesis. It was found that the T-A inferred haplotype, constructed from *COL5A1* rs12722 and *COL12A1* rs970547, was significantly associated with NON rupture risk. O'Connell et al. (89) noted similar findings, who also observed that T-A inferred haplotype was associated with ACL rupture in female Polish and South African cohorts. These finding are congruent with other genetic association studies, whereby the T and the A alleles of the *COL5A1* rs12722 and *COL12A1*

rs970547 were associated with increased risk of ACL ruptures in females (73,87,89). Furthermore, in line with previous studies investigating the *COL5A1* rs12722 (Chapter 2, section 2.7.2), the C allele has been previously implicated with reduced risk of several MSK pathologies (74,76,80).

An additional significant finding was that the C-G inferred pseudo-haplotype constructed from *COL6A1* rs35796750 and *COL12A1* rs970547 was significantly associated with ACL and NON rupture risk, respectively (89). *COL6A1* and *COL12A1* both encode for members of the non-fibrillar type VI and XII collagens, respectively (99). Both collagens are involved in fibrillogenesis and facilitate fibrillar interactions with the extra-cellular matrix through interactions with other collagens (525). Collins and Posthumus (388) previously proposed that changes within fibrillogenesis and subsequent collagen protein assembly may modulate the overall fibril architecture and structure, altering the tissue's mechanical properties. Future studies could potentially focus on the other collagen gene-gene interactions and how they potentially interact to influence ACL rupture risk.

In agreement with other non-European populations (31,236,377,378,405,521,522), the *COL1A1* rs1800012 TT genotype was rare and, therefore, not informative in the mixed ancestry cohort investigated in this study. Stratification by sex was not performed as the reduction in sample size may render the results unreliable due to a limited population size.

In conclusion, this is the first study to implicate collagen gene polymorphisms with risk of ACL rupture in a South African mixed ancestry population. In particular, the *COL12A1* rs970547 variant was independently associated with risk in non-contact ACL ruptures, in addition to identifying gene-gene interactions constructed from *COL5A1* rs12722, *COL6A1* rs35796750 and *COL12A1* rs970547 that were associated with altered risk of ACL ruptures. This study adds to the

growing body of literature investigating genetic risk factors for MSK injuries within a genetically diverse South African population. A continuous summary of the results of this thesis are presented in Table 6.3 and Table 6.4.

Table 6.3: A continuing summary of the genotype results from each chapter in this thesis. The results of Chapter 5, and preceding the preceding chapter, are shown.

Investigated Gene Variant	Investigated Phenotype			
	Chapter 4		Chapter 5	Chapter 6
	RCT	SST	EA ACL	MA ACL
COL1A1 rs1800012	n.s	n.s/Prev pub CC (30)	TT[a] (thesis)	n.s
		n.s	Prev pub TT (84)	n.s
			Prev pub GG (90)	
COL5A1 rs12722	n.s	n.s	n.s	
			Prev pub CC (79)	n.s
COL5A1 rs10628678	n.s	n.s	n.s	n.s
COL6A1 rs35796540	n.s	n.s	n.s	
			n.s Prev pub (89)	n.s
COL11A1 rs3754841	n.s	n.s	n.s	n.s
COL11A1 rs16746744	n.s	n.s	n.s	n.s
COL11A2 rs1799907	n.s	n.s	n.s	n.s
COL12A1 rs970547	n.s	n.s	n.s	GG (thesis)
			Prev pub AA (87,521)	

Green shading indicates a reduction in injury risk. Red shading indicates an increased risk of injury. Grey shading indicates no association with risk. Prev pub; Previously published. RCT; rotator cuff tendinopathy. SST; Supraspinatus tendinopathy. EA; European Ancestry. ACL; Anterior cruciate ligament rupture. MA; Mixed Ancestry.

[a] significant association in combined EA cohort but not in Swedish only cohort.

Table 6.4: A continuing summary of the haplotype results from each chapter in this thesis. The results of Chapter 6, and preceding chapters, are shown.

Phenotype		Investigated Gene variants						
		COL11A1 rs3753841 (T/C)	*COL11A1* rs1676486 (C/T)	*COL11A2* rs1799907 (A/T)	*COL5A1* rs12722 (T/C)	*COL5A1* rs10628678 (AGGG/-)	*COL6A1* rs35796750 (C/T)	*COL12A 1* rs970547 (A/G)
Chapter 4	RCT	C		A	(-) (thesis)			
	SST							
Chapter 5	EA				C	AGGG (thesis)		
	ACL				T	(-) (thesis)		
					T	AGGG (thesis)		
					Prev pub T	(-)(502)		
		Prev pub C	T	A (409)				
		Prev pub T	C	T	AGGG (409)			
		Prev pub T	C	T	(-)(409)			
		Prev pub C	T	A	AGGG (409)			
		Prev pub T	C	A	AGGG (409)			
							T	G (thesis)
							T	A (thesis)
Chapter 6	MA				T			A (thesis)
	ACL				Prev pub T			A (89)
							C	G (thesis)

Green shading indicates a reduction in injury risk. Red shading indicates an increased risk of injury. Grey shading indicates no association with risk. Prev pub; Previously published. RCT; rotator cuff tendinopathy. SST; Supraspinatus tendinopathy. EA; European Ancestry. ACL; Anterior cruciate ligament rupture. MA; Mixed Ancestry.

CHAPTER 7: SUMMARY AND PERSPECTIVES

Acute and chronic MSK injuries are considered multifactorial, with complex aetiologies that are the result of a combination of internally and externally derived risk factors (17,55,71). The current literature that suggests that inherited genetic components may influence the predisposition and susceptibility to injury (139,141,178,237,512). Although genetic sequence variants that encode for various structural and regulatory components of MSK tissues have been shown to be associated with injury risk in some capacity (507,526,527), this thesis focused on variants within specific collagen genes. The association of several of the investigated collagen gene variants in this thesis have previously been reported to associate with other injury and exercise-related phenotypes including performance, range of motion and exercise-associated muscle cramps (398,410,487,528). As such, collagen gene variants have been proposed to influence both the structural and functional aspects of tendons and ligaments and by implication alter the risk of injury (10,529).

The collagens, mainly type I, are the predominant structural proteins in tendons and ligament making between 70 to 85% of their dry mass (100,102,103). As reviewed in Chapter 2, previous literature has demonstrated that common collagen gene variants, specifically within *COL1A1*, *COL5A1*, *COL6A1*, *COL11A1*, *COL11A2* and *COL12A1*, have been reported to associate with MSK injuries, including ACL rupture (31,73,84,87–90,380,405,521), TEN (74–

77,80,82,385,386), tennis elbow (24–26), carpal tunnel syndrome (94,98), shoulder dislocations (27,174), shoulder instability**,** rotator cuff tendinopathy (RCT) (28,30,187,469) ankle injury (407,530) and/or the longitudinal ligament of the spine (395,531,532). With the exception of *COL1A1*, the association of polymorphisms within the other collagen genes have not been extensively investigated with ACL ruptures. The association of collagen gene variants with RCT has also not been extensively investigated.

Therefore, the aim of this thesis was to investigate whether the *COL1A1* rs1800012 (G/T), *COL5A1* rs12722 (T/C), *COL5A1* rs10628678 (-/AGGG), *COL6A1* rs35796750 (T/C), *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T), *COL11A2* rs1799907 (A/T) and *COL12A1* rs970547 (G/A) gene variants independently modulate risk of i) RCT in a South African cohort (Chapter 4), ii) ACL ruptures in Swedish and combined European ancestry cohorts (Chapter 5) and iii) ACL ruptures in a South African mixed ancestry population (Chapter 6). A secondary aim was to investigate hypothesis-driven collagen gene-gene interactions between variants in altering the risk of injury in the different RCT (Chapter 4) and ACL (Chapters 5 and 6) cohorts.

Finally, the non-genetic risk factors for ACL ruptures have been well described in the literature (13,20,212,219,220,231,233,247,412), and was therefore not extensively reviewed in this thesis. In contrast, despite the large volume of research, the non-genetic risk factors for RCT are still not yet fully understood or reviewed (Chapter 2, section 2.2.7). Therefore, a systematic review of the risk factors associated with RCT in swimmers was also included in this thesis (Chapter 3).

7.1 NOVEL FINDINGS OF THIS THESIS

7.1.1 Genetic Risk factors for rotator cuff tendinopathy in swimmers

7.1.1.1 Independent collagen variant interactions

In this thesis, none of the investigated collagen variants were independently associated with RCT or SST risk (Table 7.1). This contrasts with previous studies investigating other overuse musculoskeletal soft tissue injuries (21,99,519). The *COL1A1* rs1800012 variant has been previously associated with risk of shoulder dislocations (27) but not with rotator cuff disease (28,30). It is also important to note that there is no consistent definition of rotator cuff pathologies (21). Several authors have used a number of definitions that and it is therefore not surprising that there is a potential heterogeneity for this injury phenotype. Other overuse injuries such as, tennis elbow (24,25) and TEN (77) have also failed to show a significant difference in *COL1A1* rs1800012 genotype frequencies between cases and controls.

Table 7.1. Summary of independent genetic associations of several injury phenotypes.

Phenotype	*COL1A1* rs1800012	*COL5A1* rs12722	*COL5A1* rs10628678	*COL11A1* rs3754841	*COL11A1* rs16746744	*COL11A2* rs1799907	*COL6A1 rs35796540*	*COL12A1 rs970547*
RCT	n.s	n.s CC (30)	n.s	n.s	n.s	n.s	n.s	n.s
SST	n.s	n.s	n.s	n.s	n.s	n.s	n.s	n.s
SD	TT (27)							
RC Tear	n.s	n.s						
ACL	TT (Chap 5) TT (84) GG (88)	n.s CC (73)	n.s n.s	n.s n.s	n.s n.s	n.s	n.s n.s	GG (Chap 6) AA (87,521)
ATR	TT (77)	n.s	AGGG/AGGG (80)					
TEN	n.s	CC (80)		n.s	n.s	n.s	n.s	n.s
CTS	n.s	CC (94)	n.s	TT (98)		n.s		n.s
LDD	TT (383,399,406) GT (387)	n.s			CC (313)	n.s	n.s	
SLL							n.s	
TE	n.s	TT (24)		CT (26)	n.s			
AI		TT (407)						
TMJ		CT (184)						

RCT: Rotator Cuff Tendinopathy, Shoulder dislocation, SST; Supraspinatus Tendinopathy, RC; rotator cuff, ACL; Anterior Cruciate Ligament rupture, ATR; Achilles Tendon rupture, TEN; Achilles Tendinopathy, CTS; Carpal tunnel syndrome, LDD; Lumbar Disk Degeneration, SLL; spinal longitudinal ligament, TE; Tennis elbow; AI; Ankle Instability, TMJ; Temporomandibular Joint Degeneration.
n.s; Not significantly associated. Green highlight indicates reduced risk, red highlight indicates increased risk. Grey highlight indicates variants investigated in this thesis Bracket refers to specific reference.

Two case control genetic association studies, similar to the findings of this thesis, were also unable to find an association between *COL5A1* rs12722 and rotator cuff tears (187) and RCT (30). Conversely, a study by Alakhdar et al. (469), the *COL5A1* rs12722 CC genotype was significantly associated with RCT in a group consisting of 137 youth athletes. Additionally, the *COL5A1* rs12722 CC and rs10628678 AGGG/AGGG genotypes were independently associated with reduced and increased risk of ACL injury (502). In addition to the rs12722 CC genotype, the rs12722 TT genotype was associated with a reduction in risk of TEN (385). Interestingly, although

COL5A1 rs12722 and rs10628678 were not independently associated with chronic tendinopathy in a British population, inferred haplotypes constructed from three *COL5A1* polymorphisms, including rs12722 (C/T) and rs10628678 (AGGG/-), was associated with the risk of tendinopathy, where a haplotype containing the rs12722 C and rs3196378 deletion (-) alleles were associated with reduced risk (80). This underscores the crucial point that genetic regions should be scrutinized rather than focusing solely on a single polymorphism. The rs12722 CC and TT genotypes were also independently associated with reduced risk of CTS (94) with increased risk of tennis elbow (24), respectively. Recently the TT genotype has also been reported to be associated with increased risk of temporomandibular joint (TMJ) anterior disc displacement without reduction (408).

The *COL11A1* rs3754841, *COL11A1* rs16746744, *COL11A2* rs1799907 were not significantly associated with RCT. A similar finding was observed by Alakhdar et al. (469). Despite the lack of independent association with chronic tendinopathy (76), the *COL11A1* rs3754841 TT and CC genotype were implicated with an increased and decreased risk of CTS, respectively (98).

This is the first study to investigate the association of *COL5A1* rs10628678, *COL6A1* rs35796540 and *COL12A1* rs970547 variants with any shoulder injury. However, to date, none of these polymorphisms have been investigated in a rotator cuff pathology cohort. However, no independent genetic associations were observed. Type V, VI and XII collagens may have similar structure and function, gene-gene interactions have been investigated comparable overuse injury phenotypes also failed to find significant associations with the aforementioned variants.

The AGGG/AGGG variant of the *COL5A1* rs10628678 has been significantly associated with risk of chronic tendinopathy in Australian and South African populations (386) but not in a British Achilles Pathology cohort (80). Similar, the variant was not found to be significantly associated

with CTS (94,98). Whilst not independently associated, the *COL5A1* rs10628678 has been implicated in gene-gene interactions across several injury phenotypes (76,98,389,409,502). Whereas the *COL6A1* rs35796750 (T/C) variant has previously been associated with other complex phenotype conditions, such as posterior longitudinal ligament (395) and ligamentum flavum (397) ossification. The C allele of the rs35796750 variant has also been linked to a reduced risk of TEN (89), improved endurance cycling performance (398). Furthermore, the C allele has been associated with the reduced expression of *COL6A1*, which may result in a similarly reduced fibril diameter with concomitant increase in fibril density (89). Finally, the *COL12A1* rs970547 polymorphism has only been associated with ACL rupture risk (87,521) but not in AT (74) or CTS (98).

7.1.1.2. Collagen gene-gene interactions

While the association of collagen gene-gene interactions with other MSK soft tissue injuries have been investigated (76,80,89,94,98), this was the first study to examine the potential link between collagen gene-gene interactions with RCT and SST (Table 7.2). The novel finding of this thesis was that the C-A-(-) inferred haplotype constructed from *COL11A1* rs3753841(T/C), *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) was found to be significantly with reduced risk of RCT and SST. However, none of the inferred haplotypes constructed from i) the two *COL5A1* variants, ii) the two *COL11A1* variants, iii) the three *COL11A1* and *COL11A2* variants, as well as all iv) the *COL11A1, COL11A2* and *COL5A1* variants were associated with RCT or SST risk. Similarly inferred haplotypes constructed from v) *COL5A1* and *COL6A1*, vi) *COL5A1* and *COL12A1*, as well as viii) *COL6A1* and *COL12A1* were also not associated with RCT or SST.

Although Hay et al. (76) reported no independent associations between the *COL11A1* and *COL11A2* variants, an association was observed between the T-C-T inferred haplotype constructed from *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (T/A) was increased risk of Achilles tendinopathy in Australian and South African populations. The T-C inferred haplotype constructed from *COL11A1* rs3753841 and rs1676486 was associated with increased risk of CTS (98), while the C-C inferred haplotype was associated with reduced risk.

Table 7.2. Main Inferred haplotypes for chronic and overuse injuries.

Phenotype	*COL11A1* rs3753841 (T/C)	*COL11A1* rs1676486 (C/T)	*COL11A2* rs1799907 (T/A)	*COL5A1* rs10628678 (AGGG/-)	Reference
	T	C	T		(76)
	T	C	T	AGGG	(76)
	T	C			(409)
		C	A		(409)
TEN		C	T		(409)
		T	T		(409)
	T		T		(409)
	C	C	A	(-)	(409)
	C	T	A	(-)	(409)
	T	C			(98)
	T	C		(-)	(98)
CTS	C	C		AGGG	(98)
	T	C		AGGG	(98)
	C	C			(98)
RCT	C		A	-	Chapter 4
SST	C		A	-	Chapter 4

RCT: Rotator Cuff Tendinopathy, SST; Supraspinatus Tendinopathy, TEN; Achilles Tendinopathy, CTS; Carpal Tunnel Syndrome; Green highlight indicates reduced risk, red highlight indicates increased risk.

Since it has previously been shown that types V and XI collagen functionally interact during tendon development (266,298,402), the association of inferred haplotypes constructed from *COL11A1*, *COL11A2* and/or *COL5A1* variants with Achilles tendinopathy and CTS have been

investigated (76,98,409). The T-C-T-AGGG inferred haplotype constructed from *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T), *COL11A2* rs1799907 (T/A) and *COL5A1* rs10628678 (AGGG/-) was associated with increased risk of Achilles tendinopathy in a South African and Australian populations (76). Although two of inferred haplotype combinations associated weakly with reduced risk, the T-C-(AGGG) haplotype constructed from the *COL11A1* and *COL5A1* variants was strongly associated with increased risk of CTS (98). The novel finding of this thesis was that the complementary C-A-(-) inferred haplotype constructed from *COL11A1* rs3753841 (T/C), *COL11A2* rs1799907 (T/A) and COL5A1 rs10628678 (AGGG/-) was associated with reduced risk of RCT and SST.

7.1.2 Genetic Risk factors for ACL ruptures

Several collagen gene variants have been associated with ACL rupture risk in populations of European ancestry, including *COL1A1* rs1800012 (84,90), *COL5A1* rs12722 (73,89), *COL5A1* rs10628678 (389,502), *COL6A1* rs35796750 (89), *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 (76,409), and *COL12A1* rs970547 (87–89,503). An important limitation of these previous studies, which often produced conflicting results, has been small sample sizes and subsequently power. Therefore, although the association of these collagen gene variants in this thesis was investigated in an additional Swedish population, the primary aim was to investigate their association in a combined analysis consisting of previously published ACL rupture populations of European ancestry. These populations included Polish (88,508), RSA (89,409), Norwegian & Finnish (503), Swedish (507), Australian (502) and a combined cohort of RSA, Polish, Swedish, Norwegian & Finnish (500).

Furthermore, only a limited number of studies have investigated the association of collagen genes in non-European population groups (31,405,499,516,521,522). Previous investigations within a South African mixed ancestry population focused on the role of angiogenesis genes and ACL rupture risk (518). Therefore, an additional aim of this thesis was to determine if any of the collagen gene variants were associated with ACL injury risk in a South African mixed ancestry population.

Table 7.3. Main Inferred haplotypes of ACL ruptures.

Phenotype	*COL11A1* rs3753841 (T/C)	*COL11A1* rs1676486 (C/T)	*COL11A2* rs1799907 (A/T)	*COL5A1* rs12722 (T/C)	*COL5A1* rs10628678 (AGGG/-)	*COL6A1* rs35796750 (C/T)	*COL12A1* rs970547 (A/G)	Ref.
ACL (F)				C	-			(502)
ACL (A+F)				T	-			(502)
ACL, NON (A+M)	C	T	A					(409)
ACL, NON (A+M)	C	T						(409)
ACL	T	C	T		AGGG			(409)
ACL, NON (M)	T	C	T		-			(409)
NON (A)	T	C	A		AGGG			(409)
ACL				T			A	(409)
ACL	C	T	A		AGGG			

ACL; Anterior cruciate ligament rupture, NON; non-contact mechanism of ACL rupture.
Green colour indicates protection, Red colour indicates risk.
A: All participants; F: Female participants; M: Male participants. Haplotypes that were not found to be significant are not shown.

7.1.2.1 Independent collagen variant interactions

The *COL1A1* gene, more specifically, the rs1800012 polymorphism which has been one of the most extensively investigated polymorphisms across the several acute connective tissue pathologies. In agreement with the findings of this chapter, rare TT genotype has been previously shown to be associated a reduction of ACL rupture risk in Swedish (27) and South African cohorts

(84). It has been further implicated in Achilles tendon ruptures (77), ACL rupture in Polish Soccer players (90) as well as ACL ruptures from ski boot injuries (380). The T allele of *COL1A1* rs1800012 has been previously associated with greater SP-1 binding affinity and type I collagen production (374,533). As a result, this increases quantity and relative abundance of type 1 procollagen which may produce a protective effect against ligament and bone injuries (374).

The *COL1A1* rs1800012 TT genotype was found to be significantly under-represented in the ACL group of European ancestry. A novel finding of this study was that this association was only observed in the female but not male sub-groups. Although independently associated with ACL rupture in European populations, the *COL1A1* rs1800012 TT genotype was not associated with ACL rupture in the mixed ancestry cohort. The TT genotype of this variant is either absent or exceedingly rare in all non-European population groups, with prevalence ranging from 0% to 2.3% other population groups (Table 7.4).

A novel finding of this thesis was that like other non-European populations, the TT genotype (0.5%) was very rare in the South African mixed ancestry cohort. South African populations who self-identify as mixed ancestry have a varied genetic heterogeneity comprising of several ethnic and cultural groups. The complex and varied ancestry includes European, African, Middle-East and Far-Eastern populations (520). Hence, it is unsurprising that this polymorphism has not been found to be associated with ACL rupture in non-European population groups (31,405,521).

Table 7.4: The *COL1A1* rs1800012 (G>T) minor allele and genotype frequencies within different populations obtained from Ensembl (www.ensembl.org).

Population	n	Minor Allele (%)	Minor Genotype (%)
African	1302	6.4	0.5
American	694	12.4	2.3
East Asian	1000	0.1	0.0
South Asian	978	9.7	1.2
European	1000	18.9	4.4

A second novel finding was that the *COL12A1* rs970547 polymorphism was significantly associated with risk in a South African mixed ancestry ACL rupture cohort. Previously, the *COL12A1* rs970547 polymorphism has been implicated in ACL ruptures within South African females of European ancestry (87). It was noted that the AA genotype was significantly over-represented in the ACL group when compared to the CON group, in contrast to the findings of this chapter. The GG genotype was found to be significantly associated with risk of ACL rupture in the South African mixed ancestry population. Similarly, the GG genotype, and in particular the G allele, were found to be significantly associated with ACL rupture risk in an Indian population (499). However, studies that investigated East Asian populations including Chinese Han (405,521) and Korean (517) patients, the AA genotype and A allele were significantly associated with risk.

As described in Chapter 2, section 2.7.5, type XII collagen is a member of the FACIT family of collagens (99,267). The rs970547 polymorphism, according to functional analysis, undergoes a missense mutation, shifting the amino acid codon from glycine (G allele) to serine (A allele) which could alter the function of the α1(XII) collagen chain (534). Therefore, the A allele may modify the fibril diameter, potentially altering the biomechanical and structural properties of the

connective tissue (310,401). Further research is required to determine rs970547 polymorphism's role in modulating injury risk.

None of the other collagen variants were independently associated with ACL rupture in the European or mixed ancestry cohorts. Previous investigations within a South African mixed ancestry population are limited and have focused on the role of angiogenesis genes and ACL rupture risk (518) or CTS (94,98). Whilst not independently significant the *COL5A1* rs12722 and rs10628678 variants, together with two adjacent variants (rs146774622 and rs55748801) were found to be significantly under-represented in the CTS group (94). The WW+CC genotypes resulted in a 2.1-fold decrease in CTS risk, congruent with similarly reported risk decrease in European TEN cohorts (82). Secondly, Dada et al. (98) found that the CC genotype and the TT genotype of the *COL11A1* rs3753741 were significantly over-represented in the CON and CTS groups, respectively. Moreover, the authors noted that the C-C and T-C haplotypes constructed from the *COL11A1* rs3753841 and rs1676486 were significantly over-represented in the CON and CTS groups, respectively (98). Although not significantly associated with risk, this is the first study to investigate the *COL6A1* rs35796750 with ACL rupture risk in a mixed ancestry population.

7.1.2.2 Collagen gene-gene interactions

Whilst not significant in either of the independent Swedish or Combined European ancestry cohorts, the *COL5A1* rs12722 polymorphism, more specifically, the 3'-UTR, has been previously associated with risk of ACL rupture in females (73), in addition to TEN (80). However, the C-

AGGG and T-(-) inferred haplotypes constructed from the *COL5A1* rs12722 and rs10628678 variants were associated with increased risk of ACL rupture, while the T-AGGG inferred haplotype was associated with reduced risk, in the combined male and female group and female-only group. Recently, Alvarez-Romero et al (502) also reported that the T-(-) inferred haplotype was associated with increased risk when only the same South African and Australian cohorts were analysed together. There are limited studies that have investigated the association of the *COL5A1* rs10628678 variant with MSK injuries and thus should be considered for further investigation given the possibility of interaction with other collagen gene variants.

Of the several previously investigated collagen polymorphism interactions (76,98,409), it has been previously shown that type V collagen interacts with type XI collagen (103,252) during fibrillogenesis during tendon development. In contrast to previous evidence, this chapter found that none of the investigated type XI collagen gene variants were independently associated with ACL rupture in the Swedish and combined cohorts. Nor were any gene-gene interactions observed. Within a South African ACL injury cohort, Hay (409) found that the TC genotype was significantly over-represented in the CON group when compared to the ACL group. Although genotype distributions were similar between the ACL and NON group, the NON group was not significantly different to CON group. When stratified by sex, the frequency distributions remained non-significant. Similarly, *COL11A1* rs1676486 and *COL11A2* rs1799907 were not found to be significantly different between the ACL, NON and CON groups (409).

Interestingly, significant associations were found when inferred haplotypes were constructed from *COL11A1* rs3753841 and *COL11A1* rs1676486 variants. The C-T inferred haplotype was significantly over-represented in the CON group compared to the ACL group. When stratified by

sex, the C-T inferred haplotype was also significantly over-represented in CON group when compared to the NON sub-group in female, but not male, participants. Moreover, when haplotypes were constructed using *COL11A1* rs3753841, *COL11A1* rs1676486, and *COL11A2* rs1799907, the C-T-A haplotype was found to be significantly over-represented in the CON group when compared to the ACL group and NON sub-groups. This significant association was only observed in the male participants when analysed separately. Hay (409) further observed genotype-genotype interactions between *COL11A1* rs3753841 and rs1676486, *COL11A2* rs1799907 and *COL5A1* rs71746744 variants. The T-C-T-AGGG inferred haplotype was significantly over-represented in the CON group compared to the NON sub-group. Additionally, the C-T-A-AGGG haplotype was over-represented in the CON group when compared to the ACL group. When stratified by sex, significant associations were observed in the male participants where the T-C-T-AGGG and C-T-A-AGGG haplotypes were significantly over-represented in the CON group when compared to the ACL group. Whereas the distribution of the T-C-T-(-) inferred haplotype was significantly different to in ACL group and NON sub-group when compared to the CON group. No significant differences were observed in the female participants (409). As such, the possible role of Type XI collagen needs to be further investigated.

As previously discussed, *COL6A1* rs35796750 and the *COL12A1* rs970547 have been previously associated with several multifactorial conditions (397,535) including TEN and ACL rupture (87,89). While neither variant was independently associated with ACL rupture risk in the Swedish and Combined ACL groups, a significant gene-gene interaction was observed. The T-A inferred haplotype was found to be significantly over-represented in the combined ACL group compared to the healthy controls. Interestingly, the T-G inferred haplotype was significantly associated with reduced risk when only the male participants were analysed. Both collagen type VI and XII are

non-fibrillar collagens and are expressed in connective tissues such as bone and skeletal muscle (514,515). The function of type VI collagen is still being investigated; however, it has been implicated in the development and maintenance of basement membranes of connective tissues (393). Type V, VI and XII collagens can mediate cell-matrix and matrix-matrix interactions (515), which are important features regulating cell migration, adhesion, apoptosis, and are crucial for matrix bridge formation. Izu et al. (515) demonstrated that deficiencies in collagens VI or XII impair matrix bridge formation particularly in bone forming sites, potentially affecting connective tissues such as ligaments and altering risk. Future studies should focus on the potential role of *COL5A1* and *COL12A1* in injury risk.

Since type V and type XII collagen may have similar structure and function, gene-gene interactions were investigated. As observed in Chapter 6, the T-A inferred haplotype, constructed from *COL5A1* rs12722 and *COL12A1* rs970547, was significantly over-represented in the NON sub-group. Similarly, O'Connell et al. (89) also noted the T-A inferred haplotype to be associated with ACL rupture in two European ancestry populations. Whilst not significant the current mixed ancestry population, the T allele of the *COL5A1* rs12722 polymorphism was found to be significantly associated with ACL rupture in a female European ancestry population (73,87,89). It is interesting to note that the G allele of the *COL12A1* rs970547 variant was independently associated risk, but when examined in conjunction with the rs12722 polymorphism the A allele produces risk. A possible reason for the difference in allele frequency could be due to a mixed ancestry population's heterogeneity. However, several studies have failed to report a significant association with risk (80,187,502,503,507). Further robust and population-specific studies are required to fully understand the risk modulation of ACL rupture.

A further gene-gene interaction was observed between *COL12A1* rs970547 and *COL6A1* rs35796750. The C-G of the *COL6A1* rs35796750 and *COL12A1* rs970547 was significantly over-represented in the ACL group. Whilst not independently associated with risk in this thesis, the *COL6A1* rs35796750 (T/C) variant has previously been associated with other multifactorial conditions, posterior longitudinal ligament (535) ligamentum flavum (397) ossification. The C allele of the rs35796750 variant has also been implicated in reduced risk of TEN (89), improved endurance cycling performance (398). Furthermore, the C allele has been associated with reduced *COL6A1* gene expression, may result in a similarly reduced fibril diameter with concomitant increase in fibril density (411). Consequently, these structural alterations may influence the stiffness and mechanical loading ability of tissues that contain type VI collagen. The association of *COL6A1* rs35796750 and *COL12A1* rs970547 with ACL ruptures has however not been extensively investigated.

7.1.3 Non-genetic Risk factors for Shoulder injuries in swimmers

As presented in Chapter 3, only four non-genetic risk factors for shoulder injuries were determined to be of moderate certainty, with the remaining 25 risk factors being appraised as low certainty. Moderate level of certainty was determined in i) previous history of pain and injury, ii) internal/external rotation range of motion, iii) clinical joint laxity and instability and iv) internal/external rotation strength. Swimmers with a history of shoulder pain were 4.1 and 11.3 times more likely to sustain a subsequent injury for significant interfering pain in the shoulder and a significant shoulder injury, respectively (48,54,424). While the relationship between previous injury and subsequent shoulder injury is unclear, it may relate to other pre-existing factors or insufficient rehabilitation after initial injury (54,536). Whereas i) internal/external rotation range

of motion, ii) clinical joint laxity and instability and iii) internal/external rotation strength relate to shoulder joint characteristics and function. Imbalances, in both strength and range of motion may contribute to instability, joint laxity and changes in range of motion, predisposing an athlete to risk of injury (18,59,202,441).

There is no single factor that can determine joint laxity or range of motion, several intrinsic and extrinsic factors have been shown to influence the mechanical properties of tendons, ligaments and muscles. Posthumus et al. (17) reported generalised joint laxity and anterior knee laxity as potential risk factors for ACL rupture. As described in Chapter 2, section 2.7.2., the relative content of type V collagen has been proposed to influence the mechanical properties of the tendons and ligaments through the alteration of fibril diameter and packing density (487,537) which can be problematic given the limited expansion space within a shoulder joint. Similarly, Sein et al (59) proposed that repetitive movement that occurs during intensive training may leads to an increase in tendon thickness, mechanical impingement, and inflammation. Previously, the *COL1A1*, *COL5A1*, *COL11A1* and/or *COL11A2* variants investigated in this thesis were reported to associate with several measures of knee joint laxity and changes in knee ligament length in healthy physically active individuals (538).

One of the most extensively studied sport-specific risk factors (Figure 7.1) for injury is training load (537,539,540). Within this cohort, both RCT and SST groups trained significantly more hours per week than the CON group. Both the RCT and SST groups reported training significantly more hours per week in both summer and winter compared to the CON group. During training, a significant amount of repetitive load is placed on the shoulder, resulting in approximately 2500-9600 overhead rotations completed per day (493,494). As a result, several possible physiological

adaptations have been noted in the literature, which include tendon thickness (59,450,493,494), tendon degeneration (433), reduced joint stability (202), increased range of motion(51,442), muscle fatigue (495), altered scapular kinematics (436,444,457) and increased mechanical impingement (206,433,466) (Figure 4.15). These adaptations have previously associated with tendinopathy (40,151,421,496).

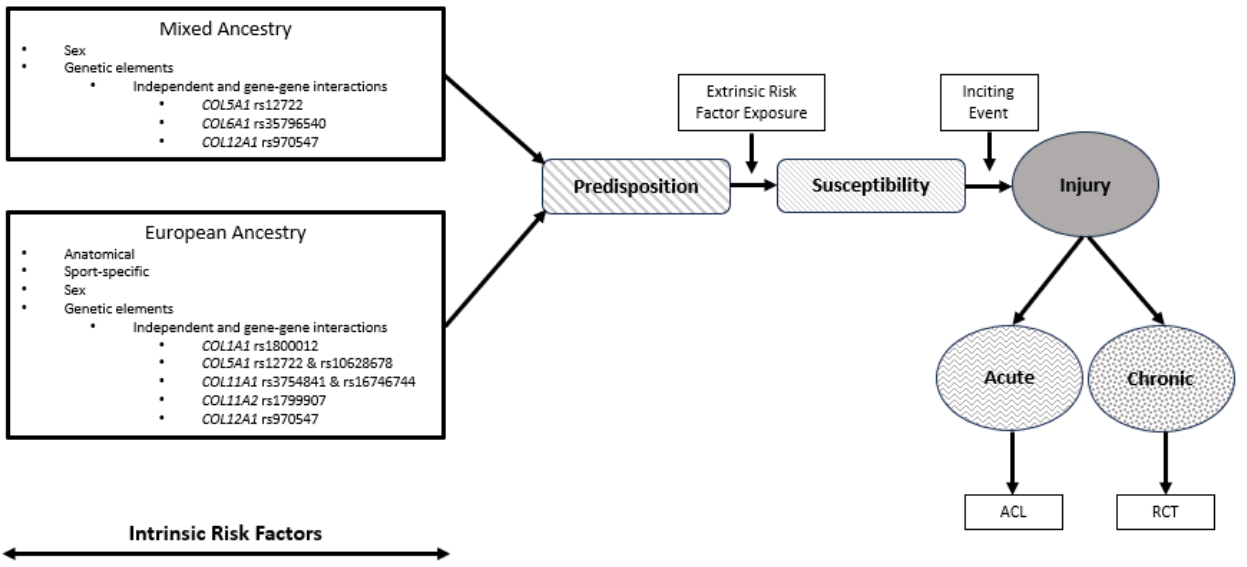


Figure 7.1. A schematic diagram, adapted from Meeuwisse's (14) proposed injury model of the dynamic and multifactorial aetiology of injury. The model outlines the complex relationship between intrinsic and extrinsic risk factors relating to injury predisposition, susceptibility, inciting events and ultimately injury outcomes. The findings of this thesis have been incorporated into the model. Independent and gene-gene interactions were identified and were associated with injury risk and therefore may influence overall risk outcomes. ACL; Anterior cruciate ligament rupture, RCT; Rotator cuff tendinopathy.

As previously discussed in Chapter 2, section 2.3 of, and in agreement with Posthumus et al. (87), participants who reported having sustained a previous joint injury or a history of multiple tendon injuries were 2.3x greater risk of RCT and 2.1x great risk of SST. Although a family history of

MSK injuries has been shown to increase the risk of injury (87,139,141,237,510), the current study did not find any significant differences between CON, RCT and SST. However, a trend towards significance was noted in the RCT and SST group and a family history of tendon injuries.

To date, no studies have investigated the role of specific collagen variants in the aetiology of RCT in swimmers. It is therefore a reasonable hypothesis that alterations of connective tissue properties may contribute to RCT in swimmers. The specific candidate genes examined in this thesis were selected based on associations with exercise-related and injury phenotypes. The hypothesis tested in this thesis was if genetic polymorphisms that have been previously linked to injury risk, were also associated with overuse injuries, such as RCT.

Whilst several risk factors were determined to be a low level of certainty, it is possible that a lack of true association may be in part, due to the lack of prospective studies of sufficient quality. The evidence available from the literature may have not been sufficient to facilitate a reliable estimate of risk due to i) the limited number of studies that investigated a particular risk factor, ii) a clear lack of consistency of the definition of shoulder pain or injury iii), limited amount of high-quality (Level 1 and 2) studies, and iv) the inconsistency of the reported findings.

7.2 LIMITATIONS AND FUTURE STUDIES

Several limitations to this thesis were noted. Firstly, the small sample size of the RCT cohort, yet it does provide insight into potential mechanism of multifactorial musculoskeletal injuries. Furthermore, the heterogeneity of the RCT, although is well described in literature, the injury is thought to be multifactorial with several external factors contributing to the onset of tendinopathy (21). However, the tests to confirm diagnosis was done by qualified and experienced orthopaedic

surgeons. Moreover, the majority participants did not undergo surgery, but were instead treated conservatively including pharmaceutical interventions, and with physical therapy.

Secondly, the pooling of several seemingly similar ancestral populations groups could have potentially resulted in biases that would require further refinement and investigation. Thirdly, part of the sourced data was obtained from the publicly available literature as opposed to the raw data. In order to accommodate for this, only genotype frequencies were pooled for combined analysis involving of South African, Polish, Swedish, Norwegian, Finish and Australian participants. Whereas haplotypes were constructed only when raw genetic data was available, as the case in South African, Swedish, and Australian participants. Future studies should consider the development and sharing of databases that would improve our analytical ability to determine risk of injury.

Fourthly, sample size of the mixed ancestry cohort was small, and in particular female participants. Several studies have shown that females are more at risk for ACL injuries. The low numbers could potentially be attributed to the recruitment method, as outlined by Rahim et al. (518) whereby less females were admitted to hospital for ACL reconstructive surgery. Secondly, the populations, like the mixed ancestry population, may not be ideal for investigation of new candidate genes, however this thesis utilised an *a priori* hypothesis and all variants had been previously investigated or implicated in other phenotypes. Further, the mixed ancestry population may contain more genetic variation and heterozygosity than ancestral European South Africans (37,94,96–98,518). Therefore, results will need to be interpreted with caution until further studies become available.

7.3 PRACTICAL IMPLICATIONS

The results of this thesis may have several significant practical implications in sports medicine and injury prevention. Two major areas investigated by this thesis include 1) critical implications for swimmers and coaches and 2) Genetic profiling of athletes and the general population.

7.3.1 Critical Implications for Swimmers and Coaches

The research findings underscore the importance of understanding the cumulative effects of a week's training loads on shoulder physical qualities in competitive swimmers. This knowledge is crucial for athletes and coaches in the swimming community as it can significantly impact shoulder health, performance, and the risk of injury.

7.3.1.1 Injury Prevention and Management

The research indicates that high training volumes can lead to acute maladaptation in shoulder physical qualities, such as decreased external rotation range of motion and isometric rotation torque (541,542). This knowledge can be used to develop targeted injury prevention programs that include specific exercises to counteract this maladaptation. Understanding the relationship between training loads and shoulder pain can help design more effective management strategies for swimmers experiencing shoulder discomfort.

7.3.1.2 Training Program Design

The findings of this study can be directly applied to the design of training programs. Coaches and trainers can adjust training volumes and intensities to mitigate the risk of shoulder injuries. For instance, incorporating high-intensity training days with lower-intensity sessions can help manage the cumulative stress on swimmers' shoulders (539,543). Furthermore, dividing

swimmers into high-volume and low-volume training groups based on their shoulder health status can lead to customizing training programs that cater to individual athlete needs.

7.3.1.3 Monitoring and Assessment

Regular monitoring of shoulder function in swimmers, using measures such as external rotation range of motion and isometric rotation torque, can serve as an early warning system for potential shoulder problems (541,542). Implementing routine wellness assessments can help track the broader impact of training loads on athletes' general health and inform adjustments to training plans.

7.3.1.4 Rehabilitation and Recovery

The findings underscore the importance of incorporating adequate recovery strategies into training schedules. This includes sufficient rest, proper nutrition, and possibly physiotherapy interventions focusing on shoulder health (Tovin, 2006). Rehabilitation programs might need to be tailored to address specific deficits identified through regular assessments, such as reduced pectoralis minor length or decreased shoulder external rotation (453,544,545).

7.3.1.5 Educational Implications

Educating swimmers about the signs of shoulder overuse and reporting these symptoms early can lead to prompt interventions and reduced downtime due to injuries. Workshops and training sessions for coaches on the latest research findings related to swimmers' shoulders can improve overall training quality and athlete care. Integrating these insights into training and rehabilitation programs can reduce the incidence of shoulder injuries, improve performance, and extend swimmers' athletic careers.

7.3.2 Implications for Genetic profiling of athletes and the general population

Although currently premature, integrating genetic information into sports injury risk assessment and personalized training strategies may significantly affect athletes and the broader sports community. While there is indication to suggest that genetic variants may be implicated in MSK risk, there still remains some uncertainty around the biological and clinical relevance of these genetic markers (546). Despite the potential benefits, the use of genetic testing for predicting injury risk and other health-related traits comes with challenges. The predictive value of these tests can be limited by the complexity of genetic and environmental interactions that contribute to traits like injury susceptibility. For instance, while genetic tests can potentially, one day, provide valuable information, they must be interpreted with caution due to the multifactorial nature of most injuries and out limited understanding of the genetic risk factors that contribute to injury susceptibility (547). Until the scientific community generates and investigates more evidence, any conclusions about genetic profiling for athletes and the public, in general, are still premature. With this in mind, we may speculate on the following implications.

7.3.2.1 Personalized Injury Prevention and Training

Genetic research has the potential to eventual provide insights that can significantly enhance personalized training and injury prevention strategies for athletes. By understanding an individual's genetic predisposition, athletes can tailor their training, nutrition, and recovery practices to better align with their genetic profile (547,548). This approach could help reduce the risk of injuries and optimize performance by enhancing the strengths and mitigating the weaknesses identified through genetic markers.

7.3.2.2 Enhanced Understanding of Injury Risks

Applying genetic information also has the potential to develop a deeper understanding of an athlete's susceptibility to specific injuries, such as ligament tears, tendon injuries and/or muscle strains. This knowledge can lead to more targeted preventive measures, potentially reducing the incidence and severity of sports-related injuries (548–550). For instance, athletes predisposed to certain muscle injuries might adopt specific strengthening exercises or adjust their training intensity and volume to prevent such injuries.

7.3.2.3 Ethical and Privacy Considerations

While the potential benefits of genetic testing in sports are clear, it also raises ethical and privacy concerns. The potential misuse of genetic data to select athletes or determine their career paths without considering other critical factors like training, psychological readiness, and environmental influences is a significant concern (551–553).

7.3.2.4 Future Research and Methodological Improvements

Current research, while promising, often suffers from tiny sample sizes and methodological limitations. Future studies with larger cohorts and improved methodologies are necessary to validate the effectiveness of genetic testing in reducing injury rates and enhancing athletic performance (552,554–557). Moreover, expanding this research to include diverse populations will help generalize the findings across different racial and ethnic groups.

In conclusion, although currently premature, using genetic information in sports offers exciting possibilities for enhancing athletic performance and reducing injury risks. However, careful

consideration of ethical standards and privacy protections would be required to ensure that the benefits are realized without compromising the rights and well-being of athletes.

7.4 SUMMARY AND CONCLUSIONS

This thesis investigated two common injury phenotypes, RCT and ACL rupture, from a genetic context by investigating candidate genes that have been implicated in specific MSK injuries, as a potential modulator of injury risk.

The onset of RCT seems to differ from other chronic injuries (TEN, Tennis Elbow, CTS), potentially due to the significant volume and load that is placed on the shoulder joint during swimming. Whilst none of the gene variants investigated in this population were significant, a significant gene-gene interaction was observed. Since the space within the shoulder joint is so limited, any structural changes to underlying connective tissue either through physical activity or genetic predisposition may alter the risk of developing RCT. Although RCT is one of the most common pathologies, the exact aetiology is still unknown. It is also interesting to note that there are non-collagen gene associations in RCT that aren't found in other injury phenotypes (30,140,179,180). It is not surprising since similarities and differences between tendon and ligament injuries may be influenced by sex (73,231,503,558,559), injury location (16,20,48,218,223,560) and type of injury (acute vs. chronic) (561–563). It is therefore expected that we would observe genetic differences between these different injury phenotypes.

Contrary to the expected independent or gene-gene interaction associations of at least the types V and XI gene variants, the collagen gene variants investigated in thesis are not associated with RCT. It is possible to theorise that individuals with these functional variants could produce type V and

XI collagens, which may be responsible for changes in the biomechanical properties of the tendon which could ultimately lead to pathology. While the selection of these candidate genes was hypothesis-driven, and despite failing to find significant associations, they are still biologically relevant to tendon biology as they have been previously implicated in other multifactorial conditions and injuries. This thesis adds to the growing evidence that suggests that the various collagens that comprise the tendon fibre of the RC may interact to effect risk outcomes of RCT. There is also the potential possibility that other polymorphisms within the selected candidate genes may also be involved in the aetiology of RCT. Future research should also focus on the functional mechanisms and biological processes that may under this complex phenotype. Further exploration of single variants as well as gene-gene interactions may assist in the understanding of the genetic contribution to risk of pathology.

Furthermore, the genetic predisposition of ligament injuries, specifically ACL ruptures in individual European ancestry populations, has been established in the literature (73,84,87–91,380,500,507,508). In particular, *COL1A1* and *COL5A1* variants have been the most extensively researched. This thesis, by investigating several other lesser studied variants, provides more evidence of the complex nature of genetic risk and MSK injuries. A common limitation in the majority of genetic studies is that of sample size, which may constrain generalisability and our ability to determine true risk. Through combining several populations of reported European ancestry, this thesis was providing further evidence of the role of *COL1A1* rs1800012 as a potential risk factor for ACL rupture. Although no other independent associations were found within the selected candidate genes, it does not rule out their potential involvement in risk modulation. Several gene-gene interactions were observed and through this identification, assist in understanding the genetic predisposition to risk. Future studies should examine the potential

functional effects of collagen encoding genes such as *COL5A1* rs10628678, *COL6A1* rs35796840, *COL11A1* rs3753841, *COL11A1* rs1676846, *COL11A2* rs1799907 and *COL12A1* rs970547.

This thesis highlights the potential role of collagen gene variants in risk of ACL rupture within European Ancestry populations. This thesis supports previous observations that the *COL1A1* variant may play a significant role in the mediation of ACL rupture risk, in addition to other implicated variants noted in this thesis. ACL rupture is a complex and multifactorial condition and is influenced by several intrinsic and extrinsic factors (564). The limited available literature highlights the importance of investigating genetic associations in genetically diverse populations and related phenotypes, which allows for a deeper exploration of the important biological processes underlying a potential susceptibility to MSK injuries (518,519). This is, to our knowledge the first research to investigate collagen gene polymorphisms, previously associated with MSK injury risk, in a South African self-reported mixed ancestry population from the Western Cape region of South Africa.

Finally, thesis further emphasises the difficulty in identifying intrinsically derived components and individual risk factors that alter predisposition and subsequent susceptibility to MSK injuries. Larger collaborative studies are desperately needed to further examine how genetic variants, particularly within collagen proteins may interact to modulate and alter risk of ACL rupture, not only within ancestral European cohorts, but in other populations as well.

REFERENCES

1. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. CMAJ. 2006 Mar;174(6):801–9.
2. Janssen I, LeBlanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. International Journal of Behavioral Nutrition and Physical Activity. 2010;7.
3. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity – a systematic review of longitudinal studies. BMC Public Health. 2013 Dec;13(1):813.
4. Mandolesi L, Polverino A, Montuori S, Foti F, Ferraioli G, Sorrentino P, et al. Effects of Physical Exercise on Cognitive Functioning and Wellbeing: Biological and Psychological Benefits. Front Psychol. 2018 Apr;9.
5. Schneider S, Seither B, Tönges S, Schmitt H. Sports injuries: population based representative data on incidence, diagnosis, sequelae, and high risk groups. Br J Sports Med. 2006 Apr;40(4):334–9; discussion 339.
6. Sreekaarini I, Eapen C, Zulfeequer CP. Prevalence of Sports Injuries in Adolescent Athletes. J Athl Enhanc. 2014;03(05).
7. Nicholl JP, Coleman P, Williams BT. The epidemiology of sports and exercise related injury in the United Kingdom. Br J Sports Med. 1995 Dec;29(4):232–8.
8. Malcolm D. Sport, Medicine and Health. New York, NY: Routledge; 2017.
9. Hauser RA. Ligament Injury and Healing: A Review of Current Clinical Diagnostics and Therapeutics. Open Rehabil J. 2013;6(1):1–20.
10. Collins M, September A V., Posthumus M. Biological variation in musculoskeletal injuries: Current knowledge, future research and practical implications. Br J Sports Med. 2015;49(23):1497–503.
11. Norris CM. Sports injuries: diagnosis and management. 3rd Ed. Oxford: Butterworth Heinemann; 2004.
12. Sayampanathan AA, Andrew THC. Systematic review on risk factors of rotator cuff tears. J Orthop Surg (Hong Kong) [Internet]. 2017;25(1):1–9. Available from: https://doi.org/10.1177/2309499016684318
13. Bahr R, Krosshaug T. Understanding injury mechanisms: A key component of preventing injuries in sport. Br J Sports Med. 2005;39(6):324–9.
14. Meeuwisse WHWH. Assessing Causation in Sport Injury: A Multifactorical Model. Clinical Journal of Sport Medicine. 1994;4(3):166–70.

15. Serpell BG, Scarvell JM, Ball NB, Smith PN. Mechanisms and risk factors for noncontact ACL injury in age mature athletes who engage in field or court sports: a summary of the literature since 1980. J Strength Cond Res. 2012 Nov;26(11):3160–76.

16. Vaughn NH, Stepanyan H, Gallo RA, Dhawan A. Genetic factors in tendon injury: A systematic review of the literature. Orthop J Sports Med. 2017;5(8):1–11.

17. Posthumus M, Collins M, September A V., Schwellnus MP. The intrinsic risk factors for ACL ruptures: An evidence-based review. Physician and Sportsmedicine. 2011;39(1):62–73.

18. Seitz AL, McClure PW, Finucane S, Boardman ND, Michener LA. Mechanisms of rotator cuff tendinopathy: Intrinsic, extrinsic, or both? Clinical Biomechanics [Internet]. 2011;26(1):1–12. Available from: http://dx.doi.org/10.1016/j.clinbiomech.2010.08.001

19. Järvinen TAH, Kannus P, Maffulli N, Khan KM. Achilles tendon disorders: Etiology and epidemiology. Foot Ankle Clin. 2005;10(2):255–66.

20. Pfeifer CE, Beattie PF, Sacko RS, Hand A. Risk Factors Associated With Non-Contact Anterior Cruciate Ligament Injury: a Systematic Review. Int J Sports Phys Ther. 2018;13(4):575–87.

21. Hill L. Genetic studies in shoulder pathology: Current status of knowledge. Vol. 70, Minerva Ortopedica e Traumatologica. Edizioni Minerva Medica; 2019. p. 137–52.

22. Collins M. Molecular Genetics An Important Tool In Elucidating The Molecular Mechanisms Underlying Tendinopathies. Br J Sports Med. 2013 Jun;47(9):e2.7-e2.

23. Posthumus M, September A, Schwellnus MP, Collins M. Identification of genetic risk factors underlying complex multifactorial phenotypes. Knee Surgery, Sports Traumatology, Arthroscopy. 2010;18(12):1810–1.

24. Altinisik J, Meric G, Erduran M, Ates O, Ulusal AE, Akseki D. The BstUI and DpnII variants of the COL5A1 gene are associated with tennis elbow. American Journal of Sports Medicine. 2015;43(7):1784–9.

25. Erduran M, Altinisik J, Meric G, Ates O, Ulusal AE, Akseki D. Is Sp1 binding site polymorphism within COL1A1 gene associated with tennis elbowα. Gene [Internet]. 2014;537(2):308–11. Available from: http://dx.doi.org/10.1016/j.gene.2013.12.014

26. Alakhdar Mohmara Y, Cook J, Benítez-Martínez JC, McPeek ER, Aguilar AA, Olivas ES, et al. Influence of genetic factors in elbow tendon pathology: a case-control study. Sci Rep. 2020;10(1):1–7.

27. Khoschnau S, Melhus H, Jacobson A, Rahme H, Bengtsson H, Ribom E, et al. Type I Collagen α1 Sp1 Polymorphism and the Risk of Cruciate Ligament Ruptures or Shoulder Dislocations - 2008.pdf. Am J Sports Med. 2008 Dec;36(12):2432–6.

28. Petrillo S, Longo UG, Margiotti K, Candela V, Fusilli C, Rizzello G, et al. Genetic factors in rotator cuff pathology: Potential influence of col 5A1 polymorphism in outcomes of rotator cuff repair. BMC Med Genet. 2020;21(1):1–7.

29. Motta G da R, Amaral MV, Rezende E, Pitta R, dos Santos Vieira TC, Duarte MEL, et al. Evidence of genetic variations associated with rotator cuff disease. J Shoulder Elbow Surg. 2014;23(2):227–35.

30. Vidal Rodríguez S, del Castillo I. COL1A1, COL5A1 AND GDF5 POLYMORPHISMS ASSOCIATED WITH ROTATOR CUFF DISEASE. Orthopaedic Proceedings. 2018;99-B(Supp 8).

31. John R, Dhillon MS, Prabhakar S AAA. Col1a1 and Col12a1 With Acl Tears- a Association of Gene Polymorphisms in Col1a1 and Col12a1 With Acl Tears - a Study in the Indian. International Journal of Conference Proceedings. 2015;2015(112).

32. Roos TR, Roos AK, Avins AL, Ahmed MA, Kleimeyer JP, Fredericson M, et al. Genome-wide association study identifies a locus associated with rotator cuff injury. PLoS One. 2017;12(12):1–15.

33. Moses B, Orchard J, Orchard J. Systematic Review: Annual Incidence of ACL Injury and Surgery in Various Populations. Research in Sports Medicine. 2012 Jul;20(3–4):157–79.

34. Marieb EN, Hoehn K. Human Anatomy and Physiology. 9th Ed. London: Pearson Education, Inc.; 2015.

35. Del Buono A, Chan O, Maffulli N. Achilles tendon: Functional anatomy and novel emerging models of imaging classification. Int Orthop. 2013;37(4):715–21.

36. Zantop T, Petersen W, Sekiya JK, Musahl V, Fu FH. Anterior cruciate ligament anatomy and function relating to anatomical reconstruction. Knee Surgery, Sports Traumatology, Arthroscopy. 2006 Sep;14(10):982–92.

37. Rahim M, Collins M, September A. Genes and Musculoskeletal Soft-Tissue Injuries. Med Sport Sci. 2016;61:68–91.

38. Renstrom P, Ljungqvist A, Arendt E, Beynnon B, Fukubayashi T, Garrett W, et al. Non-contact ACL injuries in female athletes: an International Olympic Committee current concepts statement. Br J Sports Med [Internet]. 2008 Apr 7;42(6):394–412. Available from: https://bjsm.bmj.com/lookup/doi/10.1136/bjsm.2008.048934

39. Ljungqvist A, Schwellnus MP, Bachl N, Collins M, Cook J, Khan KMM, et al. International Olympic Committee Consensus Statement: Molecular Basis of Connective Tissue and Muscle Injuries in Sport. Clin Sports Med. 2008 Jan;27(1):231–9.

40. Cook JL, Purdam CR. Is tendon pathology a continuum? A pathology model to explain the clinical presentation of load-induced tendinopathy. Br J Sports Med. 2009 Jun;43(6):409–16.

41. Alentorn-Geli E, Myer GD, Silvers HJ, Samitier G, Romero D, Lázaro-Haro C, et al. Prevention of non-contact anterior cruciate ligament injuries in soccer players. Part 1: Mechanisms of injury and underlying risk factors. Knee Surgery, Sports Traumatology, Arthroscopy. 2009 Jul;17(7):705–29.

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

43. Sharma P, Maffulli N. Tendon injury and tendinopathy: Healing and repair. Journal of Bone and Joint Surgery - Series A. 2005;87(1):187–202.

44. Gwilym SE, Watkins B, Cooper CD, Harvie P, Auplish S, Pollard TCB, et al. Genetic influences in the progression of tears of the rotator cuff. Journal of Bone and Joint Surgery - Series B. 2009;91(7):915–7.

45. Lewis JS. Rotator cuff tendinopathy. Br J Sports Med. 2009 Apr;43(4):236–41.

46. Lewis JS. Rotator cuff tendinopathy: A model for the continuum of pathology and related management. Br J Sports Med. 2010;44(13):918–23.

47. Müller SA, Todorov A, Heisterbach PE, Martin I, Majewski M. Tendon healing: an overview of physiology, biology, and pathology of tendon healing and systematic review of state of the art in tendon bioengineering. Knee Surgery, Sports Traumatology, Arthroscopy. 2015;23(7):2097–105.

48. Chase KI, Caine DJ, Goodwin BJ, Whitehead JR, Romanick MA. A prospective study of injury affecting competitive collegiate swimmers. Research in Sports Medicine. 2013;21(2):111–23.

49. Pozzi F, Plummer HA, Shanley E, Thigpen CA, Bauer C, Wilson ML, et al. Preseason shoulder range of motion screening and in-season risk of shoulder and elbow injuries in overhead athletes: systematic review and meta-analysis. Br J Sports Med [Internet]. 2020 Sep;54(17):1019–27. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31937577

50. Tate A, Sarver J, DiPaola L, Yim J, Paul R, Thomas SJ. Changes in clinical measures and tissue adaptations in collegiate swimmers across a competitive season. J Shoulder Elbow Surg [Internet]. 2020;29(11):2375–84. Available from: https://doi.org/10.1016/j.jse.2020.03.028

51. Harrington S, Meisel C, Tate A. Cross-sectional study examining shoulder pain and disability in division I female swimmers. J Sport Rehabil. 2014 Feb;23(1):65–75.

52. McMaster WC. Shoulder injuries in competitive swimmers. Clin Sports Med. 1999;18(2):349–59.

53. McMaster, W.C., Troup, J.P, Arrendondo S. The incidence of shoulder problems in developing elite swimmers. J Swim Res. 1989;5(2):11–6.

54. Walker H, Gabbe B, Wajswelner H, Blanch P, Bennell K. Shoulder pain in swimmers: A 12-month prospective cohort study of incidence and risk factors. Physical Therapy in Sport. 2012 Nov;13(4):243–9.

55. Hill L, Collins M, Posthumus M. Risk factors for shoulder pain and injury in swimmers: A critical systematic review. Vol. 43, Physician and Sportsmedicine. Taylor and Francis Ltd.; 2015. p. 412–20.

56. Rupp S, Berninger K, Hopf T. Shoulder Problems in High Level Swimmers - Impingement, Anterior Instability, Muscular Imbalance? Int J Sports Med [Internet]. 1995 Nov 9;16(08):557–62. Available from: http://www.thieme-connect.de/DOI/DOI?10.1055/s-2007-973054

57. Bak K, Faunø P. Clinical Findings in Competitive Swimmers with Shoulder Pain. Am J Sports Med [Internet]. 1997 Mar 23;25(2):254–60. Available from: http://journals.sagepub.com/doi/10.1177/036354659702500221

58. Fredericson M, Ho C, Waite B, Jennings F, Peterson J, Williams C, et al. Magnetic Resonance Imaging Abnormalities in the Shoulder and Wrist Joints of Asymptomatic Elite Athletes. PM and R [Internet]. 2009;1(2):107–16. Available from: http://dx.doi.org/10.1016/j.pmrj.2008.09.004

59. Sein MLML, Walton J, Linklater J, Appleyard R, Kirkbride B, Kuah D, et al. Shoulder pain in elite swimmers: Primarily due to swim-volume-induced supraspinatus tendinopathy. Br J Sports Med. 2010 Feb;44(2):105–13.

60. Brophy RH, Schmitz L, Wright RW, Dunn WR, Parker RD, Andrish JT, et al. Return to Play and Future ACL Injury Risk After ACL Reconstruction in Soccer Athletes From the Multicenter Orthopaedic Outcomes Network (MOON) Group. Am J Sports Med. 2012 Nov;40(11):2517–22.

61. Beynnon BD, Johnson RJ, Abate JA, Fleming BC, Nichols CE. Treatment of anterior cruciate ligament injuries, part I. American Journal of Sports Medicine. 2005;33(10):1579–602.

62. Sallis RE, Jones K, Sunshine S, Smith G, Simon L. Comparing Sports Injuries in Men and Women. Int J Sports Med. 2001 Aug;22(6):420–3.

63. Saavedra MÁ, Navarro-Zarza JE, Villaseñor-Ovies P, Canoso JJ, Vargas A, Chiapas-Gasca K, et al. Clinical Anatomy of the Knee. Reumatol Clin. 2012 Dec;8(SUPPL.2):39–45.

64. Woo SLY, Abramowitch SD, Kilger R, Liang R. Biomechanics of knee ligaments: Injury, healing, and repair. J Biomech. 2006;39(1):1–20.

65. Abulhasan J, Grey M. Anatomy and Physiology of Knee Stability. J Funct Morphol Kinesiol. 2017 Sep;2(4):34.

66. Rees JD, Wilson AM, Wolman RL. Current concepts in the management of tendon disorders. Rheumatology. 2006 May;45(5):508–21.

67. De Loës M, Dahlstedt LJ, Thomée R. A 7-year study on risks and costs of knee injuries in male and female youth participants in 12 sports. Scand J Med Sci Sports. 2000;10(2):90–7.

68. Roos EM, Lohmander LS. Health and Quality of Life Outcomes The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis The Knee injury and Osteoarthritis Outcome Score (KOOS). Health Qual Life Outcomes [Internet]. 2003;1(64):1–8. Available from: https://hqlo.biomedcentral.com/articles/10.1186/1477-7525-1-64

69. Lohmander LS, Englund PM, Dahl LL, Roos EM. The Long-term Consequence of Anterior Cruciate Ligament and Meniscus Injuries. Am J Sports Med. 2007 Oct;35(10):1756–69.

70. Dare D, Rodeo S. Mechanisms of Post-traumatic Osteoarthritis After ACL Injury. Curr Rheumatol Rep. 2014;16(10):1–5.

71. September A V., Schwellnus MP, Collins M. Tendon and ligament injuries: The genetic component. Br J Sports Med. 2007;41(4):241–6.

72. Hauser R, Dolan E. Ligament Injury and Healing : An Overview. Journal of Prolotherapy. 2011;3(4):836–46.

73. Posthumus M, September AV V., O'cuinneagain D, Van Der Merwe W, Schwellnus MPP, Collins M, et al. The COL5A1 gene is associated with increased risk of Anterior cruciate ligament ruptures in female participants. American Journal of Sports Medicine. 2009 Nov;37(11):2234–40.

74. September AV V., Posthumus M, Van Der Merwe L, Schwellnus M, Noakes TDD, Collins M. The COL12A1 and COL14A1 genes and achilles tendon injuries. Int J Sports Med. 2008;29(3):257–63.

75. Saunders CJ, Van Der Merwe L, Posthumus M, Cook J, Handley CJ, Collins M, et al. Investigation of variants within the COL27A1 and TNC genes and Achilles tendinopathy in two populations. Journal of Orthopaedic Research. 2013;31(4):632–7.

76. Hay M, Patricios J, Collins R, Branfield A, Cook J, Handley CJ, et al. Association of type XI collagen genes with chronic Achilles tendinopathy in independent populations from South Africa and Australia. Br J Sports Med. 2013;47(9):569–74.

77. Posthumus M, September A V., Schwellnus MP, Collins M. Investigation of the Sp1-binding site polymorphism within the COL1A1 gene in participants with Achilles tendon injuries and controls. J Sci Med Sport. 2009;12(1):184–9.

78. Posthumus M, Collins M, Cook J, Handley CJ, Ribbans WJ, Smith RKW, et al. Components of the transforming growth factor- family and the pathogenesis of human

Achilles tendon pathology--a genetic association study. Rheumatology [Internet]. 2010 Nov 1;49(11):2090–7. Available from: https://academic.oup.com/rheumatology/article-lookup/doi/10.1093/rheumatology/keq072

79. El Khoury L, Posthumus M, Collins M, Handley CJ, Cook J, Raleigh SM. Polymorphic variation within the ADAMTS2, ADAMTS14, ADAMTS5, ADAM12 and TIMP2 genes and the risk of Achilles tendon pathology: A genetic association study. J Sci Med Sport [Internet]. 2013;16(6):493–8. Available from: http://dx.doi.org/10.1016/j.jsams.2013.02.006

80. Brown KL, Seale KB, El Khoury LY, Posthumus M, Ribbans WJ, Raleigh SM, et al. Polymorphisms within the COL5A1 gene and regulators of the extracellular matrix modify the risk of Achilles tendon pathology in a British case-control study. J Sports Sci. 2017;35(15):1475–83.

81. Nell EM, Van Der Merwe L, Cook J, Handley CJ, Collins M, September A V. The apoptosis pathway and the genetic predisposition to Achilles tendinopathy. Journal of Orthopaedic Research. 2012;30(11):1719–24.

82. September A V., Cook J, Handley CJ, Van Der Merwe L, Schwellnus MP, Collins M. Variants within the COL5A1 gene are associated with Achilles tendinopathy in two populations. Br J Sports Med. 2009;43(5):357–65.

83. Collins M, September A V., Posthumus M. Genetic Influences on Anterior Cruciate Ligament Injury [Internet]. Second Edi. The Anterior Cruciate Ligament: Reconstruction and Basic Science: Second Edition. Elsevier; 2018. 8-12.e1 p. Available from: http://dx.doi.org/10.1016/B978-0-323-38962-4.00002-3

84. Posthumus M, September A V., Keegan M, O'Cuinneagain D, Van Der Merwe W, Schwellnus MP, et al. Genetic risk factors for anterior cruciate ligament ruptures: COL1A1 gene variant. Br J Sports Med. 2009;43(5):352–6.

85. Brophy RH, Rai MF, Zhang Z, Torgomyan A, Sandell LJ. Molecular analysis of age and sex-related gene expression in meniscal tears with and without a concomitant anterior cruciate ligament tear. Journal of Bone and Joint Surgery - Series A. 2012;94(5):385–93.

86. Posthumus M, Collins M, van der Merwe L, O'Cuinneagain D, van der Merwe W, Ribbans WJ, et al. Matrix metalloproteinase genes on chromosome 11q22 and the risk of anterior cruciate ligament (ACL) rupture. Scand J Med Sci Sports. 2012;22(4):523–33.

87. Posthumus M, September AVA V., O'Cuinneagain D, Van Der Merwe W, Schwellnus MPMP, Collins M. The association between the COL12A1 gene and anterior cruciate ligament ruptures. Br J Sports Med. 2010 Dec;44(16):1160–5.

88. Ficek K, Stepien-Slodkowska M, Kaczmarczyk M, Maciejewska-Karlowska A, Sawczuk M, Cholewinski J, et al. Does the A9285G polymorphism in collagen type XII α1 gene associate with the risk of anterior cruciate ligament ruptures? Vol. 17, Balkan Journal of Medical Genetics. 2014. p. 41–6.

89. O'Connell K, Knight H, Ficek K, Leonska-Duniec A, Maciejewska-Karlowska A, Sawczuk M, et al. Interactions between collagen gene variants and risk of anterior cruciate ligament rupture. Eur J Sport Sci. 2015;15(4):341–50.

90. Ficek K, Cieszczyk P, Kaczmarczyk M, Maciejewska-Karłowska A, Sawczuk M, Cholewinski J, et al. Gene variants within the COL1A1 gene are associated with reduced anterior cruciate ligament injury in professional soccer players. J Sci Med Sport [Internet]. 2013;16(5):396–400. Available from: http://dx.doi.org/10.1016/j.jsams.2012.10.004

91. Stępień-Słodkowska M, Ficek K, Kaczmarczyk M, Maciejewska-Karłowska A, Sawczuk M, Leońska-Duniec A, et al. Stępień-Słodkowska et al. - The Variants Within the COL5A1 Gene are Associated with Reduced Risk of Anterior Cruciate Ligament Injury in.pdf. J Hum Kinet. 2015 Mar;45(1):103–11.

92. Thakkar D, Grant TM, Hakimi O, Carr AJ. Distribution and expression of type VI collagen and elastic fibers in human rotator cuff tendon tears. Connect Tissue Res. 2014;55(5–6):397–402.

93. Chaudhury S, Xia Z, Thakkar D, Hakimi O, Carr AJ. Gene expression profiles of changes underlying different-sized human rotator cuff tendon tears. J Shoulder Elbow Surg [Internet]. 2016;25(10):1561–70. Available from: http://dx.doi.org/10.1016/j.jse.2016.02.037

94. Burger M, de Wet H, Collins M. The COL5A1 gene is associated with increased risk of carpal tunnel syndrome. Clin Rheumatol. 2015;34(4):767–74.

95. Burger MC, de Wet H, Collins M. Interleukin and growth factor gene variants and risk of carpal tunnel syndrome. Gene [Internet]. 2015;564(1):67–72. Available from: http://dx.doi.org/10.1016/j.gene.2015.03.047

96. Burger MC, De Wet H, Collins M. The BGN and ACAN genes and carpal tunnel syndrome. Gene [Internet]. 2014;551(2):160–6. Available from: http://dx.doi.org/10.1016/j.gene.2014.08.051

97. Burger MC, De Wet H, Collins M. Matrix metalloproteinase genes on chromosome 11q22 and risk of carpal tunnel syndrome. Rheumatol Int. 2016;36(3):413–9.

98. Dada S, Burger MC, Massij F, de Wet H, Collins M. Carpal tunnel syndrome: The role of collagen gene variants. Gene [Internet]. 2016;587(1):53–8. Available from: http://dx.doi.org/10.1016/j.gene.2016.04.030

99. O'Connell K, Saunders CJ, Collins M. Collagen gene sequence variants in exercise-related traits. Central European Journal of Sports Sciences and Medicine. 2013;1(1):3–17.

100. Di Lullo GA, Sweeney SM, Körkkö J, Ala-Kokko L, San Antonio JD. Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen. Journal of Biological Chemistry. 2002;277(6):4223–31.

101. Ricard-Blum S, Ruggiero F. The collagen superfamily: From the extracellular matrix to the cell membrane. Pathologie Biologie [Internet]. 2005 Sep;53(7):430–42. Available from: https://linkinghub.elsevier.com/retrieve/pii/S036981140400313X

102. Ricard-Blum S. The Collagen Family. Cold Spring Harb Perspect Biol. 2011;3(1):1–19.

103. Wenstrup RJ, Smith SM, Florer JB, Zhang G, Beason DP, Seegmiller RE, et al. Regulation of collagen fibril nucleation and initial fibril assembly involves coordinate interactions with collagens V and XI in developing tendon. Journal of Biological Chemistry. 2011;286(23):20455–65.

104. Huegel J, Williams AA, Soslowsky LJ. Rotator Cuff Biology and Biomechanics: A Review of Normal and Pathological Conditions. Curr Rheumatol Rep. 2014;17(1):1–9.

105. Ellenbecker T. Clinical Examination of the Shoulder. Clinical Examination of the Shoulder. 2004;24(1):1–210.

106. Labriola JE, Lee TQ, Debski RE, McMahon PJ. Stability and instability of the glenohumeral joint: The role of shoulder muscles. J Shoulder Elbow Surg. 2005 Jan;14(1):S32–8.

107. Michener LA, McClure PW, Karduna AR. Anatomical and biomechanical mechanisms of subacromial impingement syndrome. Clinical Biomechanics. 2003;18(5):369–79.

108. Pinkas D, Wiater JM. Functional Anatomy of the Shoulder? Orthopaedic Physical Therapy Secrets: Third Edition. 2017;18(1):318–26.

109. Landau JP, Hoenecke HR. Genetic and biomechanical determinants of glenoid version: Implications for glenoid implant placement in shoulder arthroplasty. J Shoulder Elbow Surg. 2009 Jul;18(4):661–7.

110. Beaudreuil J, Nizard R, Thomas T, Peyre M, Liotard JP, Boileau P, et al. Contribution of clinical tests to the diagnosis of rotator cuff disease: A systematic literature review. Joint Bone Spine. 2009;76(1):15–9.

111. Blaine TAA, Kim YSYS, Voloshin I, Chen D, Murakami K, Chang SSSS, et al. The molecular pathophysiology of subacromial bursitis in rotator cuff disease. J Shoulder Elbow Surg. 2005 Jan;14(1):S84–9.

112. Nourissat G, Berenbaum F, Duprez D. Tendon injury: From biology to tendon repair. Nat Rev Rheumatol [Internet]. 2015;11(4):223–33. Available from: http://dx.doi.org/10.1038/nrrheum.2015.26

113. De Maeseneer M, Van Roy P, Shahabpour M. Normal MR Imaging Anatomy of the Rotator Cuff Tendons, Glenoid Fossa, Labrum, and Ligaments of the Shoulder. Radiol Clin North Am. 2006;44(4):479–87.

114. Cook TS, Stein JM, Simonson S, Kim W. Normal and Variant Anatomy of the Shoulder on MRI. Magn Reson Imaging Clin N Am. 2011 Aug;19(3):581–94.

115. Bennett WF. Subscapularis, medial, and lateral head coracohumeral ligament insertion anatomy. Arthroscopy: The Journal of Arthroscopic & Related Surgery. 2001 Feb;17(2):173–80.

116. Dunham KS, Bencardino JT, Rokito AS. Anatomic Variants and Pitfalls of the Labrum, Glenoid Cartilage, and Glenohumeral Ligaments. Magn Reson Imaging Clin N Am. 2012;20(2):213–28.

117. Quental C, Folgado J, Ambrósio J, Monteiro J. A multibody biomechanical model of the upper limb including the shoulder girdle. Multibody Syst Dyn. 2012 Aug;28(1–2):83–108.

118. Favre P, Senteler M, Hipp J, Scherrer S, Gerber C, Snedeker JG. An integrated model of active glenohumeral stability. J Biomech. 2012 Aug;45(13):2248–55.

119. Mochizuki T, Sugaya H, Uomizu M, Maeda K, Matsuki K, Sekiya I, et al. Humeral Insertion of the Supraspinatus and Infraspinatus. The Journal of Bone and Joint Surgery-American Volume. 2008 May;90(5):962–9.

120. Prescher A. Anatomical basics, variations, and degenerative changes of the shoulder joint and shoulder girdle. Eur J Radiol. 2000;35(2):88–102.

121. Reinold MM, Escamilla R, Wilk KE. Current Concepts in the Scientific and Clinical Rationale Behind Exercises for Glenohumeral and Scapulothoracic Musculature. Journal of Orthopaedic & Sports Physical Therapy. 2009 Feb;39(2):105–17.

122. Hurov J. Anatomy and Mechanics of the Shoulder: Review of Current Concepts. Journal of Hand Therapy [Internet]. 2009;22(4):328–43. Available from: http://dx.doi.org/10.1016/j.jht.2009.05.002

123. Clark JM, Harryman DT. Tendons, ligaments, and capsule of the rotator cuff. Gross and microscopic anatomy. J Bone Joint Surg Am. 1992 Jun;74(5):713–25.

124. Jobe C, Phipatanakul W, Coen M. Gross Anatomy of the Shoulder. In: C. Rockwood & FAM, editor. The Shoulder. 4th Ed. Philadelphia, PA:; 2009. p. 33–100.

125. Petchprapa CN, Beltran LS, Jazrawi LM, Kwon YW, Babb JS, Recht MP. The Rotator Interval: A Review of Anatomy, Function, and Normal and Abnormal MRI Appearance. American Journal of Roentgenology. 2010 Sep;195(3):567–76.

126. Fallon J, Blevins FT, Vogel K, Trotter J. Functional morphology of the supraspinatus tendon. Journal of Orthopaedic Research. 2002;20(5):920–6.

127. Peat M. Functional anatomy of the shoulder complex. Phys Ther. 1986;66(12):1855–65.

128. Colas F, Nevoux J, Gagey O. The subscapular and subcoracoid bursae: descriptive and functional anatomy. J Shoulder Elbow Surg. 2004 Jul;13(4):454–8.

129. Nakajima T, Rokuuma N, Hamada K, Tomatsu T, Fukuda H. Histologic and biomechanical characteristics of the supraspinatus tendon: Reference to rotator cuff tearing. J Shoulder Elbow Surg. 1994 Mar;3(2):79–87.

130. Itoi E, Berglund LJ, Grabowski JJ, Schultz FM, Growney ES, Morrey BF, et al. Tensile properties of the supraspinatus tendon. Journal of Orthopaedic Research. 1995 Jul;13(4):578–84.

131. Perry SM, Getz CL, Soslowsky LJ. After rotator cuff tears, the remaining (intact) tendons are mechanically altered. J Shoulder Elbow Surg. 2009 Jan;18(1):52–7.

132. Bigliani LU, Kelkar R, Flatow EL, Pollock RG, Mow VC. Glenohumeral stability. Biomechanical properties of passive and active stabilizers. Clin Orthop Relat Res. 1996 Sep;(330):13–30.

133. Reinold MM, Gill TJ. Current Concepts in the Evaluation and Treatment of the Shoulder in Overhead-Throwing Athletes, Part 1: Physical Characteristics and Clinical Examination. Sports Health: A Multidisciplinary Approach. 2010 Jan;2(1):39–50.

134. Lewis J. Rotator cuff related shoulder pain: Assessment, management and uncertainties. Man Ther [Internet]. 2015;23:57–68. Available from: http://dx.doi.org/10.1016/j.math.2016.03.009

135. Magee D. Orthopedic Physical Assessment. 6th Ed. St. Louis, Missouri: Saunders Elsevier; 2014.

136. Al-Hakim W, Noorani A, Lambert S. Assessment and treatment strategies for rotator cuff tears. Shoulder Elbow. 2015 Apr;7(2):76–84.

137. Culham E, Peat M. Functional Anatomy of the Shoulder Complex. Physical Therapy Journal. 1993;18(1):342–51.

138. Magnusson SP, Hansen M, Langberg H, Miller B, Haraldsson B, Kjoeller Westh E, et al. The adaptability of tendon to loading differs in men and women. Int J Exp Pathol. 2007 Aug;88(4):237–40.

139. Tashjian RZ, Farnham JM, Granger EK, Teerlink CC, Cannon-Albright LA. Evidence for an Environmental and Inherited Predisposition Contributing to the Risk for Global Tendinopathies or Compression Neuropathies in Patients With Rotator Cuff Tears. Orthop J Sports Med. 2016;4(4):1–7.

140. Teerlink CC, Cannon-Albright LA, Tashjian RZ. Significant association of full-thickness rotator cuff tears and estrogen-related receptor-β (ESRRB). J Shoulder Elbow Surg. 2015 Feb;24(2):e31–5.

141. Tashjian RZ, Farnham JM, Albright FS, Teerlink CC, Cannon-Albright LA. Evidence for an inherited predisposition contributing to the risk for rotator cuff disease. Journal of Bone and Joint Surgery - Series A. 2009;91(5):1136–42.

142. Wright R, Baumgarten K. Shoulder Outcomes Measures. Journal of the American Academy of Orthopaedic Surgeons. 2010;18(7):436–44.

143. Baumgarten KM, Gerlach D, Galatz LM, Teefey SA, Middleton WD, Ditsios K, et al. Cigarette Smoking Increases the Risk for Rotator Cuff Tears. Clin Orthop Relat Res. 2010 Jun;468(6):1534–41.

144. Dean E, Söderlund A. What is the role of lifestyle behaviour change associated with non-communicable disease risk in managing musculoskeletal health conditions with special reference to chronic pain? BMC Musculoskelet Disord. 2015 Dec;16(1):87.

145. Lewis J, Mccreesh K, Roy JS, Ginn K. Rotator cuff tendinopathy: Navigating the diagnosis-management conundrum. Journal of Orthopaedic and Sports Physical Therapy. 2015;45(11):923–37.

146. Magnusson SP, Langberg H, Kjaer M. The pathogenesis of tendinopathy: Balancing the response to loading. Nat Rev Rheumatol. 2010;6(5):262–8.

147. Shindle MK, Chen CCT, Robertson C, DiTullio AE, Paulus MC, Clinton CM, et al. Full-thickness supraspinatus tears are associated with more synovial inflammation and tissue degeneration than partial-thickness tears. J Shoulder Elbow Surg. 2011 Sep;20(6):917–27.

148. Lundgreen K, Lian Ø, Scott A, Engebretsen L. Increased levels of apoptosis and p53 in partial-thickness supraspinatus tendon tears. Knee Surgery, Sports Traumatology, Arthroscopy. 2013 Jul;21(7):1636–41.

149. Neer CS. Anterior acromioplasty for the chronic impingement syndrome in the shoulder: a preliminary report. J Bone Joint Surg Am. 1972 Jan;54(1):41–50.

150. Neer CS. Impingement lesions. Clin Orthop Relat Res. 1983 Mar;173(173):70–7.

151. Lewis JS. Rotator cuff tendinopathy/subacromial impingement syndrome: Is it time for a new method of assessment? Br J Sports Med. 2009;43(4):259–64.

152. Hyvönen P, Melkko J, Lehto VP, Jalovaara P. Involvement of the subacromial bursa in impingement syndrome of the shoulder as judged by expression of tenascin-C and histopathology. J Bone Joint Surg Br. 2003 Mar;85-B(2):299–305.

153. Payne LZ, Deng XHH, Craig E V., Torzilli PA, Warren RF. The combined dynamic and static contributions to subacromial impingement. A biomechanical analysis. American Journal of Sports Medicine. 1997 Nov;25(6):801–8.

154. Fukuda H, Hamada K, Nakajima T, Yamada N, Tomonaga A, Goto M. Partial-thickness tears of the rotator cuff. Int Orthop [Internet]. 1996 Jul 29;20(4):257–65. Available from: http://link.springer.com/10.1007/s002640050075

155. McClincy MP, Rodosky MW. Arthroscopic Subacromial Decompression. Oper Tech Orthop. 2015;25(1):10–4.

156. Ogilvie-Harris DJJ, Demaziere A, Demazière A. Arthroscopic debridement versus open repair for rotator cuff tears. A prospective cohort study. J Bone Joint Surg Br. 1993 May;75(3):416–20.

157. Hashimoto T, Nobuhara K, Hamada T. Pathologic Evidence of Degeneration as a Primary Cause of Rotator Cuff Tear. Clin Orthop Relat Res. 2003 Oct;415:111–20.

158. Bey MJ, Brock SK, Beierwaltes WN, Zauel R, Kolowich PA, Lock TR. In vivo measurement of subacromial space width during shoulder elevation: Technique and preliminary results in patients following unilateral rotator cuff repair. Clinical Biomechanics. 2007 Aug;22(7):767–73.

159. Bey MJ, Ramsey ML, Soslowsky LJ. Intratendinous strain fields of the supraspinatus tendon: Effect of a surgically created articular-surface rotator cuff tear. J Shoulder Elbow Surg. 2002 Nov;11(6):562–9.

160. Henkus HE, de Witte PB, Nelissen RGHH, Brand R, van Arkel ERA. Bursectomy compared with acromioplasty in the management of subacromial impingement syndrome. J Bone Joint Surg Br. 2009 Apr;91-B(4):504–10.

161. Vincent K, Leboeuf-Yde C, Gagey O. Are degenerative rotator cuff disorders a cause of shoulder pain? Comparison of prevalence of degenerative rotator cuff disease to prevalence of nontraumatic shoulder pain through three systematic and critical reviews. J Shoulder Elbow Surg. 2017 May;26(5):766–73.

162. Yuan J, Murrell GAC, Wei AQ, Wang MX. Apoptosis in rotator cuff tendonopathy. Journal of Orthopaedic Research. 2002 Nov;20(6):1372–9.

163. Benson RT, McDonnell SM, Knowles HJ, Rees JL, Carr AJ, Hulley PA. Tendinopathy and tears of the rotator cuff are associated with hypoxia and apoptosis. J Bone Joint Surg Br. 2010 Mar;92-B(3):448–53.

164. Tempelhof S, Rupp S, Seil R. Age-related prevalence of rotator cuff tears in asymptomatic shoulders. J Shoulder Elbow Surg. 1999 Jul;8(4):296–9.

165. Sher JS, Uribe JW, Posada A, Murphy BJ, Zlatkin MB. Abnormal findings on magnetic resonance images of asymptomatic shoulders. J Bone Joint Surg. 1995 Jan;77(1):10–5.

166. Dabija DI, Gao C, Edwards TL, Kuhn JE, Jain NB. Genetic and familial predisposition to rotator cuff disease: a systematic review. J Shoulder Elbow Surg [Internet]. 2017;26(6):1103–12. Available from: http://dx.doi.org/10.1016/j.jse.2016.11.038

167. Milgrom C, Schaffler M, Gilbert S, Van Holsbeeck M. Rotator-cuff changes in asymptomatic adults. The effect of age, hand dominance and gender. J Bone Joint Surg Br. 1995;77-B(2).

168. Abate M, Schiavone C, Di Carlo L, Salini V. Prevalence of and risk factors for asymptomatic rotator cuff tears in postmenopausal women. Menopause. 2014 Mar;21(3):275–80.

169. Minagawa H, Yamamoto N, Abe H, Fukuda M, Seki N, Kikuchi K, et al. Prevalence of symptomatic and asymptomatic rotator cuff tears in the general population: From mass-

screening in one village. J Orthop [Internet]. 2013;10(1):8–12. Available from: http://dx.doi.org/10.1016/j.jor.2013.01.008

170. Yamaguchi K, Ditsios K, Middleton WDWD, Hildebolt CFCF, Galatz LMLM, Teefey SASA. The demographic and morphological features of rotator cuff disease: A comparison of asymptomatic and symptomatic shoulders. Journal of Bone and Joint Surgery - Series A. 2006 Jan;88(8):1699–704.

171. Abboud JA, Kim JS. The effect of hypercholesterolemia on rotator cuff disease. Clin Orthop Relat Res. 2010;468(6):1493–7.

172. Cutts S, Prempeh M, Drew S. Anterior Shoulder Dislocation. The Annals of The Royal College of Surgeons of England. 2009 Jan;91(1):2–7.

173. Yamamoto A, Takagishi K, Osawa T, Yanagawa T, Nakajima D, Shitara H, et al. Prevalence and risk factors of a rotator cuff tear in the general population. J Shoulder Elbow Surg [Internet]. 2010;19(1):116–20. Available from: http://dx.doi.org/10.1016/j.jse.2009.04.006

174. Belangero PS, Leal MF, Figueiredo EA, Cohen C, Pochini ADC, Smith MC, et al. Gene expression analysis in patients with traumatic anterior shoulder instability suggests deregulation of collagen genes. Journal of Orthopaedic Research. 2014;32(10):1311–6.

175. Zaid MB, Young NM, Pedoia V, Feeley BT, Ma CB, Lansdown DA. Anatomic shoulder parameters and their relationship to the presence of degenerative rotator cuff tears and glenohumeral osteoarthritis: a systematic review and meta-analysis. J Shoulder Elbow Surg [Internet]. 2019;28(12):2457–66. Available from: https://doi.org/10.1016/j.jse.2019.05.008

176. Allen H, Chan BY, Davis KW, Blankenbaker DG. Overuse Injuries of the Shoulder. Radiol Clin North Am [Internet]. 2019;57(5):897–909. Available from: https://doi.org/10.1016/j.rcl.2019.03.003

177. Longo UG, Candela V, Berton A, Salvatore G, Guarnieri A, Deangelis J, et al. Genetic basis of rotator cuff injury: A systematic review. BMC Med Genet. 2019;20(1):1–6.

178. Tashjian R, Granger E, Hung M, Saltzman E. Incidence of familial tendon dysfunction in patients with full-thickness rotator cuff tears. Open Access J Sports Med. 2014 May;137.

179. Tashjian RZ, Granger EK, Farnham JM, Cannon-Albright LA, Teerlink CC. Genome-wide association study for rotator cuff tears identifies two significant single-nucleotide polymorphisms. J Shoulder Elbow Surg [Internet]. 2016;25(2):174–9. Available from: http://dx.doi.org/10.1016/j.jse.2015.07.005

180. Tashjian RZZ, Granger EKK, Zhang Y, Teerlink CCC, Cannon-Albright LAA. Identification of a genetic variant associated with rotator cuff repair healing. J Shoulder Elbow Surg [Internet]. 2016;25(6):865–72. Available from: http://dx.doi.org/10.1016/j.jse.2016.02.019

181. Harvie P, Ostlere SJ, Teh J, McNally EG, Clipsham K, Burston BJ, et al. Genetic influences in the aetiology of tears of the rotator cuff. J Bone Joint Surg Br [Internet]. 2004 Jul;86-B(5):696–700. Available from: http://online.boneandjoint.org.uk/doi/10.1302/0301-620X.86B5.14747

182. Ling Y, Peng C, Liu C, Zhang N, Yue S. Gene polymorphism of IL-6 and MMP-3 decreases passive range of motion after rotator cuff repair. Int J Clin Exp Pathol. 2015;8(5):5709–14.

183. Assunção JH, Godoy-Santos AL, dos Santos MCLG, Malavolta EA, Gracitelli MEC, Ferreira Neto AA. Matrix Metalloproteases 1 and 3 Promoter Gene Polymorphism Is Associated With Rotator Cuff Tear. Clin Orthop Relat Res. 2017;475(7):1904–10.

184. Bonato LL, Quinelato V, Pinheiro ADR, Amaral MVG, De Souza FN, Lobo JC, et al. ESRRB polymorphisms are associated with comorbidity of temporomandibular disorders and rotator cuff disease. Int J Oral Maxillofac Surg. 2016;45(3):323–31.

185. Kim S, Kleimeyer JP, Ahmed MA, Avins AL, Fredericson M, Dragoo JL, et al. A Genetic Marker Associated with Shoulder Dislocation. Int J Sports Med. 2017;38(7):508–14.

186. Kluger R, Huber KR, Seely PG, Berger CE, Frommlet F. Novel Tenascin-C Haplotype Modifies the Risk for a Failure to Heal After Rotator Cuff Repair. American Journal of Sports Medicine. 2017;45(13):2955–64.

187. Longo UG, Margiotti K, Petrillo S, Rizzello G, Fusilli C, Maffulli N, et al. Genetics of rotator cuff tears: No association of col5a1 gene in a case-control study. BMC Med Genet. 2018;19(1):1–5.

188. Leal MF, Belangero PS, Cohen C, Figueiredo EA, Loyola LC, Pochini AC, et al. Identification of suitable reference genes for gene expression studies of shoulder instability. PLoS One. 2014;9(8).

189. Cohen C, Leal MF, Belangero PS, Figueiredo EA, Smith MC, Andreoli CV, et al. The roles of tenascin C and fibronectin 1 in adhesive capsulitis: A pilot gene expression study. Clinics. 2016;71(6):325–31.

190. Figueiredo EA, Loyola LC, Belangero PS, Campos Ribeiro-dos-Santos ÂK, Emanuel Batista Santos S, Cohen C, et al. Rotator Cuff Tear Susceptibility Is Associated With Variants in Genes Involved in Tendon Extracellular Matrix Homeostasis. Journal of Orthopaedic Research. 2020;38(1):192–201.

191. Peach CA, Zhang Y, Dunford JE, Brown MA, Carr AJ. Cuff tear arthropathy: Evidence of functional variation in pyrophosphate metabolism genes. Clin Orthop Relat Res. 2007 Sep;462(462):67–72.

192. Zanin V, Segat L, Bianco A, Padovan L, Tavares N, Crovella S. DEFB1 gene 5′ untranslated region (UTR) polymorphisms in inflammatory bowel diseases. Clinics. 2012 Apr;67(4):395–8.

193. Esparza R, Gortazar AR, Forriol F. Cell study of the three areas of the meniscus: Effect of growth factors in an experimental model in sheep. Journal of Orthopaedic Research. 2012 Oct;30(10):1647–51.

194. Burgess JT, Bolderson E, Adams MN, Baird AM, Zhang SD, Gately KA, et al. Activation and cleavage of SASH1 by caspase-3 mediates an apoptotic response. Cell Death Dis. 2016 Nov;7(11):e2469–e2469.

195. Osti L, Buda M, Del Buono A, Osti R, Massari L, Maffulli N. Apoptosis and rotator cuff tears: Scientific evidence from basic science to clinical findings. Br Med Bull. 2017;122(1):123–33.

196. Geiger E V., Henrich D, Wutzler S, Schneidmüller D, Jakob H, Frank JM, et al. The role of TNXB single-nucleotide polymorphisms in recurrent shoulder dislocation. Journal of Orthopaedic Research. 2013;31(2):295–9.

197. Egging D, van den Berkmortel F, Taylor G, Bristow J, Schalkwijk J. Interactions of human tenascin-X domains with dermal extracellular matrix molecules. Arch Dermatol Res. 2007 Jan;298(8):389–96.

198. Minamitani T, Ikuta T, Saito Y, Takebe G, Sato M, Sawa H, et al. Modulation of collagen fibrillogenesis by tenascin-X and type VI collagen. Exp Cell Res. 2004 Aug;298(1):305–15.

199. Johnson J, Sim F, Scott S. Musculoskeletal Injuries in Competitive Swimmers. Mayo Clin Proc. 1987 Apr;62(4):289–304.

200. Richardson AB, Jobe FW, Collins HR. The shoulder in competitive swimming. Am J Sports Med. 1980;8(3):159–63.

201. van Houwelingen J, Schreven S, Smeets JBJ, Clercx HJH, Beek PJ. Effective Propulsion in Swimming: Grasping the Hydrodynamics of Hand and Arm Movements. J Appl Biomech. 2017 Feb;33(1):87–100.

202. Borsa PA, Scibek JS, Jacobson JA, Meister K. Sonographic stress measurement of glenohumeral joint laxity in collegiate swimmers and age-matched controls. American Journal of Sports Medicine. 2005;33(7):1077–84.

203. Kennedy, J.C., Hawkins RJ. Swimmers Shoulder. Phys Sportsmed. 1974;2(4):34–8.

204. Heinlein SA, Cosgarea AJ. Biomechanical considerations in the competitive swimmer's shoulder. Sports Health. 2010;2(6):519–25.

205. Weldon EJ, Richardson AB. Upper extremity overuse injuries in swimming. A discussion of swimmer's shoulder. Clin Sports Med. 2001 Jul;20(3):423–38.

206. Brushøj C, Bak K, Johannsen H V., Faunø P. Swimmers' painful shoulder arthroscopic findings and return rate to sports. Scand J Med Sci Sports. 2007;17(4):373–7.

207. Flandry F, Hommel G. Normal anatomy and biomechanics of the knee. Sports Med Arthrosc Rev. 2011;19(2):82–92.

208. Neumann DA. Knee. In: Neumann DA, editor. Kinesiology of the Musculoskeletal System: Foundation for Rehabilitation. St. Louis: Mosby Elsevier; 2010. p. 523.

209. Deehan DJ, Cawston TE. The biology of intergration of the anterior cruciate ligament. Journal of Bone and Joint Surgery - Series B. 2005;87(7):889–95.

210. Gimber LH, Scalcione LR, Rowan A, Hardy JC, Melville DM, Taljanovic MS. Multiligamentous injuries and knee dislocations. Skeletal Radiol. 2015 Nov;44(11):1559–72.

211. Petersen W, Zantop T. Anatomy of the anterior cruciate ligament with regard to its two bundles. Clin Orthop Relat Res. 2007;(454):35–47.

212. Siegel L, Vandenakker-Albanese C, Siegel D. Anterior cruciate ligament injuries: Anatomy, physiology, biomechanics, and management. Clinical Journal of Sport Medicine. 2012;22(4):349–55.

213. Kalawadia J V., Guenther D, Irarrázaval S, Fu FH. Anatomy and Biomechanics of the Anterior Cruciate Ligament [Internet]. Second Edi. The Anterior Cruciate Ligament: Reconstruction and Basic Science: Second Edition. Elsevier; 2018. 1-7.e2 p. Available from: http://dx.doi.org/10.1016/B978-0-323-38962-4.00001-1

214. Ménétrey J, Duthon VB, Laumonier T, Fritschy D. "biological failure" of the anterior cruciate ligament graft. Knee Surgery, Sports Traumatology, Arthroscopy. 2008;16(3):224–31.

215. Duthon VB, Barea C, Abrassart S, Fasel JH, Fritschy D, Ménétrey J. Anatomy of the anterior cruciate ligament. Knee Surgery, Sports Traumatology, Arthroscopy. 2006;14(3):204–13.

216. Provenzano PP, Vanderby R. Collagen fibril morphology and organization: Implications for force transmission in ligament and tendon. Matrix Biology. 2006;25(2):71–84.

217. Provenzano P, Lakes R, Keenan T, Vanderby R. Nonlinear ligament viscoelasticity. Ann Biomed Eng. 2001;29(10):908–14.

218. Beynnon BD, Vacek PM, Newell MK, Tourville TW, Smith HC, Shultz SJ, et al. The Effects of Level of Competition, Sport, and Sex on the Incidence of First-Time Noncontact Anterior Cruciate Ligament Injury. Am J Sports Med. 2014 Aug;42(8):1806–12.

219. Koga H, Nakamae A, Shima Y, Iwasa J, Myklebust G, Engebretsen L, et al. Mechanisms for Noncontact Anterior Cruciate Ligament Injuries. Am J Sports Med. 2010 Nov;38(11):2218–25.

220. Griffin LY, Albohm MJ, Arendt EA, Bahr R, Beynnon BD, DeMaio M, et al. Understanding and Preventing Noncontact Anterior Cruciate Ligament Injuries. Am J Sports Med. 2006 Sep;34(9):1512–32.

221. McConkey JP. Anterior cruciate ligament rupture in skiing. Am J Sports Med. 1986 Mar;14(2):160–4.

222. Hewett TE, Myer GD, Ford KR, Paterno M V., Quatman CE. Mechanisms, prediction, and prevention of ACL injuries: Cut risk with three sharpened and validated tools. Journal of Orthopaedic Research. 2016;34(11):1843–55.

223. Vacek PM, Slauterbeck JR, Tourville TW, Sturnick DR, Holterman LA, Smith HC, et al. Multivariate Analysis of the Risk Factors for First-Time Noncontact ACL Injury in High School and College Athletes. Am J Sports Med. 2016;44(6):1492–501.

224. Shah F, Usman Riaz M, Hassan D, Abbas Z. Factors contributing to anterior cruciate ligament injury and pattern of presentations. Rawal Medical Journal. 2017;42(2):223–5.

225. Prodromos CC, Han Y, Rogowski J, Joyce B, Shi K. A Meta-analysis of the Incidence of Anterior Cruciate Ligament Tears as a Function of Gender, Sport, and a Knee Injury–Reduction Regimen. Arthroscopy: The Journal of Arthroscopic & Related Surgery. 2007 Dec;23(12):1320-1325.e6.

226. Hertel J, Dorfman JH, Braham RA. Lower extremity malalignments and anterior cruciate ligament injury history. J Sports Sci Med. 2004 Dec;3(4):220–5.

227. Everhart JS, Flanigan DC, Simon RA, Chaudhari AMW. Association of Noncontact Anterior Cruciate Ligament Injury With Presence and Thickness of a Bony Ridge on the Anteromedial Aspect of the Femoral Intercondylar Notch. Am J Sports Med. 2010 Aug;38(8):1667–73.

228. Shen L, Jin ZG, Dong QR, Li LB. Anatomical Risk Factors of Anterior Cruciate Ligament Injury. Chin Med J (Engl). 2018 Dec;131(24):2960–7.

229. Ireland ML. Anterior cruciate ligament injury in female athletes: epidemiology. J Athl Train. 1999 Apr;34(2):150–4.

230. Hewett TE, Myer GD, Ford KR, Heidt RS, Colosimo AJ, McLean SG, et al. Biomechanical Measures of Neuromuscular Control and Valgus Loading of the Knee Predict Anterior Cruciate Ligament Injury Risk in Female Athletes: A Prospective Study. Am J Sports Med. 2005 Apr;33(4):492–501.

231. Hewett TE, Myer GD, Ford KR. Anterior Cruciate Ligament Injuries in Female Athletes. Am J Sports Med [Internet]. 2006 Feb 30;34(2):299–311. Available from: http://journals.sagepub.com/doi/10.1177/0363546505284183

232. Hewett TE, Zazulak BT, Myer GD. Effects of the Menstrual Cycle on Anterior Cruciate Ligament Injury Risk. Am J Sports Med. 2007 Apr;35(4):659–68.

233. Smith HC, Vacek P, Johnson RJ, Slauterbeck JR, Hashemi J, Shultz S, et al. Risk Factors for Anterior Cruciate Ligament Injury. Sports Health: A Multidisciplinary Approach. 2012 Jan;4(1):69–78.

234. Adachi N, Nawata K, Maeta M, Kurozawa Y. Relationship of the menstrual cycle phase to anterior cruciate ligament injuries in teenaged female athletes. Arch Orthop Trauma Surg. 2008 May;128(5):473–8.

235. Kim SK, Roos TR, Roos AK, Kleimeyer JP, Ahmed MA, Goodlin GT, et al. Genome-wide association screens for achilles tendon and ACL tears and tendinopathy. PLoS One. 2017;12(3):1–16.

236. Yao C, Zhu B, Liu Y, Xu X, Yang X, Xu Z, et al. Association of MMP3 COL5A1 and COL1A1 gene polymorphisms with ACL geometric size in Chinese population. Int J Clin Exp Pathol. 2017;10(2):1991–7.

237. Flynn RK, Pedersen CL, Birmingham TB, Kirkley A, Jackowski D, Fowler PJ. The familial predisposition toward tearing the anterior cruciate ligament: A case control study. American Journal of Sports Medicine. 2005 Jan;33(1):23–8.

238. Mannion S, Mtintsilana A, Posthumus M, Van Der Merwe W, Hobbs H, Collins M, et al. Genes encoding proteoglycans are associated with the risk of anterior cruciate ligament ruptures. Br J Sports Med. 2014;48(22):1640–6.

239. Rahim M, Gibbon A, Hobbs H, Van Der Merwe W, Posthumus M, Collins M, et al. The association of genes involved in the angiogenesis-associated signaling pathway with risk of anterior cruciate ligament rupture. Journal of Orthopaedic Research. 2014;32(12):1612–8.

240. Feldmann D, Rahim M, Suijkerbuijk M, Laguette M, Cieszczyk P, Ficek K, et al. Investigation of multiple populations highlight VEGFA polymorphisms to modulate anterior cruciate ligament injury. Journal of Orthopaedic Research. 2021;In Press.

241. Simon RAA, Everhart JSS, Nagaraja HNN, Chaudhari AMM. A case-control study of anterior cruciate ligament volume, tibial plateau slopes and intercondylar notch dimensions in ACL-injured knees. J Biomech. 2010 Jun;43(9):1702–7.

242. Beckett ME, Massie DL, Bowers KD, Stoll DA. Incidence of Hyperpronation in the ACL Injured Knee: A Clinical Perspective. J Athl Train. 1992;27(1):58–62.

243. van Mechelen W, Hlobil H, Kemper HCG. Incidence, Severity, Aetiology and Prevention of Sports Injuries. Sports Medicine. 1992 Aug;14(2):82–99.

244. Woodford-Rogers B, Cyphert L, Denegar CR. Risk factors for anterior cruciate ligament injury in high school and college athletes. J Athl Train. 1994 Dec;29(4):343–6.

245. Uhorchak JM, Scoville CR, Williams GN, Arciero RA, Pierre P St., Taylor DC. Risk Factors Associated with Noncontact Injury of the Anterior Cruciate Ligament. Am J Sports Med. 2003 Nov;31(6):831–42.

246. Kirkendall DT, Garrett WEJ. The anterior cruciate ligament enigma. Injury mechanisms and prevention. Clin Orthop Relat Res. 2000 Mar;(372):64–8.

247. Krosshaug T, Nakamae A, Boden BP, Engebretsen L, Smith G, Slauterbeck JR, et al. Mechanisms of Anterior Cruciate Ligament Injury in Basketball. Am J Sports Med. 2007 Mar;35(3):359–67.

248. Collins M, Posthumus M, Schwellnus MP. The COL1A1 gene and acute soft tissue ruptures. Br J Sports Med. 2010;44(14):1063–4.

249. Shambaugh JP, Klein A, Herbert JH. Structural measures as predictors of injury basketball players. Med Sci Sports Exerc. 1991 May;23(5):522–7.

250. Brandon ML, Haynes PT, Bonamo JR, Flynn MI, Barrett GR, Sherman MF. The Association Between Posterior-Inferior Tibial Slope and Anterior Cruciate Ligament Insufficiency. Arthroscopy: The Journal of Arthroscopic & Related Surgery. 2006 Aug;22(8):894–9.

251. James R, Kesturu G, Balian G, Chhabra AB. Tendon: Biology, Biomechanics, Repair, Growth Factors, and Evolving Treatment Options. Journal of Hand Surgery. 2008;33(1):102–12.

252. Wenstrup RJ, Florer JB, Brunskill EW, Bell SM, Chervoneva I, Birk DE. Type V collagen controls the initiation of collagen fibril assembly. Journal of Biological Chemistry. 2004;279(51):53331–7.

253. Fichard A, Kleman JP, Ruggiero F. Another look at collagen V and XI molecules. Matrix Biol. 1995 Jul;14(7):515–31.

254. Oliva F, Zocchi L, Codispoti A, Candi E, Celi M, Melino G, et al. Transglutaminases expression in human supraspinatus tendon ruptures and in mouse tendons. Biochem Biophys Res Commun. 2009 Feb;379(4):887–91.

255. Lavagnino M, Wall ME, Little D, Banes AJ, Guilak F, Arnoczky SP. Tendon mechanobiology: Current knowledge and future research opportunities. Journal of Orthopaedic Research. 2015;33(6):813–22.

256. Arnoczky SP, Lavagnino M, Egerbacher M. The mechanobiological aetiopathogenesis of tendinopathy: is it the over-stimulation or the under-stimulation of tendon cells? Int J Exp Pathol. 2007 Aug;88(4):217–26.

257. Kannus P. Structure of the tendon connective tissue. Scand J Med Sci Sports. 2000;10(6):312–20.

258. Frank CB, Hart DA, Shrive NG. Molecular biology and biomechanics of normal and healing ligaments—a review. Osteoarthritis Cartilage. 1999 Jan;7(1):130–40.

259. Screen HRC, Berk DE, Kadler KE, Ramirez F, Young MF. Tendon functional extracellular matrix. Journal of Orthopaedic Research. 2015;33(6):793–9.

260. Badylak SF, Freytes DO, Gilbert TW. Extracellular matrix as a biological scaffold material: Structure and function. Acta Biomater. 2015 Sep;23:S17–26.

261. Zhang G, Young BB, Ezura Y, Favata M, Soslowsky LJ, Chakravarti S, et al. Development of tendon structure and function: Regulation of collagen fibrillogenesis. Journal of Musculoskeletal Neuronal Interactions. 2005;5(1):5–21.

262. Wang JHC, Guo Q, Li B. Tendon Biomechanics and Mechanobiology—A Minireview of Basic Concepts and Recent Advancements. Journal of Hand Therapy [Internet]. 2012 Apr;25(2):133–41. Available from: https://linkinghub.elsevier.com/retrieve/pii/S089411301100086X

263. Lipman K, Wang C, Ting K, Soo C, Zheng Z. Tendinopathy: Injury, repair, and current exploration. Drug Des Devel Ther. 2018;12:591–603.

264. Dean BJF, Franklin SL, Carr AJ. A systematic review of the histological and molecular changes in rotator cuff disease. Bone Joint Res. 2012 Jul;1(7):158–66.

265. Saeidi N, Karmelek KP, Paten JA, Zareian R, DiMasi E, Ruberti JW. Molecular crowding of collagen: A pathway to produce highly-organized collagenous structures. Biomaterials. 2012 Oct;33(30):7366–74.

266. Birk DE, Brückner P. Collagens, Suprastructures, and Collagen Fibril Assembly. In: The Extracellular Matrix: an Overview. Berlin, Heidelberg, Heidelberg: Springer Berlin Heidelberg; 2011. p. 77–115.

267. Gordon MK, Hahn RA. Collagens. Cell Tissue Res. 2010;339(1):247–57.

268. Frank CBCBCBCB. Ligament structure, physiology and function. J Musculoskelet Neuronal Interact. 2004 Jun;4(2):199–201.

269. Thornton GM, Oliynyk A, Frank CB, Shrive NG. Ligament creep cannot be predicted from stress relaxation at low stress: A biomechanical study of the rabbit medial collateral ligament. Journal of Orthopaedic Research. 1997 Sep;15(5):652–6.

270. Sherman VR, Yang W, Meyers MA. The materials science of collagen. J Mech Behav Biomed Mater. 2015;52:22–50.

271. Fratzl P. Collagen: Structure and Mechanics, an Introduction. In: Collagen. Boston, MA: Springer US; 2008. p. 1–13.

272. Voycheck CA, Luu K, McMahon PJ, Debski RE. Collagen fiber alignment and maximum principal strain in the glenohumeral capsule predict location of failure during uniaxial extension. Biomech Model Mechanobiol. 2014 Apr;13(2):379–85.

273. Dahlgren LA. Pathobiology of Tendon and Ligament Injuries. Clinical Techniques in Equine Practice. 2007 Sep;6(3):168–73.

274. Çabuk H, Kuşku Çabuk F. Mechanoreceptors of the ligaments and tendons around the knee. Clinical Anatomy. 2016 Sep;29(6):789–95.

275. Wan C, Hao Z, Wen S. A review on research on development of ligament constitutive relations on macro, meso, and micro levels. Acta Mechanica Solida Sinica. 2013;26(4):331–43.

276. Bates NA, Myer GD, Shearn JT, Hewett TE. Anterior cruciate ligament biomechanics during robotic and mechanical simulations of physiologic and clinical motion tasks: A systematic review and meta-analysis. Clinical Biomechanics. 2015 Jan;30(1):1–13.

277. Boden BP, Dean GS, Feagin JA, Garrett WE. Mechanisms of anterior cruciate ligament injury. Orthopedics. 2000 Jun;23(6):573–8.

278. Paschos NK, Gartzonikas D, Barkoula NMM, Moraiti C, Paipetis A, Matikas TE, et al. Cadaveric Study of Anterior Cruciate Ligament Failure Patterns Under Uniaxial Tension Along the Ligament. Arthroscopy: The Journal of Arthroscopic & Related Surgery. 2010 Jul;26(7):957–67.

279. Arnoux PJJ, Cavallero C, Chabrand P, Brunet C. Knee ligament failure under dynamic loadings. International Journal of Crashworthiness. 2002;7(3):255–68.

280. Krogsgaard MR, Dyhre-Poulsen P, Fischer-Rasmussen T. Cruciate ligament reflexes. Journal of Electromyography and Kinesiology. 2002 Jun;12(3):177–82.

281. Solomonow M. Sensory – Motor control of ligaments and associated neuromuscular disorders. Journal of Electromyography and Kinesiology. 2006 Dec;16(6):549–67.

282. Becker R, Hirschmann MT, Karlsson J. The role of ligament tension and sensomotoric system in total knee arthroplasty. Knee Surgery, Sports Traumatology, Arthroscopy. 2017 Jun;25(6):1663–5.

283. Katonis P, Papoutsidakis A, Aligizakis A, Tzanakakis G, Kontakis GM, Papagelopoulos PJ. Mechanoreceptors of the posterior cruciate ligament. Vol. 36, Journal of International Medical Research. 2008. p. 387–93.

284. Hagert E. Proprioception of the Wrist Joint: A Review of Current Concepts and Possible Implications on the Rehabilitation of the Wrist. Journal of Hand Therapy. 2010 Jan;23(1):2–17.

285. Exposito JY, Valcourt U, Cluzel C, Lethias C. The fibrillar collagen family. Int J Mol Sci. 2010;11(2):407–26.

286. Kadler KE, Hill A, Canty-Laird EG. Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. Curr Opin Cell Biol [Internet]. 2008;20(5):495–501. Available from: http://dx.doi.org/10.1016/j.ceb.2008.06.008

287. Kadler KE, Baldock C, Bella J, Boot-Handford RP. Collagens at a glance. J Cell Sci. 2007;120(12):1955–8.

288. Bella J, Hulmes DJS. Fibrillar Collagens. In: Parry D, Squire J, editors. Fibrous Proteins: Structures and Mechanisms Subcellular Biochemistry. Springer, Cham; 2017. p. 457–90.

289. Hulmes DJS. Building Collagen Molecules, Fibrils, and Suprafibrillar Structures. J Struct Biol. 2002 Jan;137(1–2):2–10.

290. Brodsky B, Persikov A V. Molecular Structure of the Collagen Triple Helix. In: Advances in Protein Chemistry. 2005. p. 301–39.

291. Brodsky B, Thiagarajan G, Madhan B, Kar K. Triple-helical peptides: An approach to collagen conformation, stability, and self-association. Biopolymers. 2008;89(5):345–53.

292. Brinckmann J. Collagens at a Glance. In: Topics in Current Chemistry. 2005. p. 1–6.

293. Wess TJ. Collagen Fibrillar Structure and Hierarchies. In: Collagen. Boston, MA: Springer US; 2008. p. 49–80.

294. Banos CC, Thomas AH, Kuo CK. Collagen fibrillogenesis in tendon development: Current models and regulation of fibril assembly. Birth Defects Res C Embryo Today. 2008;84(3):228–44.

295. Collins M, Raleigh SM. Genetic risk factors for musculoskeletal soft tissue injuries. Med Sport Sci. 2009;54(June 2014):136–49.

296. Luo Y, Sinkeviciute D, He Y, Karsdal M, Henrotin Y, Mobasheri A, et al. The minor collagens in articular cartilage. Protein Cell. 2017;8(8):560–72.

297. Eyre DR, Apon S, Wu JJ, Ericsson LH, Walsh KA. Collagen type IX: Evidence for covalent linkages to type II collagen in cartilage. FEBS Lett. 1987;220(2):337–41.

298. Smith SM, Birk DE. Focus on Molecules: Collagens V and XI. Exp Eye Res. 2012 May;98(2):105–6.

299. Birk DE, Fitch JM, Babiarz JP, Doane KJ, Linsenmayer TF. Collagen fibrillogenesis in vitro: Interaction of types I and V collagen regulates fibril diameter. J Cell Sci. 1990;95(4):649–57.

300. Connizzo BK, Freedman BR, Fried JH, Sun M, Birk DE, Soslowsky LJ. Regulatory role of collagen V in establishing mechanical properties of tendons and ligaments is tissue dependent. Journal of Orthopaedic Research [Internet]. 2015 Jun;33(6):882–8. Available from: http://doi.wiley.com/10.1002/jor.22893

301. Veit G, Kobbe B, Keene DR, Paulsson M, Koch M, Wagener R. Collagen XXVIII, a novel von Willebrand factor A domain-containing protein with many imperfections in the collagenous domain. Journal of Biological Chemistry. 2006;281(6):3494–504.

302. Grimal S, Puech S, Wagener R, Ventéo S, Carroll P, Fichard-Carroll A. Collagen XXVIII is a distinctive component of the peripheral nervous system nodes of Ranvier and surrounds nonmyelinating glial cells. Glia. 2010;58(16):1977–87.

303. Shoulders MD, Raines RT. Collagen structure and stability. Annu Rev Biochem. 2009;78:929–58.

304. Bächinger HP, Mizuno K, Vranka J, Boudko S. Collagen formation and structure. In: Ltd. E, editor. Comprehensive Natural Products II: Chemistry and Biology. 2010. p. 469–530.

305. Boudko SP, Engel J, Bächinger HP. The crucial role of trimerization domains in collagen folding. Int J Biochem Cell Biol. 2012 Jan;44(1):21–32.

306. Frank S, Boudko S, Mizuno K, Schulthess T, Engel J, Bächinger HP. Collagen triple helix formation can be nucleated at either end. Journal of Biological Chemistry. 2003;278(10):7747–50.

307. Sweeney SM, Orgel JP, Fertala A, McAuliffe JD, Turner KR, Di Lullo GA, et al. Candidate cell and matrix interaction domains on the collagen fibril, the predominant protein of vertebrates. Journal of Biological Chemistry. 2008;283(30):21187–97.

308. Boudko SP, Zientek KD, Vance J, Hacker JL, Engel J, Bächinger HP. The NC2 domain of collagen IX provides chain selection and heterotrimerization. Journal of Biological Chemistry. 2010;285(31):23721–31.

309. Chiquet M, Birk DE, Bönnemann CG, Koch M. Collagen XII: Protecting bone and muscle integrity by organizing collagen fibrils. Int J Biochem Cell Biol. 2014 Aug;53:51–4.

310. Nishiyama T, McDonough AM, Bruns RR, Burgeson RE. Type XII and XIV collagens mediate interactions between banded collagen fibers in vitro and may modulate extracellular matrix deformability. J Biol Chem. 1994 Nov 11;269(45):28193–9.

311. Wälchli C, Koch M, Chiquet M, Odermatt BF, Trueb B. Tissue-specific expression of the fibril-associated collagens XII and XIV. J Cell Sci. 1994;107(2):669–81.

312. Saika S, Miyamoto T, Tanaka T, Ishida I, Ohnishi Y, Ooshima A. Collagens XII and XIV (FACITs) in capsular opacification and in cultured lens epithelial cells. Curr Eye Res. 2001;23(6):463–8.

313. Noponen-Hietala N, Kyllönen E, Männikkö M, Ilkko E, Karppinen J, Ott J, et al. Sequence variations in the collagen IX and XI genes are associated with degenerative lumbar spinal stenosis. Ann Rheum Dis. 2003;62(12):1208–14.

314. Higashino K, Matsui Y, Yagi S, Takata Y, Goto T, Sakai T, et al. The alpha2 type IX collagen tryptophan polymorphism is associated with the severity of disc degeneration in younger patients with herniated nucleus pulposus of the lumbar spine. Int Orthop. 2007;31(1):107–11.

315. Oudart JB, Monboisse JC, Maquart FX, Brassart B, Brassart-Pasco S, Ramont L. Type XIX collagen: A new partner in the interactions between tumor cells and their microenvironment. Vols. 57–58, Matrix Biology. International Society of Matrix Biology; 2017. p. 169–77.

316. Ansorge HL, Meng X, Zhang G, Veit G, Sun M, Klement JF, et al. Type XIV collagen regulates fibrillogenesis: Premature collagen fibril growth and tissue dysfunction in null mice. Journal of Biological Chemistry. 2009;284(13):8427–38.

317. Hemmavanh C, Koch M, Birk DE, Espana EM. Abnormal corneal endothelial maturation in collagen XII and XIV null mice. Invest Ophthalmol Vis Sci [Internet]. 2013 May 7;54(5):3297–308. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23599329

318. Siebuhr AS, Karsdal MA. Type XIII Collagen. In: Biochemistry of Collagens, Laminins and Elastin. Elsevier; 2016. p. 87–91.

319. Tao G, Levay AKK, Peacock JDD, Huk DJJ, Both SNN, Purcell NHH, et al. Collagen XIV is important for growth and structural integrity of the myocardium. J Mol Cell Cardiol. 2012;53(5):626–38.

320. Listrat A, Lethias C, Hocquette JF, Renand G, Ménissier F, Geay Y, et al. Age-related changes and location of types I, III, XII and XIV collagen during development of skeletal muscles from genetically different animals. Histochemical Journal. 2000;32(6):349–56.

321. Germain L, Rémy-Zolghadri M, Auger F. Tissue engineering of the vascular system: From capillaries to larger blood vessels. Med Biol Eng Comput. 2000 Mar;38(2):232–40.

322. Grässel S, Timpl R, Tan EML, Chu ML. Biosynthesis and processing of type XVI collagen in human fibroblasts and smooth muscle cells. Eur J Biochem. 1996;242(3):576–84.

323. Kassner A, Hansen U, Miosge N, Reinhardt DP, Aigner T, Bruckner-Tuderman L, et al. Discrete integration of collagen XVI into tissue-specific collagen fibrils or beaded microfibrils. Matrix Biology. 2003 Apr;22(2):131–43.

324. Akagi A, Tajima S, Ishibashi A, Matsubara Y, Takehana M, Kobayashi S, et al. Type XVI Collagen is Expressed in Factor XIIIa+ Monocyte-Derived Dermal Dendrocytes and Constitutes a Potential Substrate for Factor XIIIa. Journal of Investigative Dermatology. 2002 Feb;118(2):267–74.

325. Koch M, Foley JE, Hahn R, Zhou P, Burgeson RE, Gerecke DR, et al. α1(XX) Collagen, a New Member of the Collagen Subfamily, Fibril-associated Collagens with Interrupted Triple Helices - 200.pdf. Journal of Biological Chemistry. 2001;276(25):23120–6.

326. Willumsen N, Karsdal MA. Type XX Collagen. In: Biochemistry of Collagens, Laminins and Elastin. Elsevier; 2016. p. 127–9.

327. Michelacci YM. Collagens and proteoglycans of the corneal extracellular matrix. Brazilian Journal of Medical and Biological Research. 2003 Aug;36(8):1037–46.

328. Myers JC, Li D, Amenta PS, Clark CC, Nagaswami C, Weisel JW. Type XIX Collagen Purified from Human Umbilical Cord Is Characterized by Multiple Sharp Kinks Delineating Collagenous Subdomains and by Intermolecular Aggregates via Globular, Disulfide-linked, and Heparin-binding Amino Termini. Journal of Biological Chemistry. 2003 Aug;278(34):32047–57.

329. Myers JC, Li D, Bageris A, Abraham V, Dion AS, Amenta PS. Biochemical and immunohistochemical characterization of human type XIX defines a novel class of basement membrane zone collagens. Am J Pathol. 1997 Dec;151(6):1729–40.

330. Li HC, Huang CC, Chen SF, Chou MY. Assembly of homotrimeric type XXI minicollagen by coexpression of prolyl 4-hydroxylase in stably transfected Drosophila melanogaster S2 cells. Biochem Biophys Res Commun. 2005 Oct;336(2):375–85.

331. Kehlet SN, Karsdal MA. Type XXII Collagen. In: Biochemistry of Collagens, Laminins and Elastin. Elsevier; 2016. p. 135–7.

332. Koch M, Schulze J, Hansen U, Ashwodt T, Keene DR, Brunken WJ, et al. A novel marker of tissue junctions, collagen XXII. Journal of Biological Chemistry. 2004;279(21):22514–21.

333. Sutmuller M, Bruijn JA, De Heer E. Collagen types VIII and X, two non-fibrillar, short-chain collagens. Structure homologies, functions and involvement in pathology. Histol Histopathol. 1997 Apr;12(2):557–66.

334. Baldock C, Sherratt MJ, Shuttleworth CA, Kielty CM. The supramolecular organization of collagen VI microfibrils. J Mol Biol. 2003 Jul;330(2):297–307.

335. Knupp C, Squire JM. Molecular packing in network-forming collagens. Adv Protein Chem. 2005;70:375–403.

336. Engvall E, Hessle H, Klier G. Molecular assembly, secretion, and matrix deposition of type VI collagen. Journal of Cell Biology. 1986;102(3):703–10.

337. Halper J. Progress in Heritable Soft Connective Tissue Diseases. Adv Exp Med Biol. 2014;77–94.

338. Kuo HJ, Maslen CL, Keene DR, Glanville RW. Type VI collagen anchors endothelial basement membranes by interacting with type IV collagen. Journal of Biological Chemistry. 1997;272(42):26522–9.

339. Pokidysheva E, Boudko S, Vranka J, Zientek K, Maddox K, Moser M, et al. Biological role of prolyl 3-hydroxylation in type IV collagen. Proc Natl Acad Sci U S A. 2014;111(1):161–6.

340. Hurskainen M, Eklund L, Hägg PO, Fruttiger M, Sormunen R, Ilves M, et al. Abnormal maturation of the retinal vasculature in type XVIII collagen/endostatin deficient mice and changes in retinal glial cells due to lack of collagen types XV and XVIII. The FASEB Journal. 2005;19(11):1564–6.

341. Ylikärppä R, Eklund L, Sormunen R, Kontiola AI, Utriainen A, Määttä M, et al. Lack of type XVIII collagen results in anterior ocular defects. The FASEB Journal. 2003;17(15):2257–9.

342. Hansen NUB, Willumsen N, Sand JMB, Larsen L, Karsdal MA, Leeming DJ. Type VIII collagen is elevated in diseases associated with angiogenesis and vascular remodeling. Clin Biochem [Internet]. 2016;49(12):903–8. Available from: http://dx.doi.org/10.1016/j.clinbiochem.2016.05.023

343. Illidge C, Kielty C, Shuttleworth A. The α1(VIII) and α2(VIII) chains of type VIII collagen can form stable homotrimeric molecules. Vol. 273, Journal of Biological Chemistry. 1998. p. 22091–5.

344. Chen M, Kasahara N, Keene DR, Chan L, Hoeffier WK, Finlay D, et al. Restoration of type VII collagen expression and function in dystrophic epidermolysis bullosa. Nat Genet. 2002;32(4):670–5.

345. Eklund L, Piuhola J, Komulainen J, Sormunen R, Ongvarrasopone C, Fässler R, et al. Lack of type XV collagen causes a skeletal myopathy and cardiovascular defects in mice. Proc Natl Acad Sci U S A. 2001;98(3):1194–9.

346. Lisignoli G, Codeluppi K, Todoerti K, Manferdini C, Piacentini A, Zini N, et al. Gene array profile identifies collagen type XV as a novel human osteoblast-secreted matrix protein. J Cell Physiol. 2009 Aug;220(2):401–9.

347. Tomono Y, Naito I, Ando K, Yonezawa T, Sado Y, Hirakawa S, et al. Epitope-defined monoclonal antibodies against multiplexin collagens demonstrate that type XV and XVIII collagens are expressed in specialized basement membranes. Cell Struct Funct. 2002;27(1):9–20.

348. Rehn M, Pihlajaniemi T. α1(XVIII), a collagen chain with frequent interruptions in the collagenous sequence, a distinct tissue distribution, and homology with type XV collagen. Vol. 91, Proceedings of the National Academy of Sciences of the United States of America. 1994. p. 4234–8.

349. Lin HC, Chang JH, Jain S, Gabison EE, Kure T, Kato T, et al. Matrilysin cleavage of corneal collagen type XVIII NC1 domain and generation of a 28-kDa fragment. Invest Ophthalmol Vis Sci. 2001 Oct;42(11):2517–24.

350. Areida SK, Reinhardt DP, Müller PK, Fietzek PP, Köwitzi J, Marinkovich MP, et al. Properties of the collagen type XVII ectodomain. Evidence for N- to C-terminal triple helix folding. Journal of Biological Chemistry. 2001;276(2):1594–601.

351. Nykvist P, Tasanen K, Viitasalo T, Käpylä J, Jokinen J, Bruckner-Tuderman L, et al. The Cell Adhesion Domain of Type XVII Collagen Promotes Integrin-mediated Cell Spreading by a Novel Mechanism. Journal of Biological Chemistry. 2001;276(42):38673–9.

352. Gatalica B, Pulkkinen L, Li K, Kuokkanen K, Ryynänen M, McGrath JA, et al. Cloning of the human type XVII collagen gene (COL17A1), and detection of novel mutations in generalized atrophic benign epidermolysis bullosa. Am J Hum Genet. 1997;60(2):352–65.

353. Ruzzi L, Posteraro P, Zambruno G, Castiglia D, D'Alessio M, Pas H, et al. A Homozygous Nonsense Mutation in Type XVII Collagen Gene (COL17A1) Uncovers an Alternatively Spliced mRNA Accounting for an Unusually Mild Form of Non-Herlitz Junctional Epidermolysis Bullosa. Journal of Investigative Dermatology. 2001 Jan;116(1):182–7.

354. Koch M, Veit G, Stricker S, Bhatt P, Kutsch S, Zhou P, et al. Expression of type XXIII collagen mRNA and protein. Journal of Biological Chemistry. 2006;281(30):21546–57.

355. Banyard J, Bao L, Zetter BR. Type XXIII collagen, a new transmembrane collagen identified in metastatic tumor cells. Journal of Biological Chemistry. 2003;278(23):20989–94.

356. Hulmes DJS. Collagen diversity, synthesis and assembly. Collagen: Structure and Mechanics. 2008;15–47.

357. Xu F, Zhang L, Koder RL, Nanda V. De Novo Self-Assembling Collagen Heterotrimers Using Explicit Positive and Negative Design. Biochemistry. 2010 Mar;49(11):2307–16.

358. Fallas JA, O'Leary LER, Hartgerink JD. Synthetic collagen mimics: self-assembly of homotrimers, heterotrimers and higher order structures. Chem Soc Rev. 2010;39(9):3510.

359. Gauba V, Hartgerink JD. Self-Assembled Heterotrimeric Collagen Triple Helices Directed through Electrostatic Interactions. J Am Chem Soc. 2007 Mar;129(9):2683–90.

360. Canty EG, Kadler KE. Procollagen trafficking, processing and fibrillogenesis. J Cell Sci. 2005;118(7):1341–53.

361. Järvinen TAH, Järvinen TLN, Kääriäinen M, Kalimo H, Järvinen M. Muscle injuries: Biology and treatment. American Journal of Sports Medicine. 2005;33(5):745–64.

362. Ricard-Blum S, Ruggiero F, van der Rest M. The Collagen Superfamily. In 2005. p. 35–84.

363. Kielty CM, Grant ME. The Collagen Family: Structure, Assembly, and Organization in the Extracellular Matrix. In: Connective Tissue and Its Heritable Disorders. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2002. p. 159–221.

364. Riley G. Tendon and ligament biochemistry and pathology. Soft tissue rheumatology. 2004;2053(January 2011):1–37.

365. Franchi M, Trirè A, Quaranta M, Orsini E, Ottani V. Collagen structure of tendon relates to function. ScientificWorldJournal. 2007;7:404–20.

366. Tresoldi I, Oliva F, Benvenuto M, Fantini M, Masuelli L, Bei R, et al. Tendon's ultrastructure. Muscles Ligaments Tendons J. 2013;3(1):2–6.

367. Hast MW, Zuskov A, Soslowsky LJ. The role of animal models in tendon research. Bone Joint Res. 2014;3(6):193–202.

368. Perumal S, Antipova O, Orgel JPRO. Collagen fibril architecture, domain organization, and triple-helical conformation govern its proteolysis. Proceedings of the National Academy of Sciences. 2008 Feb;105(8):2824–9.

369. Rein S, Hagert E, Schneiders W, Fieguth A, Zwipp H. Histological Analysis of the Structural Composition of Ankle Ligaments. Foot Ankle Int. 2015 Feb;36(2):211–24.

370. P. Foster B. Genetic Variation, Protein Composition and Potential Influences on Tendon Properties in Humans. The Open Sports Medicine Journal. 2012;6(1):8–21.

371. Christiansen DL, Huang EK, Silver FH. Assembly of type I collagen: fusion of fibril subunits and the influence of fibril diameter on mechanical properties. Matrix Biology. 2000 Sep;19(5):409–20.

372. Orgel JPRO, San Antonio JD, Antipova O. Molecular and structural mapping of collagen fibril interactions. Connect Tissue Res. 2011;52(1):2–17.

373. National Institute of Health. COL1A1 gene. 2017.

374. Mann V, Ralston SH. Meta-analysis of COL1A1 Sp1 polymorphism in relation to bone mineral density and osteoporotic fracture. Bone. 2003;32(6):711–7.

375. Li LJ, Lyu F, Song YW, Wang O, Jiang Y, Xia WB, et al. Genotype-phenotype relationship in a large cohort of osteogenesis imperfecta patients with COL1A1 mutations revealed by a new scoring system. Chin Med J (Engl). 2019 Jan;132(2):145–53.

376. Brady AF, Demirdas S, Fournel-Gigleux S, Ghali N, Giunta C, Kapferer-Seebacher I, et al. The Ehlers-Danlos syndromes, rare types. Am J Med Genet C Semin Med Genet. 2017 Mar;175(1):70–115.

377. Jin H, van't Hof RJ, Albagha OME, Ralston SH. Promoter and intron 1 polymorphisms of COL1A1 interact to regulate transcription and susceptibility to osteoporosis. Hum Mol Genet. 2009 Aug;18(15):2729–38.

378. Chen W, Meyer NC, Mckenna MJ, Pfister M, McBride DJ, Fukushima K, et al. Single-nucleotide polymorphisms in the COL1A1 regulatory regions are associated with otosclerosis. Clin Genet. 2007;71(5):406–14.

379. Harner CD, Paulos LE, Greenwald AE, Rosenberg TD, Cooley VC. Detailed Analysis of Patients with Bilateral Anterior Cruciate Ligament Injuries. Am J Sports Med. 1994 Jan;22(1):37–43.

380. Stępień-Słodkowska M, Ficek K, Eider J, Leońska-Duniec A, MacIejewska-Karłowska A, Sawczuk M, et al. The +1245G/T polymorpsisms in the collagen type I alpha 1 (Col1a1) gene in polish skiers with anterior cruciate ligament injury. Biol Sport. 2013;30(1):57–60.

381. Rouault K, Scotet V, Autret S, Gaucher F, Dubrana F, Tanguy D, et al. Do HOXB9 and COL1A1 genes play a role in congenital dislocation of the hip? Study in a Caucasian population. Osteoarthritis Cartilage. 2009;17(8):1099–105.

382. Näkki A, Videman T, Kujala UM, Suhonen M, Männikkö M, Peltonen L, et al. Candidate gene association study of magnetic resonance imaging-based hip osteoarthritis (OA): Evidence for COL9A2 gene as a common predisposing factor for hip OA and lumbar disc degeneration. Journal of Rheumatology. 2011;38(4):747–52.

383. Pluijm SMF, Van Essen HW, Bravenboer N, Uitterlinden AG, Smit JH, Pols HAP, et al. Collagen type I α1 Sp1 polymorphism, osteoporosis, and intervertebral disc degeneration in older men and women. Ann Rheum Dis. 2004;63(1):71–7.

384. Omair A, Mannion AF, Holden M, Leivseth G, Fairbank J, Hägg O, et al. Age and pro-inflammatory gene polymorphisms influence adjacent segment disc degeneration more than fusion does in patients treated for chronic low back pain. European Spine Journal. 2016;25(1):2–13.

385. Mokone GG, Schwellnus MP, Noakes TD, Collins M. The COL5A1 gene and Achilles tendon pathology. Scand J Med Sci Sports. 2006;16(1):19–26.

386. Abrahams Y, Laguette MJ, Prince S, Collins M. Polymorphisms within the COL5A1 3′-UTR That Alters mRNA Structure and the MIR608 Gene are Associated with Achilles Tendinopathy. Ann Hum Genet [Internet]. 2013 May;77(3):204–14. Available from: http://doi.wiley.com/10.1111/ahg.12013

387. Rodriguez SV. VDR, GDF5, COL1A1, THBS2 AND CHST GENE POLYMORPHISMS ASSOCIATED WITH SYMPTOMATIC LUMBAR DISC HERNIATION IN SPANISH SURGICAL PATIENTS. Orthopaedic Proceedings. 2017;99-B(Supp 1):119.

388. Collins M, Posthumus M. Type V collagen genotype and exercise-related phenotype relationships: A novel hypothesis. Exerc Sport Sci Rev. 2011;39(4):191–8.

389. Laguette MJ, Prince S, Collins M. Sequence variants within the 3'-UTR of the col5a1 gene alters mrna stability: implications for musculoskeletal soft tissue injuries. Br J Sports Med. 2011;45(4):322–322.

390. Cescon M, Gattazzo F, Chen P, Bonaldo P. Collagen VI at a glance. J Cell Sci. 2015 Oct;128(19):3525–31.

391. Fitzgerald J, Holden P, Hansen U. The expanded collagen VI family: new chains and new questions. Connect Tissue Res. 2013 Nov;54(6):345–50.

392. Liu X, Wu H, Byrne M, Krane S, Jaenisch R. Type III collagen is crucial for collagen I fibrillogenesis and for normal cardiovascular development. Proceedings of the National Academy of Sciences. 1997 Mar;94(5):1852–6.

393. Lampe AK, Bushby KMD. Collagen VI related muscle disorders. J Med Genet. 2005;42(9):673–85.

394. Butterfield RJ, Foley AR, Dastgir J, Asman S, Dunn DM, Zou Y, et al. Position of glycine substitutions in the triple helix of COL6A1, COL6A2, and COL6A3 is correlated with

severity and mode of inheritance in collagen vi myopathies. Hum Mutat. 2013;34(11):1558–67.

395. Tanaka T, Ikari K, Furushima K, Okada A, Tanaka H, Furukawa KI, et al. Genomewide Linkage and Linkage Disequilibrium Analyses Identify COL6A1, on Chromosome 21, as the Locus for Ossification of the Posterior Longitudinal Ligament of the Spine. Am J Hum Genet. 2003;73(4):812–22.

396. Bushby KMD, Collins J, Hicks D. Collagen type VI myopathies. Adv Exp Med Biol. 2014;802:185–99.

397. Kong Q, Ma X, Li F, Guo Z, Qi Q, Li W, et al. COL6A1 Polymorphisms Associated With Ossification of the Ligamentum Flavum and Ossification of the Posterior Longitudinal Ligament. Spine (Phila Pa 1976). 2007 Dec;32(25):2834–8.

398. O'Connell K, Posthumus M, Collins M. Collagen gene interactions and endurance running performance. South African Journal of Sports Medicine. 2014;26(1):9.

399. Videman T, Saarela J, Kaprio J, Näkki A, Levälahti E, Gill K, et al. Associations of 25 structural, degradative, and inflammatory candidate genes with lumbar disc desiccation, bulging, and height narrowing. Arthritis Rheum. 2009;60(2):470–81.

400. Mio F, Chiba K, Hirose Y, Kawaguchi Y, Mikami Y, Oya T, et al. A functional polymorphism in COL11A1, which encodes the α1 chain of type XI collagen, is associated with susceptibility to.pdf. Am J Hum Genet. 2007;81(6):1271–7.

401. Kania AM, Reichenberger E, Baur ST, Karimbux NY, Taylor RW, Olsen BR, et al. Structural variation of type XII collagen at its carboxyl-terminal NC1 domain generated by tissue-specific alternative splicing. Journal of Biological Chemistry. 1999;274(31):22053–9.

402. DE B. Type V collagen: heterotypic type I/V collagen interactions in the regulation of fibril assembly. Micron. 2001;32(3):223–37.

403. Young BB, Zhang G, Koch M, Birk DE. The roles of types XII and XIV collagen in fibrillogenesis and matrix assembly in the developing cornea. J Cell Biochem. 2002;87(2):208–20.

404. Connizzo BK, Yannascoli SM, Soslowsky LJ. Structure-function relationships of postnatal tendon development: A parallel to healing. Matrix Biology [Internet]. 2013;32(2):106–16. Available from: http://dx.doi.org/10.1016/j.matbio.2013.01.007

405. Zhao D, Zhang Q, Lu Q, Hong C, Luo T, Duan Q, et al. Correlations between the genetic variations in the COL1A1, COL5A1, COL12A1, and b-fibrinogen genes and anterior cruciate ligament injury in Chinese patients. J Athl Train. 2020;55(5):515–21.

406. Tilkeridis C, Bei T, Garantziotis S, Stratakis CA. Association of a COL1A1 polymorphism with lumbar disc disease in young military recruits. J Med Genet. 2005;42(7):7–9.

407. Kim SK, Kleimeyer JP, Ahmed MA, Avins AL, Fredericson M, Dragoo JL, et al. Two genetic loci associated with ankle injury. PLoS One. 2017;12(9):1–14.

408. Dalewski B, Białkowska K, Pałka Ł, Jakubowska A, Kiczmer P, Sobolewska E. COL5A1 RS12722 Is Associated with Temporomandibular Joint Anterior Disc Displacement without Reduction in Polish Caucasians. Cells. 2021 Sep 14;10(9):2423.

409. Hay M. An investigation of DNA sequence variants in genes that regulate collagen fibrillogenesis and predisposition to musculoskeletal soft tissue injuries. Thesis. 2013;(April).

410. O'Connell K, Posthumus M, Schwellnus MP, Collins M. Collagen genes and exercise-associated muscle cramping. Clinical Journal of Sport Medicine. 2013;23(1):64–9.

411. O'Connell KS, Sean K, O'connell K. Investigation of selected collagen genes in exercise-related musculoskeletal soft tissue phenotypes. Thesis. 2014.

412. Price MJ, Tuca M, Cordasco FA, Green DW. Nonmodifiable risk factors for anterior cruciate ligament injury. Curr Opin Pediatr. 2017;29(1):55–64.

413. Obremskey WT, Pappas N, Attallah-Wasif E, Tornetta P, Bhandari M. Level of evidence in orthopaedic journals. Journal of Bone and Joint Surgery - Series A. 2005;87(12 I):2632–8.

414. Sawaya, G.F., Guiruis-Blake, J., Harris, R., and Petitti D. US Preventive Services Task Force. Update on the Methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. Ann Intern Med. 2007;147(12):871–5.

415. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. PLoS Med. 2009;6(7).

416. Abrahams S, McFie S, Patricios J, Posthumus M, September A V. Risk factors for sports concussion: An evidence-based systematic review. Br J Sports Med. 2014;48(2):91–7.

417. Lynch SS, Thigpen CA, Mihalik JP, Prentice WE, Padua D. The effects of an exercise intervention on forward head and rounded shoulder postures in elite swimmers. Br J Sports Med [Internet]. 2010 Apr 1;44(5):376–81. Available from: https://bjsm.bmj.com/lookup/doi/10.1136/bjsm.2009.066837

418. Feijen S, Struyf T, Kuppens K, Tate A, Struyf F. Prediction of Shoulder Pain in Youth Competitive Swimmers: The Development and Internal Validation of a Prognostic Prediction Model. American Journal of Sports Medicine. 2021;49(1):154–61.

419. Swanik KA, Lephart SM, Swanik CB, Lephart SP, Stone DA, Fu FH. The effects of shoulder plyometric training on proprioception and selected muscle performance characteristics. J Shoulder Elbow Surg. 2002;11(6):579–86.

420. Drigny J, Gauthier A, Reboursière E, Guermont H, Gremeaux V, Edouard P. Shoulder Muscle Imbalance as a Risk for Shoulder Injury in Elite Adolescent Swimmers: A Prospective Study. J Hum Kinet. 2020;75(1):103–13.

421. Couanis G, Breidahl W, Burnham S. The relationship between subacromial bursa thickness on ultrasound and shoulder pain in open water endurance swimmers over time. J Sci Med Sport [Internet]. 2015 Jul;18(4):373–7. Available from: http://dx.doi.org/10.1016/j.jsams.2014.05.004

422. Ristolainen, L., Kettunen, J.A., Waller, B., Heinonen, A., Kujala UM. Training-related risk factors in the etiology of overuse injuries in endurance sports. J Sports Med Phys Fitness. 2014;54(1):78–87.

423. Tessaro M, Granzotto G, Poser A, Plebani G, Rossi A. Shoulder Pain in Competitive Teenage Swimmers and It'S Prevention: a Retrospective Epidemiological Cross Sectional Study of Prevalence. Int J Sports Phys Ther. 2017;12(5):798–811.

424. Tate A, Turner GN, Knab SE, Jorgensen C, Strittmatter A, Michener LA. Risk factors associated with shoulder pain and disability across the lifespan of competitive swimmers. J Athl Train. 2012;47(2):149–58.

425. McMaster WC, Roberts A, Stoddard T. A Correlation Between Shoulder Laxity and Interfering Pain in Competitive Swimmers. Am J Sports Med. 1998 Jan;26(1):83–6.

426. Greipp J. Swimmer's Shoulder: The influence of flexibity and weight training. Phys Sportsmed. 1985;13(8):92–105.

427. Beach, M.L., Whitney, S.L., Dickoff-Hoffman S. Relationship of shoulder flexibility, strength, and endurance to shoulder pain in competitive swimmers. J Orthop Sports Phys Ther. 1992 Jan;16(6):262–8.

428. Stocker D, Pink M, Jobe F. Comparison of shoulder injury in collegiate-and master's-level swimmers. Clinical Journal of Sport Medicine. 1995;

429. Krüger PE, Dressler A, Botha M. Incidence of shoulder injuries and related risk factors among master swimmers in South Africa. Afr J Phys Health Educ Recreat Dance. 2012;(supplement):57–76.

430. Wymore, L., Reeve, R.E., Chaput CD. No correlation between stroke specialty and rate of shoulder pain in NCAA men swimmers. Int J Shoulder Surg. 2012;6(3):71–5.

431. Bansal S, Sinha AGK, Sandhu JS. Shoulder impingement syndrome among competitive swimmers in India - Prevalence, evaluation and risk factors. J Exerc Sci Fit. 2007;5(2):102–8.

432. Puckree, T., Thomas KJ. Shoulder injuries in competitive swimmers in KwaZulu-Natal. South African Journal of Sports Medicine. 2006;18(1):10–2.

433. Rodeo SA, Nguyen JT, Cavanaugh JT, Patel Y, Adler RS. Clinical and Ultrasonographic Evaluations of the Shoulders of Elite Swimmers. American Journal of Sports Medicine [Internet]. 2016;44(12):3214–21. Available from: https://doi.org/10.1177/0363546516657823

434. Welbeck AN, Amilo NR, Le DT, Killelea CM, Kirsch AN, Zarzour RH, et al. Examining the link between thoracic rotation and scapular dyskinesis and shoulder pain amongst college swimmers. Physical Therapy in Sport. 2019;40:78–84.

435. Boettcher C, Halaki M, Holt K, Ginn KA. Is the Normal Shoulder Rotation Strength Ratio Altered in Elite Swimmers? Med Sci Sports Exerc. 2020;52(3):680–4.

436. McLaine SJ, Ginn KA, Fell JW, Bird ML. Scapular upward rotation position is symmetrical in swimmers without current shoulder pain. Physical Therapy in Sport [Internet]. 2018;29:9–13. Available from: https://doi.org/10.1016/j.ptsp.2017.09.003

437. McLaine SJ, Bird ML, Ginn KA, Hartley T, Fell JW. Shoulder extension strength: a potential risk factor for shoulder pain in young swimmers? J Sci Med Sport [Internet]. 2019;22(5):516–20. Available from: https://doi.org/10.1016/j.jsams.2018.11.008

438. de Almeida MO, Hespanhol LC, Lopes AD. PREVALENCE OF MUSCULOSKELETAL PAIN AMONG SWIMMERS IN AN ELITE NATIONAL TOURNAMENT. Int J Sports Phys Ther [Internet]. 2015 Dec;10(7):1026–34. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26676276

439. Wymore L, Fronek J. Shoulder Functional Performance Status of National Collegiate Athletic Association Swimmers: Baseline Kerlan-Jobe Orthopedic Clinic Scores. American Journal of Sports Medicine. 2015;43(6):1513–7.

440. McLaine SJ, Ginn KA, Fell JW, Bird ML. Isometric shoulder strength in young swimmers. J Sci Med Sport [Internet]. 2018;21(1):35–9. Available from: https://doi.org/10.1016/j.jsams.2017.05.003

441. Zemek, M.J., Magee DJ. Comparison of glenohumeral joint laxity in elite and recreational swimmers. Clin J Sport Med. 1996;6(1):40–7.

442. Bak K, Magnusson SP. Shoulder Strength and Range of Motion in Symptomatic and Pain-Free Elite Swimmers. Am J Sports Med. 1997 Jul;25(4):454–9.

443. Santos MJ, Belangero WD, Almeida GL. The effect of joint instability on latency and recruitment order of the shoulder muscles. Journal of Electromyography and Kinesiology. 2007;17(2):167–75.

444. Su KPE, Johnson MP, Gracely EJ, Karduna AR. Scapular Rotation in Swimmers with and without Impingement Syndrome: Practice Effects. Med Sci Sports Exerc. 2004 Jul;36(7):1117–23.

445. Hidalgo-Lozano A. Elite Swimmers With Unilateral Shoulder Pain Demonstrate Altered Pattern of Cervical Muscles Activation During a Functional Upper Limb Task. Journal of Orthopaedic and Sports Physical Therapy. 2012;42(6):552–9.

446. Hidalgo-Lozano A, Fernández-de-las-Peñas C, Calderón-Soto C, Domingo-Camara A, Madeleine P AMM. Elite swimmers with and without unilateral shoulder pain: mechanical hyperalgesia and active/latent muscle trigger points in neck-shoulder muscles. Scand J Med Sci Sports. 2013;23(1):66–73.

447. Ozcaldiran B. A relation between static flexibility and shoulder pain in competitive age group swimmers. The Pain Clinic. 2002;14(2):159–63.

448. Vila Dieguez O, Barden JM. Body roll differences in freestyle swimming between swimmers with and without shoulder pain. Sports Biomech [Internet]. 2020;00(00):1–14. Available from: https://doi.org/10.1080/14763141.2020.1760923

449. Sabzehparvar E, Khaiyat OA, Ganji Namin B, Minoonejad H. Electromyographic analysis in elite swimmers with shoulder pain during a functional task. Sports Biomech [Internet]. 2019;00(00):1–11. Available from: https://doi.org/10.1080/14763141.2019.1597153

450. McKenna LJ, de Ronde M, Le M, Burke W, Graves A, Williams SA. Measurement of muscle thickness of the serratus anterior and lower trapezius using ultrasound imaging in competitive recreational adult swimmers, with and without current shoulder pain. J Sci Med Sport [Internet]. 2018;21(2):129–33. Available from: https://doi.org/10.1016/j.jsams.2017.06.022

451. Suzuki Y, Maeda N, Sasadai J, Kaneda K, Shirakawa T, Urabe Y. Ultrasonographic evaluation of the shoulders and its associations with shoulder pain, age, and swim training in masters swimmers. Medicina (Lithuania). 2021;57(1):1–13.

452. Wolf BR, Ebinger AE, Lawler MP, Britton CL. Injury patterns in Division I collegiate swimming. American Journal of Sports Medicine. 2009;37(10):2037–42.

453. Tate A, Harrington S, Buness M, Murray S, Trout C, Meisel C. Investigation of In-Water and Dry-Land training programs for competitive swimmers in the United States. J Sport Rehabil. 2015;24(4):353–62.

454. Borstad JD, Ludewig PM. The Effect of Long Versus Short Pectoralis Minor Resting Length on Scapular Kinematics in Healthy Individuals. Journal of Orthopaedic & Sports Physical Therapy [Internet]. 2005 Apr;35(4):227–38. Available from: http://www.jospt.org/doi/10.2519/jospt.2005.35.4.227

455. Scovazzo, M.L., Browne, A., Pink, M., Jobe, F.W., Kerrigan J. The painful shoulder during freestyle swimming. An electromyographic cinematographic analysis of twelve muscles. Am J Sports Med. 1991;19(6):577–82.

456. Ruwe, P.A., Pink, M., Jobe, F.W., Perry, J., Scovazzo ML, Ruwe P, Pink M, Jobe F. The normal and the painful shoulders during the breaststroke. Electromyographic and cinematographic analysis of twelve muscles. Am J Sports Med. 1994;22(6):789–96.

457. Wadsworth, D.J., Bullock-Saxton JE. Recruitment patterns of the scapular rotator muscles in freestyle swimmers with subacromial impingement. Int J Sports Med. 1997;18(8):618–24.

458. Yoma M, Herrington L, MacKenzie TA, Almond TA. Training intensity and shoulder musculoskeletal physical quality responses in competitive swimmers. J Athl Train. 2021;56(1):54–63.

459. Marques MC, Yáñez-García JM, Marinho DA, González-Badillo JJ, Rodríguez-Rosell D. In-Season Strength Training in Elite Junior Swimmers: The Role of the Low-Volume, High-Velocity Training on Swimming Performance. J Hum Kinet [Internet]. 2020 Aug;74:71–84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/33312277

460. Pollen TR, Ebaugh D, Warren M, Milner CE, Taylor JA, Silfies SP. The Relationship Between Workload and Non-Contact Musculoskeletal Injury in College Swimmers: A Prospective Cohort Study. J Athl Train [Internet]. 2021 Jun 15; Available from: https://meridian.allenpress.com/jat/article/doi/10.4085/1062-6050-0135.21/466428/The-Relationship-Between-Workload-and-Non-Contact

461. Gabbett T, Sancho I, Dingenen B, Willy RW. When progressing training loads, what are the considerations for healthy and injured athletes? Br J Sports Med. 2021;(April).

462. Jones J V, Pyne DB, Haff GG, Newton RU. Comparison of ballistic and strength training on swimming turn and dry-land leg extensor characteristics in elite swimmers. Int J Sports Sci Coach [Internet]. 2018 Apr 25;13(2):262–9. Available from: http://journals.sagepub.com/doi/10.1177/1747954117726017

463. Boltz AJ, Robison HJ, Morris SN, D'Alonzo BA, Collins CL, Chandran A, et al. Epidemiology of Injuries in National Collegiate Athletic Association Men's Swimming and Diving: 2014–2015 Through 2018–2019. J Athl Train [Internet]. 2021 Jul 1;56(7):711–8. Available from: https://meridian.allenpress.com/jat/article/56/7/711/467958/Epidemiology-of-Injuries-in-National-Collegiate

464. Lucas D, Neiva H, Marinho D, Ferraz R, Rolo I, Duarte-Mendes P. Functional Movement Screen® evaluation: comparison between elite and non-elite young swimmers. Cuadernos de Psicología del Deporte [Internet]. 2021 Apr 20;21(2):163–73. Available from: https://revistas.um.es/cpd/article/view/438401

465. Dominguez RH. Shoulder Pain in Swimmers. Phys Sportsmed. 1980 Jul;8(7):35–42.

466. Dorssen E van, Whiteley R, Mosler A, Cebrian SO, Dijkstra P. Shoulder Injuries in Sport. Sports Rehabilitation and Injury Prevention. 2020;3(3):571–80.

467. McMaster WC, Troup J. A survey of interfering shoulder pain in United States competitive swimmers. Am J Sports Med. 1993 Jan;21(1):67–70.

468. Allegrucci M, Whitney SL, Irrgang JJ. Clinical Implications of Secondary Impingement of the Shoulder in Freestyle Swimmers. Journal of Orthopaedic & Sports Physical Therapy. 1994 Dec;20(6):307–18.

469. Alakhdar Y, Cook J, Gallego D, Querol F, Chulvi-Medrano I, Alberola A, et al. Association Between COL5a1, COL11a1, and COL11a2 Gene Variations and Rotator Cuff Tendinopathy in Young Athletes. Clinical Journal of Sport Medicine. 2021;Publish Ah(00).

470. Little J, Higgins JPT, Ioannidis JPA, Moher D, Gagnon F, von Elm E, et al. Strengthening the reporting of genetic association studies (STREGA): An extension of the STROBE Statement. Hum Genet. 2009;125(2):131–51.

471. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. International Journal of Surgery [Internet]. 2014;12(12):1495–9. Available from: http://dx.doi.org/10.1016/j.ijsu.2014.07.013

472. General Assembly of the World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. J Am Coll Dent. 2014;81(3):14–8.

473. Kolz CW, Suter T, Henninger HB. Regional mechanical properties of the long head of the biceps tendon. Clinical Biomechanics [Internet]. 2015;30(9):940–5. Available from: http://dx.doi.org/10.1016/j.clinbiomech.2015.07.005

474. Apple AE, O'Brien MJ, Savoie FH. Long Head of the Biceps Tendinopathy. In: Muscle and Tendon Injuries. Berlin, Heidelberg: Springer Berlin Heidelberg; 2017. p. 215–21.

475. Singaraju VM, Kang RW, Yanke AB, McNickle AG, Lewis PB, Wang VM, et al. Biceps tendinitis in chronic rotator cuff tears: A histologic perspective. J Shoulder Elbow Surg. 2008 Nov;17(6):898–904.

476. Redondo-Alonso L, Chamorro-Moriana G, Jiménez-Rejano JJ, López-Tarrida P, Ridao-Fernández C. Relationship between chronic pathologies of the supraspinatus tendon and the long head of the biceps tendon: systematic review. BMC Musculoskelet Disord [Internet]. 2014 Nov 18;15(4):377. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25408141

477. Burkhart SS, Morgan CD, Kibler W Ben. The disabled throwing shoulder: Spectrum of pathology part III: The SICK scapula, scapular dyskinesis, the kinetic chain, and rehabilitation. Arthroscopy - Journal of Arthroscopic and Related Surgery. 2003;19(6):641–61.

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

479. Mokone GG, Gajjar M, September A V., Schwellnus MP, Greenberg J, Noakes TD, et al. The guanine-thymine dinucleotide repeat polymorphism within the tenascin-C gene is associated with Achilles tendon injuries. American Journal of Sports Medicine. 2005;33(7):1016–21.

480. Warnes G, Gorjanc G, Leisch F, Al. E. Genetics: a package for population genetics. R Package (version 1.3.6). 2011;

481. González JR, Armengol L, Solé X, Guinó E, Mercader JM, Estivill X, et al. SNPassoc: An R package to perform whole genome association studies. Bioinformatics. 2007 Mar;23(5):644–5.

482. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score Tests for Association between Traits and Haplotypes when Linkage Phase Is Ambiguous. The American Journal of Human Genetics. 2002 Feb;70(2):425–34.

483. Sinnwell JP, Schaid DJ. haplo.stats: {S}tatistical {A}nalysis of {H}aplotypes with {T}raits and {C}ovariates when {L}inkage {P}hase is {A}mbiguous. 2005.

484. Nyholt DR. A Simple Correction for Multiple Testing for Single-Nucleotide Polymorphisms in Linkage Disequilibrium with Each Other. The American Journal of Human Genetics. 2004 Apr;74(4):765–9.

485. Perneger T. What's wrong with Bonferroni adjustments. Br Med J. 1998;316:1236–8.

486. Rahim M, Mannion S, Klug B, Hobbs H, van der Merwe W, Posthumus M, et al. Modulators of the extracellular matrix and risk of anterior cruciate ligament ruptures. J Sci Med Sport [Internet]. 2017;20(2):152–8. Available from: http://dx.doi.org/10.1016/j.jsams.2016.07.003

487. O'connell K, Posthumus M, Collins M. No association between COL3A1, COL6A1 or COL12A1 gene variants and range of motion. J Sports Sci. 2013;31(2):181–7.

488. Niyibizi C, Kavalkovich K, Yamaji T, Woo SLY. Type V collagen is increased during rabbit medial collateral ligament healing. Knee Surgery, Sports Traumatology, Arthroscopy. 2000 Sep;8(5):281–5.

489. Laguette MJN, Collins M, Prince S. Characterisation of the 3'-UTR of the COL5A1 gene: Implications for musculoskeletal soft tissue injuries. 2014;286. Available from: https://open.uct.ac.za/bitstream/item/16805/thesis_hsf_2015_laguette_mary_jessica_nancy.pdf?sequence=1

490. Laguette MJN, Barrow K, Firfirey F, Dlamini S, Saunders CJ, Dandara C, et al. Exploring new genetic variants within COL5A1 intron 4-exon 5 region and TGF-β family with risk of anterior cruciate ligament ruptures. Journal of Orthopaedic Research. 2020;38(8):1856–65.

491. Bray RC, Salo PT, Lo IK, Ackermann P, Rattner JB, Hart DA. Normal ligament structure, physiology and function. Sports Med Arthrosc Rev. 2005;13(3):127–35.

492. Matsuura Y, Hangai M, Koizumi K, Ueno K, Hirai N, Akuzawa H, et al. Injury trend analysis in the Japan national swim team from 2002 to 2016: Effect of the lumbar injury prevention project. BMJ Open Sport Exerc Med. 2019;5(1):1–8.

493. Porter KN, Blanch PD, Walker HM, Shield AJ. The effect of previous shoulder pain on supraspinatus tendon thickness changes following swimming practice. Scand J Med Sci Sports. 2020;30(8):1442–8.

494. Porter KN, Talpey S, Pascoe D, Blanch PD, Walker HM, Shield AJ. The effect of swimming volume and intensity on changes in supraspinatus tendon thickness. Physical Therapy in Sport [Internet]. 2021;47:173–7. Available from: https://doi.org/10.1016/j.ptsp.2020.12.002

495. Hibberd EE, Laudner KG, Kucera KL, Berkoff DJ, Yu B, Myers JB. Effect of Swim Training on the Physical Characteristics of Competitive Adolescent Swimmers. American Journal of Sports Medicine [Internet]. 2016;44(11):2813–9. Available from: https://doi.org/10.1177/0363546516669506

496. Matzkin E, Suslavich K, Wes D. Swimmer's shoulder: Painful shoulder in the competitive swimmer. Vol. 24, Journal of the American Academy of Orthopaedic Surgeons. 2016. p. 527–36.

497. Factor D, Dale B. Current Concepts of Rotator Cuff Tendinopathy. Int J Sports Phys Ther. 2014;9(2):274–88.

498. Cook JL, Khan KM, Kiss ZS, Griffiths L. Patellar tendinopathy in junior basketball players: a controlled clinical and ultrasonographic study of 268 patellar tendons in players aged 14-18 years. Scand J Med Sci Sports. 2000 Aug;10(4):216–20.

499. John R, Prabhakar S, Singh Dhillon M, Anand A, Minhas G. Association of ACL tears and single nucleotide polymorphisms in the collagen 12 A1 gene in the Indian population - a preliminary case-control study. Muscles Ligaments Tendons J [Internet]. 2016;6(2):253–7. Available from: http://www.mltj.org/common/php/portiere.php?ID=c3ef28d7c2656a146e0e0fe0a6112b16

500. Gibbon A, Raleigh SM, Ribbans WJ, Posthumus M, Collins M, September A V. Functional COL1A1 variants are associated with the risk of acute musculoskeletal soft tissue injuries. Journal of Orthopaedic Research [Internet]. 2020;38(10):2290–8. Available from: http://dx.doi.org/10.1002/jor.24621

501. Stępień-Słodkowska M, Ficek K, Eider J, Leońska-Duniec A, Maciejewska-Karłowska A, Sawczuk M, et al. THE +1245G/T POLYMORPHISMS IN THE COLLAGEN TYPE I ALPHA 1 (COL1A1) GENE IN POLISH SKIERS WITH ANTERIOR CRUCIATE LIGAMENT INJURY. Biol Sport [Internet]. 2013 Jan 21;30(1):57–60. Available from: http://biolsport.com/abstracted.php?level=5&ICID=1029823

502. Alvarez-Romero J, Laguette MJN, Seale K, Jacques M, Voisin S, Hiam D, et al. Genetic variants within the COL5A1 gene are associated with ligament injuries in physically active populations from Australia, South Africa, and Japan. Eur J Sport Sci [Internet]. 2021 Nov 25;1–21. Available from: https://www.tandfonline.com/doi/full/10.1080/17461391.2021.2011426

503. Sivertsen EA, Haug KBF, Kristianslund EK, Trøseid AMS, Parkkari J, Lehtimäki T, et al. No Association Between Risk of Anterior Cruciate Ligament Rupture and Selected Candidate Collagen Gene Variants in Female Elite Athletes From High-Risk Team Sports. American Journal of Sports Medicine. 2019;47(1):52–8.

504. Tengman E, Grip H, Stensdotter A, Häger CK. Anterior cruciate ligament injury about 20 years post-treatment: A kinematic analysis of one-leg hop. Scand J Med Sci Sports. 2015;25(6):818–27.

505. Stensdotter AK, Tengman E, Olofsson LB, Häger C. Deficits in single-limb stance more than 20 years after ACL injury. Eur J Physiother. 2013;15(2):78–85.

506. Tengman E, Brax Olofsson L, Nilsson KG, Tegner Y, Lundgren L, Häger CK. Anterior cruciate ligament injury after more than 20 years: I. Physical activity level and knee function. Scand J Med Sci Sports. 2014;24(6):e491-500.

507. Suijkerbuijk MAM, Ponzetti M, Rahim M, Posthumus M, Häger CK, Stattin E, et al. Functional polymorphisms within the inflammatory pathway regulate expression of extracellular matrix components in a genetic risk dependent model for anterior cruciate ligament injuries. J Sci Med Sport [Internet]. 2019;22(11):1219–25. Available from: https://doi.org/10.1016/j.jsams.2019.07.012

508. Lulińska-Kuklik E, Rahim M, Domańska-Senderowska D, Ficek K, Michałowska-Sawczyn M, Moska W, et al. Interactions Between COL5A1 Gene and Risk of the Anterior Cruciate Ligament Rupture. J Hum Kinet [Internet]. 2018 Jun 13;62(1):65–71. Available from: https://content.sciendo.com/view/journals/hukin/62/1/article-p65.xml

509. L H. Retrospective Analysis of Non-Contact ACL Injury Risk: A Case Series Review of Elite Female Athletes. Journal of Sports Medicine and Therapy. 2017;2(1):001–5.

510. Myer GD, Heidt RS, Waits C, Finck S, Stanfield D, Posthumus M, et al. Sex comparison of familial predisposition to anterior cruciate ligament injury. Knee Surgery, Sports Traumatology, Arthroscopy. 2014;22(2):387–91.

511. Ahn HS, Lee DH, Kazmi SZ, Kang T, Lee YS, Sung R, et al. Familial Risk and Its Interaction With Body Mass Index and Physical Activity in Anterior Cruciate Ligament Injury Among First-Degree Relatives: A Population-Based Cohort Study. American Journal of Sports Medicine. 2021;49(12):3312–21.

512. Magnusson K, Turkiewicz A, Hughes V, Frobell R, Englund M. High genetic contribution to anterior cruciate ligament rupture: Heritability ~69%. Br J Sports Med. 2021;55(7):385–9.

513. Gibbon A, Saunders CJ, Collins M, Gamieldien J, September A V. Defining the molecular signatures of achilles tendinopathy and anterior cruciate ligament ruptures: A whole-exome sequencing approach. PLoS One. 2018;13(10):1–20.

514. Gerecke DR, Olson PF, Koch M, Knoll JHM, Taylor R, Hudson DL, et al. Complete primary structure of two splice variants of collagen XII, and assignment of α1(XII) collagen (COL12A1), α1(IX) collagen (COL9A1), and α1(XIX) collagen (COL19a1) to human chromosome 6q12-q13. Vol. 41, Genomics. 1997. p. 236–42.

515. Izu Y, Ezura Y, Koch M, Birk DE, Noda M. Collagens VI and XII form complexes mediating osteoblast interactions during osteogenesis. Cell Tissue Res [Internet]. 2016;364(3):623–35. Available from: http://dx.doi.org/10.1007/s00441-015-2345-y

516. Shukla M, Gupta R, Pandey V, Tiwari PK, Amrathlal RS. COLIA1 + 1245 G > T Sp1 Binding Site Polymorphism is Not Associated with ACL Injury Risks Among Indian Athletes. Indian J Orthop [Internet]. 2020 Sep 28;54(5):647–54. Available from: https://doi.org/10.1007/s43465-020-00119-1

517. Kang BY, Eun SP, Kim H, Lee JK. Distributions of genetic polymorphisms in COL12A1 and MMP12 genes in Korean patients with anterior cruciate ligament rupture by recreational sports activities. Medicina dello Sport [Internet]. 2020 Jan;72(4). Available from: https://www.minervamedica.it/index2.php?show=R26Y2019N04A0667

518. Rahim M, Hobbs H, van der Merwe W, Posthumus M, Collins M, September A V. Investigation of angiogenesis genes with anterior cruciate ligament rupture risk in a South African population. J Sports Sci [Internet]. 2018;36(5):551–7. Available from: https://doi.org/10.1080/02640414.2017.1322710

519. Rahim M, Gibbon A, Collins M, September A V. Genetics of musculoskeletal soft tissue injuries: Current status, challenges, and future directions [Internet]. Sports, Exercise, and Nutritional Genomics: Current Status and Future Directions. Elsevier Inc.; 2019. 317–339 p. Available from: http://dx.doi.org/10.1016/B978-0-12-816193-7.00015-4

520. Roberts L, Greenberg J, Ramsay M, Soodyall H. South African Society for Human Genetics. 2013. Guidelines from the SASHG Committee for publication purposes regarding: Nomenclature for South African populations.

521. CHEN H, SHI L, ZHANG J, ZHAO D, SHI L, LU Q, et al. Association of COL1A1, COL5A1 and COL12A1 genes with anterior cruciate ligament injury 前交叉韧带损伤与胶原蛋白基因COL1A1、COL5A1及COL12A1的相关性. Chinese Journal of Tissue Engineering Research. 2017;21(12):1849–54.

522. Prabhakar S, John R, Dhillon MS, Anand A, Sharma K, Bammidi S. Are COL1A1 gene polymorphisms associated with anterior cruciate ligament tear in the Indian population? Results of a preliminary case-control study. Muscles Ligaments Tendons J. 2018;8(1):15–22.

523. Flück M, Giraud MN, Tunç V, Chiquet M. Tensile stress-dependent collagen XII and fibronectin production by fibroblasts requires separate pathways. Biochim Biophys Acta Mol Cell Res. 2003;1593(2–3):239–48.

524. Kubo K, Yata H, Tsunoda N. Effect of gene polymorphisms on the mechanical properties of human tendon structures. Springerplus. 2013;2(1):1–6.

525. Orth T, Paré J, Froehlich JE. Current Concepts on the Genetic Factors in Rotator Cuff Pathology and Future Implications for Sports Physical Therapists. Int J Sports Phys Ther [Internet]. 2017;12(2):273–85. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28515982%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5380870

526. Astur DC, Novaretti JV, Cohen M. Genetic and molecular factors and anterior cruciate ligament injuries: current concepts. J ISAKOS. 2017;2(3):123–6.

527. Astur DC, Andrade E, Arliani GG, Debieux P, Loyola LC, dos Santos SEB, et al. The potential European genetic predisposition for non-contact anterior cruciate ligament injury. Knee Surgery, Sports Traumatology, Arthroscopy [Internet]. 2018 Dec 4;26(12):3532–6. Available from: http://link.springer.com/10.1007/s00167-018-4974-7

528. Miyamoto-Mikami E, Miyamoto N, Kumagai H, Hirata K, Kikuchi N, Zempo H, et al. COL5A1 rs12722 polymorphism is not associated with passive muscle stiffness and sports-related muscle injury in japanese athletes. BMC Med Genet. 2019;20(1):1–9.

529. Brazier J, Antrobus M, Stebbings GK, Day SH, Heffernan SM, Cross MJ, et al. Tendon and Ligament Injuries in Elite Rugby: The Potential Genetic Influence. Sports. 2019;7(6):138.

530. Arbeeva L, Yau M, Mitchell BD, Jackson RD, Ryan K, Golightly YM, et al. Genome-wide meta-analysis identified novel variant associated with hallux valgus in Caucasians. J Foot Ankle Res. 2020;13(1):1–11.

531. Wang P, Liu X, Zhu B, Ma Y, Yong L, Teng Z, et al. Association of IL17RC and COL6A1 genetic polymorphisms with susceptibility to ossification of the thoracic posterior longitudinal ligament in Chinese patients. J Orthop Surg Res. 2018;13(1):1–6.

532. Koga H, Hayashi K, Taketomi E, Matsunaga S, Yashiki S, Fujiyoshi T, et al. Restriction Fragment Length Polymorphism of Genes of the α2(XI) Collagen, Bone Morphogenetic Protein-2, Alkaline Phosphatase, and Tumor Necrosis Factor-α among Patients with Ossification of Posterior Longitudinal Ligament and Controls from the Japanese Po. Vol. 21, Spine. 1996. p. 469–73.

533. Ryan-Moore E, Mavrommatis Y, Waldron M. Systematic Review and Meta-Analysis of Candidate Gene Association Studies With Fracture Risk in Physically Active Participants. Front Genet. 2020 Jun 16;11.

534. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009 Jul 25;4(7):1073–81.

535. Wang P, Teng Z, Liu X, Liu X, Kong C, Lu S. The COL6A1 rs201153092 single nucleotide polymorphism, associates with thoracic ossification of the posterior longitudinal ligament. Mol Med Rep. 2020;21(1):191–200.

536. Emery CA. Risk Factors for Injury in Child and Adolescent Sport: A Systematic Review of the Literature. Clinical Journal of Sport Medicine [Internet]. 2003 Jul;13(4):256–68. Available from: http://journals.lww.com/00042752-200307000-00011

537. Toubekis AG, Drosou E, Gourgoulis V, Thomaidis S, Douda H, Tokmakidis SP. Competitive performance, training load and physiological responses during tapering in young swimmers. J Hum Kinet. 2013;38(1):125–34.

538. Bell RD, Shultz SJ, Wideman L, Henrich VC. Collagen Gene Variants Previously Associated With Anterior Cruciate Ligament Injury Risk Are Also Associated With Joint Laxity. Sports Health. 2012;4(4):312–8.

539. Barry L, Lyons M, McCreesh K, Powell C, Comyns T. The relationship between training load and pain, injury and illness in competitive swimming: A systematic review. Phys Ther Sport [Internet]. 2021 Mar;48:154–68. Available from: http://www.ncbi.nlm.nih.gov/pubmed/33486408

540. Jones CM, Griffiths PC, Mellalieu SD. Training Load and Fatigue Marker Associations with Injury and Illness: A Systematic Review of Longitudinal Studies. Vol. 47, Sports Medicine. Springer International Publishing; 2017. 943–974 p.

541. McKenzie A, Larequi S, Hams A, Headrick J, Whiteley R, Duhig S. Shoulder pain and injury risk factors in competitive swimmers: A systematic review. Scand J Med Sci Sports. 2023 Dec 28;33(12):2396–412.

542. McLaine SJ, Bird ML, Ginn KA, Hartley T, Fell JW. Shoulder extension strength: a potential risk factor for shoulder pain in young swimmers? J Sci Med Sport. 2019 May;22(5):516–20.

543. Higson E, Herrington L, Butler C, Horsley I. The short-term effect of swimming training load on shoulder rotational range of motion, shoulder joint position sense and pectoralis minor length. Shoulder Elbow. 2018;10(4):285–91.

544. Hibberd EE, Oyama S, Spang JT, Prentice W, Myers JB. Effect of a 6-week strengthening program on shoulder and scapular-stabilizer strength and scapular kinematics in division I collegiate swimmers. J Sport Rehabil [Internet]. 2012 Aug;21(3):253–65. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22387875

545. Matsuura Y, Fujimoto T, Koizumi K, Mise T, Edama M, Shimoyama Y. Posterior deltoid shoulder tightness and greater contralateral lower limb muscle strength are associated with swimmers' shoulder pain. J Sci Med Sport. 2023 Dec;26(12):694–9.

546. Collins M, September A V. Are commercial genetic injury tests premature? Scand J Med Sci Sports [Internet]. 2023 Sep 27;33(9):1584–97. Available from: https://onlinelibrary.wiley.com/doi/10.1111/sms.14406

547. McAuley ABT, Hughes DC, Tsaprouni LG, Varley I, Suraci B, Roos TR, et al. A Systematic Review of the Genetic Predisposition to Injury in Football. Journal of Science in Sport and Exercise. 2023 May 28;5(2):97–115.

548. Vlahovich N, Hughes DC, Griffiths LR, Wang G, Pitsiladis YP, Pigozzi F, et al. Genetic testing for exercise prescription and injury prevention: AIS-Athlome consortium-FIMS joint statement. BMC Genomics. 2017 Nov 14;18(S8):818.

549. Goodlin GT, Roos TR, Roos AK, Kim SK. The Dawning Age of Genetic Testing for Sports Injuries. Clinical Journal of Sport Medicine. 2015 Jan;25(1):1–5.

550. Varillas-Delgado D, Del Coso J, Gutiérrez-Hellín J, Aguilar-Navarro M, Muñoz A, Maestro A, et al. Genetics and sports performance: the present and future in the identification of talent for sports based on DNA testing. Eur J Appl Physiol. 2022 Aug 16;122(8):1811–30.

551. Patel S, Varley I. Exploring the Regulation of Genetic Testing in Sport. Entertainment and Sports Law Journal. 2019 Apr 1;17(1).

552. Vlahovich N, Fricker PA, Brown MA, Hughes D. Ethics of genetic testing and research in sport: a position statement from the Australian Institute of Sport. Br J Sports Med. 2017 Jan;51(1):5–11.

553. Wells D. Genetic engineering in athletes. BMJ. 2008 Jul 7;337(jul07 1):a607–a607.

554. Pickering C, Kiely J. The Development of a Personalised Training Framework: Implementation of Emerging Technologies for Performance. J Funct Morphol Kinesiol. 2019 May 16;4(2):25.

555. Jaskulski S, Nuszbaum C, Michels KB. Components, prospects and challenges of personalized prevention. Front Public Health. 2023 Feb 16;11.

556. Semenova EA, Hall ECR, Ahmetov II. Genes and Athletic Performance: The 2023 Update. Genes (Basel). 2023 Jun 8;14(6):1235.

557. Guth LM, Roth SM. Genetic influence on athletic performance. Curr Opin Pediatr. 2013 Dec;25(6):653–8.

558. Johnson JS, Morscher MA, Jones KC, Moen SM, Klonk CJ, Jacquet R, et al. Gene expression differences between ruptured anterior cruciate ligaments in young male and female subjects. Journal of Bone and Joint Surgery - American Volume. 2015;97(1):71–9.

559. Kim HC, Park KJ. Injuries in male and female elite aquatic sports athletes: An 8-year prospective, epidemiological study. J Sports Sci Med. 2020;19(2):390–6.

560. Shultz SJ, Schmitz RJ, Benjaminse A, Collins M, Ford K, Kulas AS. ACL research retreat VII: An update on anterior cruciate ligament injury risk factor identification, screening, and prevention March 19-21, 2015; Greensboro, nc. J Athl Train. 2015;50(10):1076–93.

561. Hägglund M, Waldén M. Risk factors for acute knee injury in female youth football. Knee Surgery, Sports Traumatology, Arthroscopy. 2016;24(3):737–46.

562. Åman M, Forssblad M, Larsén K. Incidence and body location of reported acute sport injuries in seven sports using a national insurance database. Scand J Med Sci Sports. 2018;28(3):1147–58.

563. Seale K, Burger M, Posthumus M, Häger CK, Stattin E, Nilsson KG, et al. The Apoptosis Pathway and CASP8 Variants Conferring Risk for Acute and Overuse Musculoskeletal Injuries. Journal of Orthopaedic Research [Internet]. 2020;38(3):680–8. Available from: http://dx.doi.org/10.1002/jor.24504

564. September A, Posthumus M. Application of Genomics in the Prevention, Treatment and Management of Achilles Tendinopathy and Anterior Cruciate Ligament Ruptures. Recent Pat DNA Gene Seq [Internet]. 2012 Aug 1;6(3):216–23. Available from: http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1872-2156&volume=6&issue=3&spage=216

565. Posthumus M, September A V., Keegan M, O'Cuinneagain D, Van Der Merwe W, Schwellnus MP, et al. Genetic risk factors for anterior cruciate ligament ruptures: COL1A1 gene variant. Br J Sports Med. 2009;43(5):352–6.

566. Ficek K, Cieszczyk P, Kaczmarczyk M, Maciejewska-Karłowska A, Sawczuk M, Cholewinski J, et al. Gene variants within the COL1A1 gene are associated with reduced anterior cruciate ligament injury in professional soccer players. J Sci Med Sport. 2013;16(5):396–400.

567. John R, Dhillon MS, Prabhakar S AAA. Col1a1 and Col12a1 With Acl Tears- a Association of Gene Polymorphisms in Col1a1 and Col12a1 With Acl Tears - a Study in the Indian. International Journal of Conference Proceedings. 2015;2015(112).

568. Sivertsen EA, Haug KBF, Kristianslund EK, Trøseid AMS, Parkkari J, Lehtimäki T, et al. No Association Between Risk of Anterior Cruciate Ligament Rupture and Selected Candidate Collagen Gene Variants in Female Elite Athletes From High-Risk Team Sports. American Journal of Sports Medicine. 2019;47(1):52–8.

569. Shukla M, Gupta R, Pandey V, Tiwari PK, Amrathlal RS. COLIA1 + 1245 G > T Sp1 Binding Site Polymorphism is Not Associated with ACL Injury Risks Among Indian Athletes. Indian J Orthop. 2020 Sep 28;54(5):647–54.

570. Gibbon A, Raleigh SM, Ribbans WJ, Posthumus M, Collins M, September A V. Functional COL1A1 variants are associated with the risk of acute musculoskeletal soft tissue injuries. Journal of Orthopaedic Research. 2020;38(10):2290–8.

571. Zhao D, Zhang Q, Lu Q, Hong C, Luo T, Duan Q, et al. Correlations between the genetic variations in the COL1A1, COL5A1, COL12A1, and b-fibrinogen genes and anterior cruciate ligament injury in Chinese patients. J Athl Train. 2020;55(5):515–21.

572. Posthumus M, September A V., Schwellnus MP, Collins M. Investigation of the Sp1-binding site polymorphism within the COL1A1 gene in participants with Achilles tendon injuries and controls. J Sci Med Sport. 2009;12(1):184–9.

573. Dada S, Burger MC, Massij F, de Wet H, Collins M. Carpal tunnel syndrome: The role of collagen gene variants. Gene. 2016;587(1):53–8.

574. Khoschnau S, Melhus H, Jacobson A, Rahme H, Bengtsson H, Ribom E, et al. Type I Collagen α1 Sp1 Polymorphism and the Risk of Cruciate Ligament Ruptures or Shoulder Dislocations - 2008.pdf. Am J Sports Med. 2008 Dec;36(12):2432–6.

575. Vidal Rodríguez S, del Castillo I. COL1A1, COL5A1 AND GDF5 POLYMORPHISMS ASSOCIATED WITH ROTATOR CUFF DISEASE. Orthopaedic Proceedings. 2018;99-B(Supp 8).

576. Rodriguez SV. VDR, GDF5, COL1A1, THBS2 AND CHST GENE POLYMORPHISMS ASSOCIATED WITH SYMPTOMATIC LUMBAR DISC HERNIATION IN SPANISH SURGICAL PATIENTS. Orthopaedic Proceedings. 2017;99-B(Supp 1):119.

577. Pluijm SMF, Van Essen HW, Bravenboer N, Uitterlinden AG, Smit JH, Pols HAP, et al. Collagen type I α1 Sp1 polymorphism, osteoporosis, and intervertebral disc degeneration in older men and women. Ann Rheum Dis. 2004;63(1):71–7.

578. Tilkeridis C, Bei T, Garantziotis S, Stratakis CA. Association of a COL1A1 polymorphism with lumbar disc disease in young military recruits. J Med Genet. 2005;42(7):7–9.

579. Videman T, Saarela J, Kaprio J, Näkki A, Levälahti E, Gill K, et al. Associations of 25 structural, degradative, and inflammatory candidate genes with lumbar disc desiccation, bulging, and height narrowing. Arthritis Rheum. 2009;60(2):470–81.

580. Erduran M, Altinisik J, Meric G, Ates O, Ulusal AE, Akseki D. Is Sp1 binding site polymorphism within COL1A1 gene associated with tennis elbowα. Gene. 2014;537(2):308–11.

581. Näkki A, Videman T, Kujala UM, Suhonen M, Männikkö M, Peltonen L, et al. Candidate gene association study of magnetic resonance imaging-based hip osteoarthritis (OA): Evidence for COL9A2 gene as a common predisposing factor for hip OA and lumbar disc degeneration. Journal of Rheumatology. 2011;38(4):747–52.

582. Omair A, Lie BA, Reikeras O, Brox JI. An Association Study of Interleukin 18 Receptor Genes (IL18R1 and IL18RAP) in Lumbar Disc Degeneration. Open Orthop J. 2012;6(1):164–71.

583. Deng Y, Tan XT, Wu Q, Wang X. Correlations between COL2A and aggrecan genetic polymorphisms and the risk and clinicopathological features of intervertebral disc degeneration in a Chinese han population: A case-control study. Genet Test Mol Biomarkers. 2017;21(2):108–15.

584. O'Connell K, Knight H, Ficek K, Leonska-Duniec A, Maciejewska-Karlowska A, Sawczuk M, et al. Interactions between collagen gene variants and risk of anterior cruciate ligament rupture. Eur J Sport Sci. 2015;15(4):341–50.

585. Stepien-Slodkowska M, Ficek K, Maciejewska-Karłowska A, Sawczuk M, Ziętek P, Król P, et al. Overrepresentation of the COL3A1 AA genotype in Polish skiers with anterior cruciate ligament injury. Biol Sport. 2015;32(2):143–7.

586. Omair A, Mannion AF, Holden M, Leivseth G, Fairbank J, Hägg O, et al. Age and pro-inflammatory gene polymorphisms influence adjacent segment disc degeneration more than fusion does in patients treated for chronic low back pain. European Spine Journal. 2016;25(1):2–13.

587. Posthumus M, September AV V., O'cuinneagain D, Van Der Merwe W, Schwellnus MPP, Collins M, et al. The COL5A1 gene is associated with increased risk of Anterior cruciate ligament ruptures in female participants. American Journal of Sports Medicine. 2009 Nov;37(11):2234–40.

588. McCabe K, Collins C. Can Genetics Predict Sports Injury? The Association of the Genes GDF5, AMPD1, COL5A1 and IGF2 on Soccer Player Injury Occurrence. Sports. 2018;6(1):21.

589. September A V., Cook J, Handley CJ, Van Der Merwe L, Schwellnus MP, Collins M. Variants within the COL5A1 gene are associated with Achilles tendinopathy in two populations. Br J Sports Med. 2009;43(5):357–65.

590. Brown KL, Seale KB, El Khoury LY, Posthumus M, Ribbans WJ, Raleigh SM, et al. Polymorphisms within the COL5A1 gene and regulators of the extracellular matrix modify the risk of Achilles tendon pathology in a British case-control study. J Sports Sci. 2017;35(15):1475–83.

591. Burger M, de Wet H, Collins M. The COL5A1 gene is associated with increased risk of carpal tunnel syndrome. Clin Rheumatol. 2015;34(4):767–74.

592. Longo UG, Margiotti K, Petrillo S, Rizzello G, Fusilli C, Maffulli N, et al. Genetics of rotator cuff tears: No association of col5a1 gene in a case-control study. BMC Med Genet. 2018;19(1):1–5.

593. Alakhdar Y, Cook J, Gallego D, Querol F, Chulvi-Medrano I, Alberola A, et al. Association Between COL5a1, COL11a1, and COL11a2 Gene Variations and Rotator Cuff Tendinopathy in Young Athletes. Clinical Journal of Sport Medicine. 2021;Publish Ah(00).

594. Paassilta P, Lohiniva J, Göring HHH, Perälä M, Räinä SS, Karppinen J, et al. Identification of a novel common genetic risk factor for lumbar disk disease. J Am Med Assoc. 2001;285(14):1843–9.

595. Altinisik J, Meric G, Erduran M, Ates O, Ulusal AE, Akseki D. The BstUI and DpnII variants of the COL5A1 gene are associated with tennis elbow. American Journal of Sports Medicine. 2015;43(7):1784–9.

596. Alakhdar Mohmara Y, Cook J, Benítez-Martínez JC, McPeek ER, Aguilar AA, Olivas ES, et al. Influence of genetic factors in elbow tendon pathology: a case-control study. Sci Rep. 2020;10(1):1–7.

597. Alvarez-Romero J, Laguette MJN, Seale K, Jacques M, Voisin S, Hiam D, et al. Genetic variants within the COL5A1 gene are associated with ligament injuries in physically active populations from Australia, South Africa, and Japan. Eur J Sport Sci. 2021 Nov 25;1–21.

598. Abrahams Y, Laguette MJ, Prince S, Collins M. Polymorphisms within the COL5A1 3′-UTR That Alters mRNA Structure and the MIR608 Gene are Associated with Achilles Tendinopathy. Ann Hum Genet. 2013 May;77(3):204–14.

599. Laguette MJN, Barrow K, Firfirey F, Dlamini S, Saunders CJ, Dandara C, et al. Exploring new genetic variants within COL5A1 intron 4-exon 5 region and TGF-β family with risk of anterior cruciate ligament ruptures. Journal of Orthopaedic Research. 2020;38(8):1856–65.

600. O'Connell KS, Sean K, O'connell K. Investigation of selected collagen genes in exercise-related musculoskeletal soft tissue phenotypes. Thesis. University of Cape Town; 2014.

601. Kim KH, Kuh SU, Park JY, Lee SJ, Park HS, Chin DK, et al. Association between BMP-2 and COL6A1 gene polymorphisms with susceptibility to ossification of the posterior longitudinal ligament of the cervical spine in Korean patients and family members. Genetics and Molecular Research. 2014;13(1):2240–7.

602. Tanaka T, Ikahari K, Furushima K, Okada A, Harata S, Nakajima T, et al. Identification of Col6a1 As a Susceptible Gene for the Ossification of the Posterior Longitudinal Ligament of the Spine From the Linkage Region on Chromosome 21. 50th Annual Meeting of the Orthopaedic Research Society. 2004;1086.

603. Seki S, Kawaguchi Y, Mori M, Mio F, Chiba K, Mikami Y, et al. Association study of COL9A2 with lumbar disc disease in the Japanese population. J Hum Genet. 2006;51(12):1063–7.

604. Zhang Z, Zhang J, Ding L, Teng X. Meta-analysis of the association between COL9A2 genetic polymorphisms and lumbar disc disease susceptibility. Spine (Phila Pa 1976). 2014;39(20):1699–706.

605. Annunen S, Paassilta P, Lohiniva J, Perälä M, Karppinen J, Tervonen O, et al. An Allele of COL9A2 Associated with Intervertebral Disc Disease Published by : American Association for the Advancement of Science Stable URL : http://www.jstor.org/stable/2898848 digitize , preserve and extend access to Science An Allele of COL9A2 Associ. 2017;14–8.

606. Jim JJT, Noponen-Hietala N, Cheung KMC, Ott J, Karppinen J, Sahraravand A, et al. The TRP2 allele of COL9A2 is an age-dependent risk factor for the development and severity of intervertebral disc degeneration. Spine (Phila Pa 1976). 2005;30(24):2735–42.

607. Karppinen J, Pääkkö E, Räinä S, Tervonen O, Kurunlahti M, Nieminen P, et al. Magnetic resonance imaging findings in relation to the COL9A2 tryptophan allele among patients with sciatica. Spine (Phila Pa 1976). 2002;27(1):78–82.

608. Karppinen J, Pääkkö E, Paassilta P, Lohiniva J, Kurunlahti M, Tervonen O, et al. Radiologic Phenotypes in Lumbar MR Imaging for a Gene Defect in the COL9A3 Gene of Type IX Collagen. Radiology. 2003 Apr;227(1):143–8.

609. Eskola PJ, Kjaer P, Daavittila IM, Solovieva S, Okuloff A, Sorensen JS, et al. Genetic risk factors of disc degeneration among 12-14-year-old Danish children: A population study. Int J Mol Epidemiol Genet. 2010;1(2):158–65.

610. Hay M, Patricios J, Collins R, Branfield A, Cook J, Handley CJ, et al. Association of type XI collagen genes with chronic Achilles tendinopathy in independent populations from South Africa and Australia. Br J Sports Med. 2013;47(9):569–74.

611. Mio F, Chiba K, Hirose Y, Kawaguchi Y, Mikami Y, Oya T, et al. A functional polymorphism in COL11A1, which encodes the α1 chain of type XI collagen, is associated with susceptibility to.pdf. Am J Hum Genet. 2007;81(6):1271–7.

612. Kelempisioti A, Eskola PJ, Okuloff A, Karjalainen U, Takatalo J, Daavittila I, et al. Genetic susceptibility of intervertebral disc degeneration among young Finnish adults. BMC Med Genet. 2011;12(1):153.

613. Posthumus M, September AVA V., O'Cuinneagain D, Van Der Merwe W, Schwellnus MPMP, Collins M. The association between the COL12A1 gene and anterior cruciate ligament ruptures. Br J Sports Med. 2010 Dec;44(16):1160–5.

614. September AV V., Posthumus M, Van Der Merwe L, Schwellnus M, Noakes TDD, Collins M. The COL12A1 and COL14A1 genes and achilles tendon injuries. Int J Sports Med. 2008;29(3):257–63.

615. CHEN H, SHI L, ZHANG J, ZHAO D, SHI L, LU Q, et al. Association of COL1A1, COL5A1 and COL12A1 genes with anterior cruciate ligament injury 前交叉韧带损伤与胶原蛋白基因COL1A1、COL5A1及COL12A1的相关性. Chinese Journal of Tissue Engineering Research. 2017;21(12):1849–54.

616. Wei W, He HL, Chen CY, Zhao Y, Jiang HL, Liu WT, et al. Whole exome sequencing implicates PTCH1 and COL17A1 genes in ossification of the posterior longitudinal ligament of the cervical spine in Chinese patients. Genetics and Molecular Research. 2014;13(1):1794–804.

617. Saunders CJ, Van Der Merwe L, Posthumus M, Cook J, Handley CJ, Collins M, et al. Investigation of variants within the COL27A1 and TNC genes and Achilles tendinopathy in two populations. Journal of Orthopaedic Research. 2013;31(4):632–7.

618. Stępień-Słodkowska M, Ficek K, Eider J, Leońska-Duniec A, MacIejewska-Karłowska A, Sawczuk M, et al. The +1245G/T polymorpsisms in the collagen type I alpha 1 (Col1a1) gene in polish skiers with anterior cruciate ligament injury. Biol Sport. 2013;30(1):57–60.

619. Stępień-Słodkowska M, Ficek K, Kaczmarczyk M, Maciejewska-Karłowska A, Sawczuk M, Leońska-Duniec A, et al. Stępień-Słodkowska et al. - The Variants Within the COL5A1 Gene are Associated with Reduced Risk of Anterior Cruciate Ligament Injury in.pdf. J Hum Kinet. 2015 Mar;45(1):103–11.

620. Hay M. An investigation of DNA sequence variants in genes that regulate collagen fibrillogenesis and predisposition to musculoskeletal soft tissue injuries. Thesis. 2013;(April).

SUPPLEMENTARY MATERIAL

A – Ethics Approval and Recruitment Forms

A1.1 UCT HREC Ethics Approval: Shoulder Injuries in Swimming

A1.2 Recruitment Information Sheet: Shoulder Injuries in Swimming

A1.3 Informed Consent Form: Shoulder Injuries in Swimming

A1.4 Participant Questionnaire: Shoulder Injuries in Swimming

A1.5 UCT HREC Ethics Approval: Anterior Cruciate Ligament Injuries and a Mixed Ancestry Cohort

A1.6 Recruitment Information and Informed Consent Form Sheet: Anterior Cruciate Ligament Injuries and a Mixed Ancestry Cohort

A1.7 Participant Questionnaire: Anterior Cruciate Ligament Injuries and a Mixed Ancestry Cohort

A1.8 UCT HREC Ethics Approval: Anterior Cruciate Ligament Injuries in a Swedish Cohort

A1.9 UU RERB Ethics Approval: Punkt 1 Fortsatt handläggning

A1.10 Samtycke kontroll: Informed consent controls

A1.11 Samtycke korsband: Informed consent patients

B – Supplementary Results

A1.1 UCT HREC Ethics Approval: Shoulder Injuries in Swimming

UNIVERSITY OF CAPE TOWN

FACULTY OF HEALTH SCIENCES
Human Research Ethics Committee

HUMAN RESEARCH ETHICS COMMITTEE
06 NOV 2021
UNIVERSITY OF CAPE TOWN

FHS016: Annual Progress Report / Renewal

HREC office use only (FWA00001637; IRB00001938) This serves as notification of annual approval, including any documentation described below.			
☑ Approved	Annual progress report	Approved until/next renewal date	30.10.22
☐ Not approved	See attached comments		
Signature Chairperson of the HREC/ Designee		Date Signed	9/11/2021

Note: Please email this form and supporting documents (if applicable) in a combined pdf-file to hrec-enquiries@uct.ac.za.
Please clarify your plan for research-related activities during COVID-19 lockdown.
Please use the latest form found on our website:
http://www.health.uct.ac.za/fhs/research/humanethics/forms

Comments to PI from the HREC

Principal Investigator to complete the following:

1. Protocol information

Date (when submitting this form)	25 Oct 2021		
HREC REF Number	421/2013	Current Ethics Approval was granted until	30/10/2021
Protocol title	Factors associated with shoulder injuries in swimmers		
Protocol number (if applicable)	N/A		
Are there any sub-studies linked to this study?		☐ Yes	✓ No
If yes, could you please provide the HREC Reference number for all sub-studies? Note: A separate FHS016 must be submitted for each sub-study.			
Principal Investigator	Prof. Malcolm Collins		

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

A1.2 Recruitment Information Sheet: Shoulder Injuries in Swimming

Division of Exercise Science and 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

RISK FACTORS ASSOCIATED WITH SHOULDER INJURIES IN SWIMMERS

Swimmers can perform up to ONE MILLION arm rotations per arm per year whilst they are training. This puts a large amount of stress on the shoulder and can lead to an injury. However, some swimmers have never been injured and others are plagued by chronic pain and injury. WHY IS THIS? That is exactly the question we are asking at the Division of Exercise Science and Sports Medicine. Some researchers have suggested that there is a genetic component to injuries that can put a person at risk of sustaining an injury. We would like to investigate how your genes can put you at risk for injury and whether or not we will be able to one day predict injuries in swimmers. **BUT, WE CANNOT DO THIS WITHOUT YOUR HELP.**

Requirements for Participants

- Any swimmers over the age of 18 **(MALE and FEMALE)**
- We are looking for both **INJURED** and **UNINJURED** swimmers
- A single testing session of 15-30 minutes

Where does testing take place?

Testing will take place either at the Sports Science Institute (Boundary Road, Newlands) or at your club or at a competition (I am flexible and will fit in with your time). The testing will take approximately 15-30 minutes of your time. However, if necessity dictates, I am will to travel to you.

What to expect during the study?

- You will also be required to complete a single questionnaire documenting personal particulars, sporting participation, personal and family medical history and swimming history questionnaire.
- You will be asked to donate 5 ml (1 teaspoon) of a blood sample or depending on the circumstances, a saliva sample for DNA analysis.

What are we specifically investigating?

The genes we will be investigating are those that are involved in tendon growth, inflammation and healing in response to exercise. The specific genes we are interested in are variations within the type V collagen (molecules that are involved in your connective tissues) gene (*COL5A1*), the Tenacin-C gene (*TNC*), the matrix metalloproteinase-3 gene (*MMP3*), the growth and differentiation factor-5 gene (*GDF5*), *CASP8* gene and *IL-1β, IL-1RN* and *IL-6* genes as well as additional genes, which may become relevant during the course of the investigation. These particular genes have been shown to either increase or decrease your risk of injury.

What are the benefits of taking part?

- There will be no direct benefit to you for participating in this study however, if a genetic predisposition for skeletal muscle injuries can be established, then future generations will be able to establish their risk for this condition.
- This may allow better prevention and treatment options for your fellow swimmers in the future.
- You will not be reimbursed or compensated if you participate.
- You will not receive your personal genetic results.
- You will however, receive overall results of the study and its findings.

Important study Ethics

- All the information retrieved from this study will be treated with the strictest confidentiality and will be used only for scientific research purposes.
- Your name and personal particulars will not be released under any circumstances and all data will be analysed anonymously.

- Your DNA sample will be destroyed on completion of the study on the genetic risk factors associated with shoulder tendinopathy; however, due to the nature of genetic research, new candidate genes may be identified during the course of the study which may be investigated in addition to the genes we are already investigating.
- Your DNA sample will be stored in coded form so that you remain anonymous and will not have your personal particulars linked to your DNA sample.
- You are also free to request that your DNA sample be destroyed before the completion of the study.
- Any research data provided by you during the project may be included in a thesis, presented at conferences and published in journals on the condition that neither your name nor any other identifying information is used.

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:

Mr Lee Hill, B.Sc (Med)(Hons) *Exercise Science*
0832668650
uctswimming@gmail.com

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

Prof Malcolm Collins, Phd
Malcolm.collins@uct.ac.za

This study has obtained ethical approval from the UCT Faculty of Health Sciences Research Ethic Committee (HREC REF: 421/2013). If you have any complaints or queries that the investigator has not been able to answer to your satisfaction, you may contact **Prof Marc Blockman** from the FHS REC on telephone number **021 406 6452**.

A1.3 Informed Consent Form: Shoulder Injuries in Swimming

Division of Exercise Science and 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

RISK FACTORS ASSOCIATED WITH SHOULDER INJURIES IN SWIMMERS

INFORMED CONSENT

I, the undersigned, have been fully informed about the Department of Exercise Science and Sports Medicine within the Faculty of Health Science at the University of Cape Town's study to identify genetic risk factors associated with rotator cuff tendinopathy.

I have agreed to give five millilitres of venous blood or a saliva mouthwash/swab sample, which will be used for analysis of genetic material (DNA). I have also agreed to complete personal particulars, sporting participation, personal and family medical history, and muscle cramping questionnaires and understand that all the information that is collected during the study will be treated with the strictest confidentiality and will only be used for scientific research purposes. I also understand that my name and personal particulars will be not released under any circumstances and that all data will be analysed anonymously.

I give permission that the study investigators may access my medical records (doctor/physiotherapist/biokineticist) in order to confirm my diagnosis (either previous or current).

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

I understand the potential risks (having the risks fully explained) associated with blood collection technique from the forearm veins and give permission for a sample to be taken.

I understand that the DNA will be genotyped and analysed for variations within the type V collagen gene (*COL5A1*), the Tenacin-C gene (*TNC*), the matrix metalloproteinase-3 gene (*MMP3*), the growth and differentiation factor-5 gene (*GDF5*), *CASP8* gene and *IL-1β, IL-1RN* and *IL-6* genes as well as additional genes, which may become relevant during the course of the investigation.

I understand that there is no direct benefit to myself. I understand that I will receive the overall results of the study. I have read (or where appropriate, have had read to me) and understand the information about this study, and any questions I have asked have been answered to my satisfaction.

I agree to participate in the study and that my DNA (in coded form) will be stored until the completion of the study. 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 Investigators: **Dr Mike Posthumus** on telephone number **021 650 4572** or e-mail michael.posthumus@uct.ac.za and **Mr Lee Hill** on cell phone number 0832668650 or e-mail uctswimming@gmail.com.

This study has obtained ethical approval from the UCT Faculty of Health Sciences Research Ethic Committee (FHS REC). If you have any complaints or queries that the investigator has not been able to answer to your satisfaction, you may contact **Prof Marc Blockman** from the FHS REC on telephone number **021 406 6452**.

Initial:_____

I have read and received an oral explanation of the informed consent form and would like to take part in the current study.

FULL NAME OF SUBJECT: _____

SUBJECT'S SIGNATURE: _____

DATE: _____

INVESTIGATOR : _____

INVESTIGATOR'S SIGNATURE: _____

A1.4 Participant Questionnaire: Shoulder Injuries in Swimming

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
Lee Hill Email: uctswimming@gmail.com
Cell: 0832668650

SWIMMING INJURY QUESTIONNAIRE

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

1

B. OCCUPATIONAL DETAILS		
What is your current occupation?		
What was your occupation prior to your shoulder injury (if applicable)?		
Prior to injury, did your occupation involve lifting using your arms?	Right arm ☐ Both arms ☐ Left arm ☐ None ☐	
If yes:	Approximate load: Light (0-5kg) ☐ Medium (5-10kg) ☐ Heavy (over 10kg) ☐	
	Hours per week of lifting:	Number of years:
Prior to injury, did your occupation involve lifting *above the shoulder?*	Right arm ☐ Both arms ☐ Left arm ☐ None ☐	
If yes:	Approximate load: Light (0-5kg) ☐ Medium (5-10kg) ☐ Heavy (over 10kg) ☐	
	Hours per week of lifting:	Number of years:

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

C. **SPORTING DETAILS** **Please record your sporting activities in order of importance**			
Type of sport(s) you have participated in (please name)	**Main sport 1**	**Other sport 2**	**Other sport 3**
Current or past participation	Current ☐ Past ☐	Current ☐ Past ☐	Current ☐ Past ☐
Year started participation			
Number of years involved in the sport			
Hours of training per week in the last 12 months			
Hours of training per week in the last 13-24 months			

SPORTING DETAILS Continued **Please record your sporting activities in order of importance**			
Type of sport(s) you have participated in (please name)	**Other sport 4**	**Other sport 5**	**Other sport 6**
Current or past participation	Current ☐ Past ☐	Current ☐ Past ☐	Current ☐ Past ☐
Year started participation			
Number of years involved in the sport			
Hours of training per week in the last 12 months			
Hours of training per week in the last 13-24 months			

D. SWIMMING HISTORY

How many hours a week do you **currently** train for swimming?		
How long have you been swimming (years)?		
What is your highest level that you have competed at?	☐ International (Represented Country) ☐ Provincial (Attended national competition) ☐ Club ☐ School ☐ Social ☐ Never competed ☐ Other……	
Are you currently competing in competitions?	☐ Yes ☐ No	
If yes, please select at what level you are currently competing.	☐ International (Represented Country) ☐ Provincial (Attended nationals) ☐ Club ☐ School ☐ Social ☐ Never competed ☐ Other……	
How often do you compete?	Currently	Previously
	☐ Once a year ☐ Once every 6 months ☐ Once a month ☐ Once a week ☐ None ☐ Other (specify)……	☐ Once a year ☐ Once every 6 months ☐ Once a month ☐ Once a week ☐ None ☐ Other (specify)……

What are **your hours of training per week in the periods below:** (Please specify)	**Hours per Week**	**Did you take a break during the period? If Yes, How long was the break(months or weeks)?**
Hours of training per week in the last 0-3 months		☐ Yes _____
Hours of training per week in the last 4-12 months		☐ Yes _____
Hours of training per week in the last 13-24 months		☐ Yes _____
Hours of training per week in the last 3-5 years		☐ Yes _____
Hours of training per week in the last 6-10 years		☐ Yes _____
Hours of training per week in the last 11-20 years		☐ Yes _____
Hours of training per week in the last >20 years		☐ Yes _____

Are you registered with a club or swim team? If yes please state.	☐ Yes _____	☐ No
Are you instructed by a coach or qualified trainer?	☐ Yes	☐ No
Do you swim in a pool or open water (Sea or fresh water)?	Pool ☐ Open water (Fresh) ☐ Open Water (Sea) ☐	
How many times a week do you train with paddles, kickboard or poolbouy?	Paddles	_____per week
	Kickboard	_____per week
	Poolbouy	_____per week
What is the length of the pool you <u>train</u> in?	☐ 25m ☐ 50m ☐ Other (please specify)_____	
How many **hours per week** do you train during summer and winter?	Summer (in-season) _____hrs/week	Winter (off-season) _____hrs/week
What is the length of the pool you <u>Race</u> in?	☐ 25m ☐ 50m ☐ Other (please specify)_____	

Do you **CURRENTLY** participate in any of the sports listed? If yes, please state the duration of the activity.	Waterpolo	☐ Yes	☐ No	____Hrs/week
	Life-Saving	☐ Yes	☐ No	____Hrs/week
	Synchronized Swimming	☐ Yes	☐ No	____Hrs/week
	Surfing	☐ Yes	☐ No	____Hrs/week
	Canoeing	☐ Yes	☐ No	____Hrs/week
	Kayaking	☐ Yes	☐ No	____Hrs/week
	Surf Ski	☐ Yes	☐ No	____Hrs/week
	Paddle Ski	☐ Yes	☐ No	____Hrs/week
	Rowing	☐ Yes	☐ No	____Hrs/week
	Other (specify)	☐ Yes	☐ No	____Hrs/week

Have you **PREVIOUSLY** participated in any other sport listed? If **Yes**, please the state the duration of the activity.	Waterpolo	☐ Yes	☐ No	____Hrs/week
	Life-Saving	☐ Yes	☐ No	____Hrs/week
	Synchronized Swimming	☐ Yes	☐ No	____Hrs/week
	Surfing	☐ Yes	☐ No	____Hrs/week
	Canoeing	☐ Yes	☐ No	____Hrs/week
	Kayaking	☐ Yes	☐ No	____Hrs/week
	Surf Ski	☐ Yes	☐ No	____Hrs/week
	Paddle Ski	☐ Yes	☐ No	____Hrs/week
	Rowing	☐ Yes	☐ No	____Hrs/week
	Other (specify)	☐ Yes	☐ No	____Hrs/week

	Training			**Competition**		
Have you had an injury that forced you to take time off training and or competition?	☐ Yes	☐ No	☐ N/A	☐ Yes	☐ No	☐ N/A

Have you had any injury or a shoulder injury in periods described below?	**Any injury**			**Shoulder Injury**		
If yes, did it occur in the last 0-3 months	☐ Yes	☐ No	☐ N/A	☐ Yes	☐ No	☐ N/A
If yes, did it occur in the last 4-12 months	☐ Yes	☐ No	☐ N/A	☐ Yes	☐ No	☐ N/A
If yes, did it occur in the last 13-24 months	☐ Yes	☐ No	☐ N/A	☐ Yes	☐ No	☐ N/A
If yes, did it occur in the last 3-5 years	☐ Yes	☐ No	☐ N/A	☐ Yes	☐ No	☐ N/A
If yes, did it occur in the last 5-10 years	☐ Yes	☐ No	☐ N/A	☐ Yes	☐ No	☐ N/A
If yes, did it occur in the last 10-20 years	☐ Yes	☐ No	☐ N/A	☐ Yes	☐ No	☐ N/A
If yes, did it occur in the last >20 years	☐ Yes	☐ No	☐ N/A	☐ Yes	☐ No	☐ N/A

Please select the strokes you **train and compete in** (You may select more than one)	**Train**	**Compete**
Butterfly	☐ Yes ☐ No	☐ Yes ☐ No
Backstroke	☐ Yes ☐ No	☐ Yes ☐ No
Breaststroke	☐ Yes ☐ No	☐ Yes ☐ No
Freestyle	☐ Yes ☐ No	☐ Yes ☐ No
Individual Medley	☐ Yes ☐ No	☐ Yes ☐ No

Please mark **(with an X)** the distances and strokes you **previously** competed in.	Butterfly	Backstroke	Breaststroke	Freestyle	Individual Medley
50m					
100m					
200m					
400m					
800m					
1500m					
1600m/ Mile					
3000m					
5000m					
Other (specify)					

Please mark **(with an X)** the distances and strokes you **currently** compete in.	Butterfly	Backstroke	Breaststroke	Freestyle	Individual Medley
50m					
100m					
200m					
400m					
800m					
1500m					
1600m/ Mile					
3000m					
5000m					
Other (specify)					

E. SHOULDER INJURY DETAILS

Have you ever experienced shoulder pain?	☐ Yes	☐ No	☐ Don't know
Do you currently experience any shoulder pain?	☐ Yes	☐ No	☐ Don't know
Do you currently expereince shoulder pain whilst swimming?	☐ Yes	☐ No	☐ Don't know
Have you previously experienced shoulder pain whilst swimming?	☐ Yes	☐ No	☐ Don't know
Was your shoulder pain diagnosed by a medical professional/physician?	☐ Yes	☐ No	☐ Don't know
Who diagnosed your shoulder pain?	☐ Orthopaedic Surgeon ☐ General Practitioner ☐ Sports Physician ☐ Physiotherapist ☐ Biokineticist ☐ Other (please specify):_____		
Who diagnosed/treated your pain; name and practice (If Known)?			
Did you have any treatment for your shoulder pain?	☐ Yes	☐ No	☐ Don't know
If yes, what treatment was done?	☐ Surgery ☐ Rehabilitation ☐ Rest ☐ Medication ☐ Other: _____		

Have you had a **specific rotator cuff injury**?	☐ Yes	☐ No	☐ Don't know
Have you ever been diagnosed with rotator cuff **tendinopathy or tendinitis**?	☐ Yes	☐ No	☐ Don't know

If Yes (you have had rotator cuff tendinopathy / tendinitis), please complete the questions below for EACH injury.

	Injury 1	Injury 2	Injury 3
Date of Injury	DD / MM / YY	DD / MM / YY	DD / MM / YY
Please specify which **rotator cuff tendon**	☐ Supraspinatus ☐ Infraspinatus ☐ Subscapularis ☐ Teres Major ☐ Teres Minor ☐ Biceps Tendon ☐ Unknown	☐ Supraspinatus ☐ Infraspinatus ☐ Subscapularis ☐ Teres Major ☐ Teres Minor ☐ Biceps Tendon ☐ Unknown	☐ Supraspinatus ☐ Infraspinatus ☐ Subscapularis ☐ Teres Major ☐ Teres Minor ☐ Biceps Tendon ☐ Unknown
Which side was injured	Left ☐ Right ☐	Left ☐ Right ☐	Left ☐ Right ☐
Which practitioner diagnosed this injury	☐ Orthopaedic Surgeon ☐ General Practitioner ☐ Physiotherapist ☐ Sports Physician ☐ Biokineticist ☐ Other (please specify)_____	☐ Orthopaedic Surgeon ☐ General Practitioner ☐ Physiotherapist ☐ Sports Physician ☐ Biokineticist ☐ Other (please specify)_____	☐ Orthopaedic Surgeon ☐ General Practitioner ☐ Physiotherapist ☐ Sports Physician ☐ Biokineticist ☐ Other (please specify)_____
Who was the practitioner; name and practice (Name if known)			
What Investigation was done to confirm the diagnosis	☐ MRI/CT Scan ☐ Surgery ☐ Ultrasound ☐ Clinical Observation ☐ Other (please specify)_____	☐ MRI/CT Scan ☐ Surgery ☐ Ultrasound ☐ Clinical Observation ☐ Other (please specify)_____	☐ MRI/CT Scan ☐ Surgery ☐ Ultrasound ☐ Clinical Observation ☐ Other (please specify)_____
What was the final treatment for injury you have indicated above?	☐ Surgery ☐ Rehabilitation ☐ Rest ☐ Medication ☐ Other: _____	☐ Surgery ☐ Rehabilitation ☐ Rest ☐ Medication ☐ Other: _____	☐ Surgery ☐ Rehabilitation ☐ Rest ☐ Medication ☐ Other: _____

If you ticked Surgery, Please describe **(if known)** what surgery was performed.			
May we contact your practitioner(s) to confirm the diagnosis of injury?	☐ Yes	☐ No	

What was your training load **prior** to your **INJURY 1**?	0-3 months before injury	_____ hrs/week
	3-6 months before injury	_____ hrs/week
What was your training load **prior** to your **INJURY 2**?	0-3 months before injury	_____ hrs/week
	3-6 months before injury	_____ hrs/week
What was your training load **prior** to your **INJURY 3**?	0-3 months before injury	_____ hrs/week
	3-6 months before injury	_____ hrs/week

Have you had **any other shoulder-related** injuries?	☐ Yes	☐ No	☐ Don't know
If **yes**, please tick from the list below.			
Impingement Syndrome(s)			
☐ Infraspinatus Impingement ☐ Supraspinatus Impingement ☐ Rotator cuff disease	☐ Internal impingement ☐ Anterio-inferior acromion impingement	☐ Posterior Superior impingement ☐ Other (Please Specify): _____	
Musculo-Skeletal			
☐ Long head of bicep Rupture ☐ Full thickness rotator cuff tears ☐ Bursal side rotator cuff tears ☐ Exertional compartment syndrome ☐ Adhesive capsulitis ☐ Bursal hypertrophy ☐ Superior labral injury ☐ Glenohumeral Dislocation ☐ Bone Fracture ☐ Bone tumor	☐ Rib stress fractures ☐ Pectoral (major and minor) muscle tear ☐ Bicep muscle tear ☐ Tricep muscle tear ☐ Deltoid muscle tear ☐ Lattimus Dorsi muscle tear ☐ Rhomboid muscle tear ☐ Serratus muscle tear ☐ Levator Scapulae muscle tear ☐ Trapezius muscle tear	☐ Glenoid Labral tears ☐ Glenohumoral joint arthritis ☐ AC joint arthrosis ☐ AC joint injury ☐ Bankart Lesion ☐ Hill-Sachs Lesion ☐ Coracoacromial ligament ossification ☐ Os Acromionale ☐ Other (Please Specify): _____	

Nerve and Vascular		
☐ Brachial plexus neuropathy ☐ Proximal vascular Obstruction ☐ Neuropraxia ☐ Neuritis	☐ Thoracic outlet syndrome ☐ Cervical radiculopathy ☐ Nerve entrapment (Suprascapular or Long thoracic) ☐ Axillary vein thrombosis ☐ Other (Please Specify): _____	
Instability and Sublaxation		
☐ Glenohumeral Instability ☐ Levator scapulae syndrome ☐ Scapular dyskinesis	☐ Glenohumeral internal rotation deficit (GIRD) ☐ AC joint instability ☐ Involuntary dislocation	☐ Sternoclavicular joint sublaxation ☐ Other (Please Specify): _____
Inflammatory and Other		
☐ Synovitis ☐ Calcific tendinopathy	☐ Bursitis ☐ Other (Please Specify): _____	
What was the final treatment for injury you have indicated above?	☐ Surgery ☐ Rehabilitation ☐ Rest	☐ Medication ☐ Other (please specify) _____
If you ticked Surgery, Please describe **(if known)** what surgery was performed.		

F. **FLEXIBILITY AND STRENGTH TRAINING HISTORY**	
Do you perform **flexibility training** (regular stretching exercises)?	Yes ☐ No ☐
Do you perform **strength training** (regular weight lifting or other exercise to increase strength)?	Yes ☐ No ☐
Is your strength and flexibility training administrated or designed by a qualified trainer, biokineticist or physiotherapist?	Yes ☐ No ☐
On average, how many <u>days a week</u> do you perform a **stretching** session?	_____days/week
On average, how many days a week do you perform **strength training** session?	_____days/week
Please tick <u>which muscle groups</u> do you include in your **stretching session**?	☐ Shoulders ☐ Biceps ☐ Triceps ☐ Chest ☐ Deltoids ☐ Lower back ☐ Upper back (Trapezius, Latimuss Dorsi, Rhomboids) ☐ Lower Limbs ☐ Other: _____
Please tick <u>which muscle groups</u> do you include in your **strength training** sessions?	☐ Shoulders ☐ Biceps ☐ Triceps ☐ Chest ☐ Deltoids ☐ Lower back ☐ Upper back (Trapezius, Latimuss Dorsi, Rhomboids) ☐ Lower Limbs ☐ Other: _____
Please tick when you stretch? (Before, during and/or after exercising. You can tick more than one box)	☐ Before Exercise ☐ During Exercise ☐ After Exercise ☐ Separate Sessions ☐ Administrated stretch classes (i.e. Yoga or Pilates) ☐ Other:_____
When you stretch an individual muscle group, on average, <u>**how long do you hold the stretch**</u> for?	_____seconds

When you stretch an individual muscle group, on average, **how <u>many times do you stretch the muscle for</u>**?	☐ Once ☐ Twice ☐ 3 times ☐ 4 times ☐ 5 times ☐ 6 or more times
Do you perform any **<u>swimming specific</u>** prehabilitation or rehabilitation exercise?	☐ Theraband (Elastic) ☐ Weights ☐ Stretch ☐ Other _____
How many **times per week** do you perform the swimming specific prehabilitation or rehabilitation?	_____per/week
Was the swimming specific prehabilitation or rehabilitation designed by a qualified Physiotherapist, Biokineticist, or Physician?	Yes ☐ No ☐

G. **HISTORY OF OTHER LIGAMENT AND TENDON INJURIES IN THE PAST (NON-SHOULDER)**				
Have you ever injured a ligament in the past?	Yes ☐ No ☐			
If yes, please specify which ligaments? (You may tick more than one block, please select either L (left) or R (right))		L R		L R
	Knee (ACL)	☐ ☐	Wrist ligaments	☐ ☐
	Knee (MCL)	☐ ☐	Finger ligaments	☐ ☐
	Ankle lateral ligaments	☐ ☐	Knee (PCL)	☐ ☐
	Elbow ligaments	☐ ☐	Knee (LCL)	☐ ☐
	Ankle medial ligament	☐ ☐	Other ligaments	☐ ☐
To your knowledge, have any other members of your family suffered from any ligament injury?	Yes ☐ No ☐	If Yes, please specify the family member ☐ Mother ☐ Father ☐ Sibling ☐ Son / daughter ☐ Other family member_____ and condition: Please choose ligament injury from the list above _____		
Have you ever injured a tendon in the past?	Yes ☐ No ☐			
If yes, please specify which tendon?	Foot and ankle:	L R		

(You may tick more than one block, please select either L (left) or R (right))		Achilles tendon	☐ ☐
		Tibialis posterior	☐ ☐
		Plantar fascia	☐ ☐
	Knee:	Patellar tendon	☐ ☐
	Elbow and wrist:	Wrist extensor tendons	☐ ☐
	Other:_____		
To your knowledge, have any other members of your family suffered from any tendon pathology?	Yes ☐ No ☐	If Yes, please specify the family member ☐ Mother ☐ Father ☐ Sibling ☐ Son / daughter ☐ Other family member_____ and condition: Please choose ligament injury from the list above _____	

H. MEDICAL HISTORY

Do you currently suffer from any of these medical conditions:

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

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

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

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

What **SURGICAL OPERATIONS** have you had? (please list and give dates)	Operation	Date

If female:	
At what age did you start menstruating? (years)	
Are you currently using any type of contraception?	☐ Yes ☐ No
If Yes, what type of contraception are you using?	☐ Pill ☐ Injection ☐ IUD
Are you currently?	☐ Pre-menopausal (±12 cycles per year at intervals of 23– 33 days & bleeding lasts 3-7 days) ☐ Menopausal (cycles are irregular and less frequent) ☐ Post-menopausal (no longer menstruating)

J. Family History		
Do any other members of your family suffer from elevated blood cholesterol?	Yes ☐ No ☐	If Yes, which relative? ☐ Mother ☐ Father ☐ Sibling ☐ Son / daughter ☐ Other relative:_____
Is there any history of arthritis in your family?	Yes ☐ No ☐	If Yes, which relative? ☐ Mother ☐ Father ☐ Sibling ☐ Son / daughter ☐ Other relative:_____ & What type of arthritis? Rheumatoid ☐ Osteoarthritis ☐ Other ☐

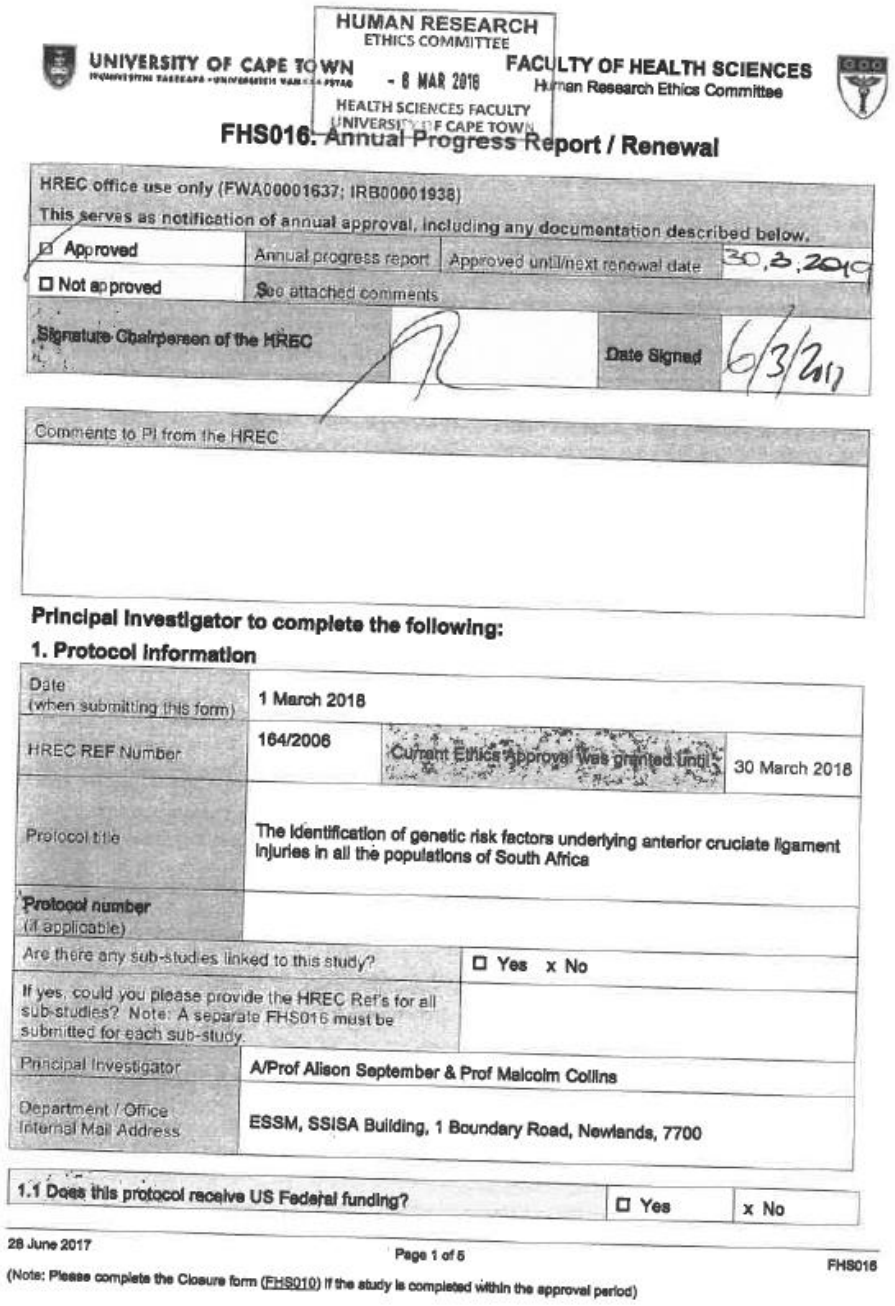
I. Drug and Allergy History		If yes, how long ago (or how many times, where applicable) did you use the medication?
Have you ever used oral corticosteroids (cortisone tablets)?	Yes ☐ No ☐	☐ 3 months ☐ 6 months ☐ 12 months ☐ 24 or more months
Was it prior to your injury?	Yes ☐ No ☐	☐ 3 months ☐ 6 months ☐ 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
Was it prior to your injury?	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
Was it prior to your injury?	Yes ☐ No ☐	☐ 3 months ☐ 6 months ☐ 12 months ☐ 24 or more months
Have you ever used anabolic steroids?	Yes ☐ No ☐	☐ 3 months ☐ 6 months ☐ 12 months ☐ 24 or more months
Was it prior to your injury?	Yes ☐ No ☐	☐ 3 months ☐ 6 months ☐ 12 months ☐ 24 or more months
Have you ever used fluoroquinoline antibiotics (Anti-malarials, Anti-STI)?	Yes ☐ No ☐	☐ 3 months ☐ 6 months ☐ 12 months ☐ 24 or more months
Was it prior to your injury?	Yes ☐ No ☐	☐ 3 months ☐ 6 months ☐ 12 months ☐ 24 or more months

If **yes (I have taken corticosteroids),** please select from the list below (if known):		
☐ ADCO-CIPRIN ☐ AVELON ☐ BACTIDRON ☐ CIFLOC ☐ CIFRAN ☐ CIPLA-CIPROFLOXACIN ☐ CIPLOXX ☐ CIPRO-HEXAL ☐Other_____ _	☐ CIPROBAY ☐ CIPROGEN ☐ CPL ALLIANCE CIPROFLOXACIN ☐ DYNAFLOC ☐ FLOXIN ☐ MAXAQUIN ☐ NOROXIN ☐ ORPIC	☐ SANDOZ CIPROFLOXACIN ☐ TAFLOC ☐ TARIVID ☐ TAVANIC ☐ TEQUIN ☐ UNIQUIN ☐ UTN-400 ☐ ZANOCIN

What medication, if any, are you currently using? (please list)	
What allergies do you have? (please list)	

Thank you for your time. ☺

A1.5 UCT HREC Ethics Approval: Anterior Cruciate Ligament Injuries and a Mixed Ancestry Cohort



HUMAN RESEARCH ETHICS COMMITTEE
- 8 MAR 2018
HEALTH SCIENCES FACULTY
UNIVERSITY OF CAPE TOWN

UNIVERSITY OF CAPE TOWN
FACULTY OF HEALTH SCIENCES
Human Research Ethics Committee

FHS016: Annual Progress Report / Renewal

HREC office use only (FWA00001637; IRB00001938) This serves as notification of annual approval, including any documentation described below.			
☑ Approved	Annual progress report	Approved until/next renewal date	30.3.2019
☐ Not approved	See attached comments		
Signature Chairperson of the HREC		Date Signed	6/3/2017

Comments to PI from the HREC

Principal Investigator to complete the following:

1. Protocol Information

Date (when submitting this form)	1 March 2018		
HREC REF Number	164/2006	Current Ethics Approval was granted until	30 March 2018
Protocol title	The Identification of genetic risk factors underlying anterior cruciate ligament injuries in all the populations of South Africa		
Protocol number (if applicable)			
Are there any sub-studies linked to this study?	☐ Yes x No		
If yes, could you please provide the HREC Ref's for all sub-studies? Note: A separate FHS016 must be submitted for each sub-study.			
Principal Investigator	A/Prof Alison September & Prof Malcolm Collins		
Department / Office Internal Mail Address	ESSM, SSISA Building, 1 Boundary Road, Newlands, 7700		

1.1 Does this protocol receive US Federal funding?	☐ Yes	x No

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

A1.6 Recruitment Information Sheet: Anterior Cruciate Ligament Injuries and a Mixed Ancestry Cohort

PARTICIPANT QUESTIONAIRES AND INFORMED CONSENT

Department of Human Biology

Faculty of Health Sciences
University of Cape Town
Private Bag X3
Observatory 7935
South Africa
Phone: +27-21-406 6235
Fax: +27-21-448 7226

THE GENETIC BASIS OF EXERCISE-INDUCED ANTERIOR CRUCIATE LIGAMENT INJURY

Although there is a high incidence of anterior cruciate ligament (ACL) injuries as a result of participation in exercise and sporting activities, the mechanism and cause of this injury is still poorly understood. There is scientific evidence to suggest that there is a genetic component to exercise-induced ACL injuries. In an attempt to determine whether this is true, we at the Division of Exercise Science and Sports Medicine within the Department of Human Biology of the University of Cape Town, are interested in studying whether certain genes are associated with ACL injuries.

In order to accomplish this, we need to analyse the genetic material, called the DNA. The DNA in our body makes all our proteins and you have protein-coding DNA (exons) and non-protein coding DNA (introns). This project will screen both the exon and intron regions of the genes, by using a whole genome sequencing study (WGS) design which basically gives us the sequence of your entire DNA.

Aims of the research:

1. Analyse the entire DNA sequence to determine if there are any variations which could contribute to ACL injuries.
2. To assist physicians in developing multifactorial models to identify individuals at an increased risk of ACL injury.

Your possible involvement:

Should you agree to participate, you would be asked to do the following:

- Complete a questionnaire relating to your personal particulars, sporting details, and medical history. This information will be anonymous and only a coding system will be used to identify you.
- Donate a 5ml (1 teaspoon) blood sample from a vein in your arm. This will be used for the extraction and analysis of genetic material (DNA).

The DNA will only be used for scientific research purposes. We will perform analyses on your entire genome by sequencing your DNA.

To ensure complete confidentiality of your specific genetic information, the following procedures will be adopted: 1) all the blood samples will be labelled on collection using a numerical coding system that is linked to your details on a master list that will be placed in a sealed envelope, 2) this sealed master list will then be kept in a secure facility and in a separate location, 3) only the principle investigator and co-investigators will have access to this master list, 4) the master list will only be opened if a sample needs to be destroyed, should a participant request this. All data will be analysed anonymously and DNA samples will be stored or destroyed as indicated by you on completion of the study. Please be aware that you are free to request that your DNA sample be destroyed before the completion of the study, and furthermore you may withdraw, with no penalties or consequences, at any point during the study.

Furthermore, due to ethical reasons, we cannot disclose any individual genetic results from this study, however a summary of the overall findings will be communicated to all the participants on completion of the research.

Potential Risks of the study:

- The completion of a questionnaire or a physical examination is not associated with any risk. Questionnaire and other clinical data (paper and electronic) will be kept confidential and secure, and will not be made available to any party other than the research team without the consent of the individual participants.
- The potential risks to participants of blood collection are minimal and are related to 1) blood sample collection technique, and 2) the volume of blood. The potential risks associated with blood collection technique from the ante-cubital veins are: infection, delayed healing, haematoma, physical pain, mental discomfort and injury to a nerve or a vessel. These risks are small and will be minimized by the use of trained phlebotomists, use of sterile techniques and the use of disposable, single-use materials.
- Your personal and genetic information will be kept secure, anonymous and will only be used for research

Potential Benefits of the study:

The research questions that will be addressed by this study have been identified to have a direct impact on improving the understanding of ACL injuries. This could impact on the diagnosis, medical treatment and/or physical activity modifications. The anticipated benefits of this study are that the results will further our understanding of the possible cause/s of ACL injuries and the genetic alterations linked to the injury.

We look forward to working with you, and are most appreciative of your contribution to this novel medical research. If you have any questions or concerns regarding any aspect of the study, please do not hesitate to contact us.

A1.7 Participant Questionnaire: Anterior Cruciate Ligament Injuries and a Mixed Ancestry Cohort

Department of Human Biology

Faculty of Health Sciences
University of Cape Town
Private Bag X3
Observatory 7935
South Africa
Phone: +27-21-406 6235
Fax: +27-21-448 7226

GENETIC BASIS OF ANTERIOR CRUCIATE LIGAMENT INJURY

Date of birth		Age	
Height (cm)		Gender	Male ☐ Female ☐
Weight (kg)	Pre-Injury:	Current:	
Ethnic group (Only required and used for the genetic aspect of the research)	Black/African ☐ White ☐ Indian ☐ Mixed Ancestry (Coloured) ☐ Asian ☐ Other ☐		
Ancestry: Tribal or national background (eg Xhosa, Dutch, Zulu, German, Italian)	Father Unknown ☐ Mother Unknown ☐		
Country of Birth			
Dominant Hand	Left ☐ Right ☐ Ambi ☐	Dominant Leg	Left ☐ Right ☐ Ambi ☐
Smoker	Yes (Current) ☐	Yes (Ex smoker) ☐	No, never ☐
	If yes, Number of years _____	If stopped, when _____	
	If yes, number per day _____		
B. OCCUPATIONAL DETAILS			
What is your current occupation?			

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

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

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

Type of sport(s) you have participated in (please name)	Other sport 4	Other sport 5	Other sport 6
Current or past participation	Current □ Past □	Current □ Past □	Current □ Past □
Year started participation			
Number of years involved in the sport			

Position played prior to injury (if appropriate)			
Playing level prior to injury (if appropriate)			
Number of years played prior to the injury.			

D. ANTERIOR CRUCIATE LIGAMENT INJURY DETAILS	
Date of ACL injury?	
Which side was injured?	□ Left □ Right □ Both
To what extent was your ligament ruptured?	□ Complete □ Partial □ None □ Unknown
Investigation done to confirm the diagnosis	□ MRI □ Surgery
How bad is your pain today? (mark line: e.g.l----------I----------l)	I----------I No pain Pain as bad as it can be
How was the ACL ruptured?	□ Direct impact (directly to the injured knee) □ Twisting and bending with indirect contact (i.e. contact elsewhere on the body) □ Twisting and bending without contact (no external contact) □ Skiing □ Other (please specify)……
Please describe the exactly how the injury occurred (If uncertain please state that you do not know)	…… …… …… ……

What was the initial treatment? (You may tick more than one block.)	□ Ice application □ Compression □ Immobilisation □ Medication □ Other…….
What was the final treatment?	□ Surgery □ Rehabilitation □ Other……..
What are your current symptoms? (You may tick more than one block.)	□ Pain □ Swelling □ Instability □ Weakness □ Other……
What is your current sports participation?	□ None □ Limited to non-weight bearing exercise □ Limited, not to same level as pre-injury □ Full participation
If you are able to recall, what were the weather and pitch conditions like at the time of injury?	□ Wet and soft ground □ Dry, but soft ground □ Dry and firm ground □ Wet, but firm ground □ Other……..

Associated injuries?	□ Meniscal tear □ MCL tear □ Other ligament tear □ Bone bruising □ Other……..

E. HISTORY OF OTHER LIGAMENT AND TENDON INJURIES IN THE PAST						
Have you ever injured a ligament in the past?	Yes □			No □		
If yes, please specify which ligaments? (You may tick more than one block, please select either L (left) or R (right))		L	R		L	R
	Knee (ACL)			Wrist ligaments		
	Knee (MCL)			Finger ligaments		
	Ankle lateral ligaments			Knee (PCL)		
	Spinal ligaments			Knee (LCL)		
	Shoulder ligaments			Ankle medial ligaments		
	Elbow ligaments			Other ligaments		

To your knowledge, have any other members of your family suffered from any ligament injury?	Yes □ No □	If Yes, please specify the family member □ Mother □ Father □ Sibling □ Son / daughter □ Other family member……..… and condition: Please choose ligament injury from the list above ……

Have you ever injured a tendon in the past?	Yes □		No □	
If yes, please specify which tendon? (You may tick more than one block, please select either L (left) or R (right)			L	R
	Foot and ankle	Achilles tendon		
		Tibialis posterior		
		Plantar fascia		
	Knee	Patellar tendon		
	Elbow and wrist	Wrist extensor tendons		
	Shoulder	Subscapularis		
		Supraspinatus		
		Infraspinatus		
		Teres minor		
	Other			

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

F. MEDICAL HISTORY
Do you currently suffer from any of these medical conditions:

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

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

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

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

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

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

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

	If yes, please select from the list below:	
□ ADCO-CIPRIN □ AVELON □ BACTIDRON □ CIFLOC □ CIFRAN □ CIPLA-CIPROFLOXACIN □ CIPLOXX □ CIPRO-HEXAL □ Other	□ CIPROBAY □ CIPROGEN □ CPL ALLIANCE CIPROFLOXACIN □ DYNAFLOC □ FLOXIN □ MAXAQUIN □ NOROXIN □ ORPIC	□ SANDOZ CIPROFLOXACIN □ TAFLOC □ TARIVID □ TAVANIC □ TEQUIN □ UNIQUIN □ UTN-400 □ ZANOCIN
What medication, if any, are you currently using? (please list)		
What allergies do you have? (please list)		

A1.8 UCT HREC Ethics Approval: Anterior Cruciate Ligament Injuries in a Swedish Cohort

HUMAN RESEARCH ETHICS COMMITTEE
22 MAY 2019
HEALTH SCIENCES FACULTY UNIVERSITY OF CAPE TOWN

UNIVERSITY OF CAPE TOWN
FACULTY OF HEALTH SCIENCES
Human Research Ethics Committee

FHS016: Annual Progress Report / Renewal

HREC office use only (FWA00001637; IRB00001938) This serves as notification of annual approval, including any documentation described below.			
☑ Approved	Annual progress report	Approved until/next renewal date	30.04.2020
☐ Not approved	See attached comments		
Signature Chairperson of the HREC		Date Signed	22/5/2019

Comments to PI from the HREC
Thank you for the deviation document

Principal Investigator to complete the following:

1. Protocol Information

Date (when submitting this form)	21/05/2019		
HREC REF Number	269/2014	Current Ethics Approval was granted until	30/04/2019
Protocol title	The identification of genetic susceptibility – a 20 year follow up of patientswith ACL injury		
Protocol number (if applicable)	N/A		
Are there any sub-studies linked to this study?	✓ Yes ☐ No		
If yes, could you please provide the HREC Ref's for all sub-studies? **Note:** A separate FHS016 must be submitted for each sub-study.	619/2015 – Risk factors associated with injury in swimmers		
Principal Investigator	Prof M Collins		
Department / Office Internal Mail Address	3rd Floor, Sport Science Institute of South Africa, Division of Exercise Science and Sports Medicine, 1 Boundary Rd, Newlands, 7725		

1.1 Does this protocol receive US Federal funding?	☐ Yes	✓ No

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

A1.9 UU RERB Ethics Approval: Punkt 1 Fortsatt handläggning

Regionala
etikprövningsnämnden
i Umeå
Avdelningen för medicinsk
forskning

PROTOKOLLSUTDRAG
sammanträdesdag
2011-09-06

Ärende, beslut eller annan åtgärd

Dnr
2011-200-31M

Punkt 1 Fortsatt handläggning
Föredragande: Erik Lundgren
Forskningshuvudman
Umeå universitet
Företrädare
Gunnevi Sundelin, prefekt, Inst. för samhällsmedicin och rehabilitering, Umeå universitet
Forskare
Charlotte Häger, Inst. för samhällsmedicin och rehabilitering, Enh. för sjukgymnastik, Umeå universitet
Projekttitel
Genetiska varianters betydelse för främre korsbandsskada.

Ärendet behandlades vid sammanträdet 2010-06-07, Punkt 15, varvid nämnden beslöt att begära in komplettering (1 original och 17 kopior) från ansvarig forskare och att, när komplettering inkommit, ta upp ärendet till förnyad bedömning vid ett kommande sammanträde.

Komplettering har nu inkommit.

Beslut
Projektet godkänns.

Föreslagen provhantering är förenlig med Biobankslagen.

Vid protokollet
Gunnel Eriksson

Justeras
Eric Lowén
Erik Lundgren

Bestyrkes
Gunnel Eriksson

Kopia till
Behörig företrädare

A1.10 Samtycke korsband: Informed consent form for patients

Samtyckeshandling personer med korsbandsskada
Bilaga 4 (i två kopior, varav en till forskningsperson)

Umeå 2011-05-13

Samtyckeshandling för deltagande i studien " Genetiska variationers betydelse för främre korsbandsskada"

För Ditt deltagande i ovan studie behövs enligt god forskningsetisk praxis ett dokument som visar att Du har fått information om den studie Du deltar i samt att Du beretts möjlighet att ställa frågor och få eventuella oklarheter förklarade för Dig. Du har också fått information om till vem Du ska vända Dig om frågor uppstår framgent. Den skriftliga informationen om studien återfinns i dokumentet "Information till patienter" daterat Umeå 2011-04-04.

Du är medveten om att de enkätuppgifter Du lämnat kommer att registreras i en forskningsdatabas och behandlas så att inga utomstående kan ta del av dessa. Du har fått information om att Du som studiedeltagare enligt personuppgiftslagen (1998:204) har rätt att få s.k. registerutdrag och även rätt att få eventuella felaktigheter korrigerade.

Du är införstådd med att de forskningsprover Du lämnat sparas i en biobank på klinisk genetik, Norrlands universitetssjukhus, Umeå. Du vet också att du när som helst, och utan att lämna någon särskild förklaring, kan avbryta Ditt deltagande i studien och begära att kvarvarande prover ska förstöras eller anonymiseras.

Med Din namnteckning nedan samtycker Du till att delta i studien " Genetiska variationers betydelse för främre korsbandsskada "

Med Din namnteckning bekräftar Du även att Du har läst och förstått den information som skriftligen delgetts Dig och frivilligt samtycker till att delta.

Provgivarens namnunderskrift	Personnummer
Namnförtydligande	Ort och datum (dag/mån/år)

OBS Ta med denna ifyllda samtyckeshandling till sjukhuset eller vårdcentralen där du tar ditt blodprov om du väljer att deltaga!

A1.11 Samtycke kontroll: Informed consent form for controls

Samtyckeshandling
Friska kontrollpersoner
Bilaga 4 (i två kopior varav 1 till forskningspersonen)

Umeå 2011-05-13

Samtyckeshandling för deltagande i studien " Genetiska varianters betydelse för främre korsbandsskada"

För Ditt deltagande i ovan studie behövs enligt god forskningsetisk praxis ett dokument som visar att Du har fått information om den studie Du deltar i samt att Du beretts möjlighet att ställa frågor och få eventuella oklarheter förklarade för Dig. Du har också fått information om till vem Du ska vända Dig om frågor uppstår framgent. Den skriftliga informationen om studien återfinns i dokumentet "Information till patient och anhöriga" daterat Umeå 2011-03-22.

Du är medveten om att de enkätuppgifter Du lämnat kommer att registreras i en forskningsdatabas och behandlas så att inga utomstående kan ta del av dessa. Du har fått information om att Du som studiedeltagare enligt personuppgiftslagen (1998:204) har rätt att få s.k. registerutdrag och även rätt att få eventuella felaktigheter korrigerade.

Du är införstådd med att de forskningsprover Du lämnat sparas i en biobank på klinisk genetik, Norrlands universitetssjukhus, Umeå. Du vet också att du när som helst, och utan att lämna någon särskild förklaring, kan avbryta Ditt deltagande i studien och begära att kvarvarande prover ska förstöras eller anonymiseras.

Med Din namnteckning nedan samtycker Du till att delta i studien " Genetiska varianters betydelse för främre korsbandsskada "

Med Din namnteckning bekräftar Du även att Du har läst och förstått den information som skriftligen och muntligen delgetts Dig och frivilligt samtycker till att delta.

Provgivarens namnunderskrift	Personnummer
Namnförtydligande	Ort och datum (dag/mån/år)

OBS Ta med denna ifyllda samtyckeshandling till sjukhuset eller vårdcentralen där du tar ditt blodprov om du väljer att deltaga!

APPENDIX B1: CHAPTER 2 SUMMARY OF INVESTIGATED GENOTYPES AND SPECIFIC MSK PHENOTYPES

Table B1.1: Polymorphisms previously associated with soft tissue injuries; including anterior cruciate ligament ruptures (ACL), Achilles pathology including Achilles (AT) and Achilles Rupture (AR), carpal tunnel syndrome (CTS), rotator cuff pathology (RCP) including rotator cuff tendinopathy (RCT), Rotator cuff Disease (RCD) and Shoulder Dislocations (SD), Lumbar disk pathology (LDP) including lumbar disk herniation (LDH) and Lumbar Disk Degeneration (LDD), Longitudinal ligament of spine (LLS), tennis elbow (TE) and Ankle injuries (AI). Table reports pathology as specifically defined by the authors of each study.

Gene	Protein	Polymorphism	Location	ACL	AP	CTS	RTP	LDP	LLS	TE	AI
COL1A1	α1(I) chain Type I collagen	rs1800012 (G/T)	17q21.33	TT Genotype ↓ Risk In SA ACL(565) GG Genotype ↓ risk in Polish ACL(566) Not significant(567–571)	Not significantly associated with AT(570,572) TT Genotype ↓ Risk of AR(572)	Not significant(573)	TT ↓ risk of SD(574) Not associated with RCD(575)·(576)	TT ↑ risk of LDD(577)-(578)·(579) GT ↑ risk of LDH(576)		Not significant(580)	
		rs1107946 (–1997 G/T)		Not significant(566,568,570)	Not significant(570)						
		rs2075555 (T/G)						Not significant(581)·(582)			
		rs1007086 (C/T)						Not significant(581)			
		rs909102 (C/T)						Not significant(581)			
COL1A2	α2(I) chain Type I collagen	rs3763468 (C/T)	7q21.3					Not significant(581)			
		rs388625 (A/G)						Not significant(581)			
		rs412777 (A/C/T)						Not significant(581)			
		rs400218 (A/G)						Not significant(581)			
		rs1034620 (C/T)						Not significant(581)			
COL2A1	α2(I) chain Type I collagen	rs917055 (A/G)	12q13.11					Not significant(581)·(582)			
		r1859443 (C/T)						Not significant(581)			
		rs1635529 (A/C)						Not significant(581)			
		rs2276454 (G/T)						Not significant(581) TT ↑ risk of IVDD(583) TC and CC ↓ risk of IVDD in Chinese(583)			
		rs2276458 (A/C/G/T)						Not significant(581)			
		rs6823 (C/G)						Not significant(581)			
		rs1793953 (G/A)						AA ↑ risk of IVDD in Chinese(583)			
		rs1793937 (C/G)						Not significant(583)			
COL3A1	α1(III) chain Type III collagen	rs1800255 (G/A)	2q31	Not significant(568,584) AA ↑ risk In Polish ACL(585)				Not significant(581) Not significant(579)			
		rs2056156 (C/T)						Not significant(581)·(586)·(579)·(582)			
		rs2203601 (A/T)						Not significant(581)·(579)			
COL5A1	α1(V) chain Type V collagen	rs12722 (C/T)	9q34.2–q34.3	CC ↓ risk in females(587) TT ↓ risk(588) Not significant(568,571)	CC ↓ risk in SA AT(589) Not significant in AUS AT[5] Not significant in British AP(590)	CC ↓ risk(591)	Not associated with RCD(575,576) Not associated RC tears (592) CC ↑ risk RC tendinopathy(593)	Not significant in LDH(594)		TT ↑ risk(595) Not significant(596)	TT ↓ risk(588)
		rs10628678(AGGG/-)		Not significant in SA and AUS(597)	AGGG/AGGG ↑ risk in AUS and SA AT(598) Not significant in British AP(590)	Not significant(591)					

		rs13946 (C/T)		Not significant(568,571,587)	Not significant in AUS and SA AT[5]	TT ↓ risk(591)				TT ↓ risk(595)	
		rs16399 (ATCT/-)			-/- ↑ risk of AT(598)						
		rs1134170 (A/T)			TT ↑ risk of AT(598)						
		rs3196378 (C/A AciI RFLP)			AC ↑ risk in AUS AT(589)						
		rs3922912 (G/A)		Not significant(599)							
		rs4841926 (C/T)		Not significant(599)							
		rs3124299 (C/T)		Not significant(599)							
COL6A1	α1(VI) chain Type VI collagen	rs35796750 (T/C)	21q22.3	Not significant(584)	Not significant(600)				Not significant(601,602)		
COL6A2	αII(VI) chain Type VI collagen	rs914246 (A/C)	21q22.3								
COL9A1	α1(IX) chain Type IX collagen	rs696990 (C/T)	6q13.					Not significant(581)·(582)			
		rs564031 (A/G)						Not significant(581)			
		rs592121 (A/G)						Not significant(581)			
		rs2076816 (C/T)						Not significant(581)			
		rs1200564 (C/T)						Not significant(581)			
		rs997953 (A/G)						Not significant(581)			
COL9A2	α2(IX) chain Type IX collagen	rs209923 (–2066 C/A)	1p33–p32					CC ↑ risk(603)			
		rs7533552 (+977 A/G)						AA ↑ risk(581) Not significant in Norwegian LLD·(582)			
		rs2076696 (G/C)						Not significant(604)			
		rs12077871 (Trp2; +976 C/T; R326W)						CC ↑ risk(581)·(605)·(606)·(607)			
		rs364281 (A/C)						Not significant(581)			
		rs449541 (C/T)						Not significant(581)			
COL9A3	α3(IX) chain Type IX collagen	rs61734651 (C/T; Trp3; R103W)	20q13.3					TT ↑ risk(581)·(594)·(608) Not significant in LLD·(582)·(609)			
		rs3891033 (A/G)						Not significant(581)			
		rs1046789 (A/G)						Not significant(581)			
COL10A1	α1(X) chain Type X collagen	rs549332 (C/T)	6q22.1					Not significant(581)			
		rs1064583 (C/T)						Not significant(581)			
		rs3812111 (A/T)						Not significant(581)			
		rs568725 (C/T)						Not significant(581)			
COL11A1	α1(XI) chain Type XI collagen	rs3753841 (T/C)	1p21		Not significant(610)	TT ↑ risk(573)		Not significant(581)		CT ↑ risk(596)	
		rs1676486 (C/T)			Not significant(610)	Not significant(573)		Not significant(581)·(582) CC ↓ risk in LDH(611)		Not significant(596)	
		rs1463035 (C/T)						Not significant(581)·(582)			
		rs1337185 (A/C/G)						Not significant(581)			
		rs1415359 (A/G)						Not significant(581)			
COL11A2	α2(XI) chain Type XI collagen	rs1799907 (T/A)	6p21.3		Not significant(610)	Not significant(591)		Not significant(612)·(609)		Not significant(596)	
		rs2072915 (A/T)						Not significant(581)·(582)			
		rs9277933 (A/T)						Not significant(581)			
		rs2076311 (G/T)						Not significant(581)			
		rs2855432 (C/T)						Not significant(581)			
		rs2257126 (C/T)						Not significant(581)			
		rs734181 (C/G)						Not significant(581)			

COL12A1	α1(XII) chain Type XII collagen	rs970547 (A/G)	6q12–q13	AA↑risk in Females(613) Not significant(568,571,613)	Not significant(614)	Not significant(573)					
				AA ↑ in Chinese ACL injuries(615)							
		rs240736 (T/C)		Not significant in SA ACL(565,571) Not significant in Indian ACL(567)	Not significant in SA AT(614)						
COL14A1	α1(XIV) chain Type XIV collagen	rs4870723 (A/C)	8q23		Not significant(614)						
		rs1563392 (A/T)			Not significant(614)						
COL17A1	α1(XVII) chain Type XVII collagen	Rs805708 (C/T)							Not significant t(616)		
		rs805701 (C/T)							Not significant t(616)		
		rs805698 (G/A)							GG ↑ in LLS t(616)		
		rs805722 (A/G)							Not significant t(616)		
		rs4918079 (C/T)							CC ↑ in LSS t(616)		
		rs2274100 (A/G)							Not significant(616)		
COL18A1	α1(XVIII) chain Type XVIII collagen		21q22.3								A/AG and AG/AG ↑ risk(601)
COL27A1	α1(XXVII) collagen chain	rs946053 (G/T)	9q32		Not significant(617)						
		rs1249744 (A/G)			Not significant(617)						
		rs753085 (C/T)			Not significant(617)						
		rs946053 (G/T)			Not significant(617)						
Gene-Gene Interactions											
COL1A1		rs1107946 (G/T) rs1800012 (G/T)		G-T haplotype ↓ risk(618) Not significant(570)	G-T haplotype ↓ risk(570)						
COL3A1 *COL6A1*		rs1800255 (G/A) rs35796750 (T/C)			A-C haplotype ↓ risk(600)						
COL3A1 *COL5A1* *COL6A1*		rs1800255 (G/A) rs12722 (T/C) rs35796750 (T/C)			A-C-C haplotype ↓ risk(600)						
COL5A1		rs12722 (C/T) rs13946 (C/T)		T-C haplotype ↓ risk(619)		T-W-C haplotype ↓ risk(591)					
		rs12722 (C/T) rs10628678 (AGGG/-)	T-(-) haplotype haplotype ↑ risk(597)								
COL5A1 *COL11A1* *COL11A2*		rs10628678 (AGGG/AGGG) rs3753841 (T/C) rs1676486 (C/T) rs1799907 (T/A)			AGGG-T-C-T haplotype ↑ risk(610)	(-)-T-C haplotype ↑ risk(573) AGGG-T-C haplotype ↑ risk(573)					

						AGGG-C-C haplotype ↓ risk(573)					
COL5A1 *COL12A1*		rs12722 (C/T) rs970547 (A/G)		T-A ↑ risk in females(584)							
COL6A1 *COL12A1*											
COL9A2		rs209923 (–2066 C/A) rs7533552 (+977 A/G) rs2076696 (+1147G/C)						A-G-C haplotype ↑ risk			
COL11A1 *COL11A2*		rs3753841 (T/C) rs1676486 (C/T) rs1799907 (T/A)			T-C-T ↑ risk(610)						
		rs3753841 (T/C) rs1676486 (C/T)		C-T ↑ risk(620)		T-C ↑ risk(573)					

APPENDIX B2: CHAPTER 3 SUMMARY OF INCLUDED STUDIES

Table B2.1. Summary of included studies, level of evidence and level of certainty.

Author	Study Design	Population	Sample Size (pain/injury)	Level of Evidence	Level of Certainty
CLINICAL JOINT LAXITY AND INSTABILITY					
Increased Risk					
McMaster et al (1998)	Cross Sectional	Elite USA Swimmers	40 Elite Competitive Swimmers (27/40).	III	Moderate
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	
Bansal et al (2012)	Cross Sectional	Elite Male Swimmers	161 National and International Level Swimmers (State level swimmers 7/61, national 18/91 and international 3/9 swimmers had impingement syndrome).	III	
No Association					
Borsa et al (2005)	Case Control	Collegiate Swimmers and matched Controls	42 Swimmers and 44 Controls (27/42).	III	
Santos et al 2007	Case Control	Elite Competitive Swimmers	8 Swimmers with no history of pain or injury and 8 swimmers with diagnosed impingement syndrome.	III	
Walker et al (2009)	Prospective	Male and Female Competitive Swimmers	74 Swimmers (28/74 reported significant interfering pain and 17/74 reported significant interfering injury in the shoulder).	I	
INTERNAL/EXTERNAL ROTATION					
Increased Risk					
Walker et al (2009)	Prospective	Male and Female Competitive Swimmers	74 Swimmers (28/74 reported significant interfering pain and 17/74 reported significant interfering injury in the shoulder).	I	Low
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	
Bansal et al (2007)	Cross Sectional	Elite Male Swimmers	161 National and International Level Swimmers (State level swimmers 7/61, national 18/91 and international 3/9 swimmers had impingement syndrome).	III	
No Association					

Beach et al (1995)	Cross Sectional	Division I Collegiate Swimmers	32 Swimmers (69% reported previous pain and 31% reported current interfering pain).	III	
Rupp et al (1995)	Case Control	Competitive Swimmers and Matched Controls	22 Cases and 22 Control (14/22 reported previous shoulder pain and 5/22 reported current shoulder pain).	III	
Harrington et al (2014)	Cross Sectional	Female Collegiate Swimmers	37 Competitive Swimmers (12/37 dominant arm, 14/37 non-dominant arm).	III	
McLaine et al (2017)	Cross Sectional	Competitive swimmers	68 Competitive swimmers (27/68 reported history of shoulder pain).	III	
YEARS OF SWIMMING EXPERIENCE					
Increased Risk					
Sein et al (2010)	Cross Sectional	Competitive Swimmers	80 Competitive Swimmers (73/80 report significant shoulder pain and 84% had a positive impingement sign).	III	
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	
Rodeo et al (2016)	Cross sectional	Elite swimmers	42 Elite swimmers (22 male and 20 female). A history of shoulder pain was reported by 29 of 42 (69%) athletes.	III	
Wymore et al (2015)	Cross sectional	NCAA Collegiate swimmers	99 collegiate swimmers assessed using the KJOC Score	III	
No Association					
Hidalgo-Lozano et al (2012)	Case Control	Elite Competitive Swimmers	17 healthy swimmers without shoulder pain and 17 swimmers with shoulder pain.	III	
Chase et al (2013)	Prospective	Male and Female Collegiate Swimmers	34 Swimmers (38% of swimmers were diagnosed with a shoulder injury during the season).	II	Low
Hidalgo-Lozano et al (2013)	Case Control	Elite Competitive Swimmers	17 Swimmers with shoulder pain, 18 swimmers without shoulder pain and 15 healthy athlete-control.	III	
Harrington et al (2014)	Cross Sectional	Female Collegiate Swimmers	37 Competitive Swimmers (12/37 dominant arm, 14/37 non-dominant arm).	III	
Feijen et al (2021)	Prospective	Club to elite Swimmers	201 Club to elite level swimmers without shoulder pain.	I	
McKenna et al (2018)	Case controls	Adult swimmers	26 symptomatic and 26 asymptomatic swimmers.	III	
De Almeida et al (2015)	Cross sectional	National level swimmers	257 national level swimmers of which 21% reported current presence of shoulder pain.	III	
Tessaro et al (2017)	Retrospective	Club swimmers	197 club swimmers. 51% of swimmers reported previous history of shoulder pain	III	

Suzuki et al (2020)	Case control	Masters swimmers	20 masters swimmers with shoulder pain, 20 asymptomatic and 20 age-matched controls	III	
TRAINING LOAD, VOLUME AND INTENSITY					
Increased Risk					
Sein et al (2008)	Cross Sectional	Competitive Swimmers	80 Competitive Swimmers (73/80 report significant shoulder pain and 84% had a positive impingement sign).	III	
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	
Kruger et al (2012)	Cross Sectional	Competitive Masters Swimmers	282 Masters Swimmers (62.4% of swimmers reported pain over the 3 year period. 28.7% reported recurring pain and 37.2% reported that pain was severe enough to cease training).	III	
Ristolainen et al (2014)	Retrospective	Finnish Elite level Swimmers, Cross-Country Skiers and Long Distance Runners	268 Swimmers (52.1% of swimmers reported a significant shoulder injury).	II	
No Association					Low
Su et al (2004)	Case Control Pre and Post testing	Competitive Age Group Swimmers	20 Swimmers with diagnosed shoulder Impingement and 20 swimmers without.	I	
Walker et al (2009)	Prospective	Male and Female Competitive Swimmers	74 Swimmers (28/74 reported significant interfering pain and 17/74 reported significant interfering injury in the shoulder).	I	
Hidalgo-Lozano et al (2012)	Case Control	Elite Competitive Swimmers	17 healthy swimmers without shoulder pain and 17 swimmers with shoulder pain.	III	
Hidalgo-Lozano et al (2013)	Case Control	Elite Competitive Swimmers	17 Swimmers with shoulder pain, 18 swimmers without shoulder pain and 15 healthy athlete-control.	III	
Chase et al (2013)	Prospective	Male and Female Competitive Swimmers	34 Swimmers (38% of swimmers were diagnosed with a shoulder injury during the season).	I	
Harrington et al (2014)	Cross Sectional	Female Collegiate Swimmers	37 Competitive Swimmers (12/37 dominant arm, 14/37 non-dominant arm).	III	
McKenna et al (2018)	Case controls	Adult swimmers	26 symptomatic and 26 asymptomatic swimmers.	III	
De Almeida et al (2015)	Cross sectional	National level swimmers	257 national level swimmers of which 21% reported current presence of shoulder pain.	III	
McLaine et al (2018)	Cross sectional	Age group club swimmers	85 club swimmers. A history of pain was reported in 27 swimmer, 58 reported no history of pain.	III	

McLaine et al (2019)	Cross sectional	Age group club swimmers	85 club swimmers. A history of pain was reported in 27 swimmers, 58 reported no history of pain.	III	
Tessaro et al (2017)	Retrospective	Club swimmers	197 club swimmers. 51% of swimmers reported previous history of shoulder pain	III	
Couanis et al (2015)	Prospective	Open water swimmers	22 open water swimmers measured for sub-acromial bursa thickness.	II	
Suzuki et al (2020)	Case control	Masters swimmers	20 masters swimmers with shoulder pain, 20 asymptomatic and 20 age-matched controls	III	
Feijen et al (2021)	Prospective	Club to Elite level swimmers	201 club to elite level swimmers (42/201 reported shoulder pain over the period)	I	
HISTORY OF PREVIOUS PAIN AND INJURY					
Increased Risk					
Walker et al (2009)	Prospective	Male and Female Competitive Swimmers	74 Swimmers (28/74 reported significant interfering pain and 17/74 reported significant interfering injury in the shoulder).	I	
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	
Bansal et al (2007)	Cross Sectional	Elite Male Swimmers	161 National and International Level Swimmers (State level swimmers 7/61, national 18/91 and international 3/9 swimmers had impingement syndrome).	III	Moderate
No Association					
Kruger et al (2012)	Cross Sectional	Male and Female Competitive Masters Swimmers	282 Masters Swimmers (62.4% of swimmers reported pain over the 3 year period. 28.7% reported recurring pain and 37.2% reported that pain was severe enough to cease training).	III	
McKenna et al (2018)	Case controls	Adult swimmers	26 symptomatic and 26 asymptomatic swimmers.	III	
De Almeida et al (2015)	Cross sectional	National level swimmers	257 national level swimmers of which 21% reported current presence of shoulder pain.	III	
McLaine et al (2019)	Cross sectional	Age group club swimmers	85 club swimmers. A history of pain was reported in 27 swimmers, 58 reported no history of pain.	III	
COMPETITIVE LEVEL					
Increased Risk					
Zemek & Magee (1996)	Case Control	Elite Swimmers and Recreational Controls	30 Elite Swimmers and 30 Recreational Swimmers Elite swimmers demonstrated 67% Shoulder Overuse Dysfunction (pain and injury) than 13% of controls).	III	Moderate

Sein et al (2008)	Cross Sectional	Male and Female Competitive Swimmers	80 Competitive Swimmers (73/80 report significant shoulder pain and 84% had a positive impingement sign).	III	
Kruger et al (2012)	Cross Sectional	Male and Female Competitive Masters Swimmers	282 Masters Swimmers (62.4% of swimmers reported pain over the 3 year period. 28.7% reported recurring pain and 37.2% reported that pain was severe enough to cease training).	III	
No Association					
Su et al (2004)	Case Control Pre and Post testing	Competitive Age Group Swimmers	20 Swimmers with diagnosed shoulder Impingement and 20 swimmers without.	I	
Feijen et al (2021)	Prospective	Club to elite Swimmers	201 Club to elite level swimmers without shoulder pain.	I	
SEX					
Increased Risk					
Bak & Faunø (1997)	Retrospective	Competitive Swimmers with history of therapy resistant shoulder pain.	23 swimmers (number of injured)	II	Low
Sallis et al (2001)	Retrospective	Male and Female Seven Competitive Collegiate Sports	3767 Collegiate Athletes (n for swimmers not available, however it is reported that female swimmers had 21.05 shoulder injuries compared to male swimmers with 6.55 injuries).	II	
Tessaro et al (2017)	Retrospective	Club swimmers	197 club swimmers. 51% of swimmers reported previous history of shoulder pain	II	
No Association					
Griepp (1985)	Prospective	Male and Female Competitive Swimmers (12-23 years)	168 Swimmers (54% of all females reported shoulder pain over the season, 71.9% of males reported shoulder pain over the season).	II	
Su et al (2004)	Case Control Pre and Post testing	Competitive Age Group Swimmers	20 Swimmers with diagnosed shoulder Impingement and 20 swimmers without.	I	
Kruger et al (2012)	Cross Sectional	Male and Female Competitive Masters Swimmers	282 Masters Swimmers (62.4% of swimmers reported pain over the 3 year period. 28.7% reported recurring pain and 37.2% reported that pain was severe enough to cease training).	III	
Chase et al (2013)	Prospective	Male and Female Competitive Swimmers	34 Swimmers (38% of swimmers were diagnosed with a shoulder injury during the season).	I	
Boettcher et al (2020)	Cross sectional	Elite swimmers	68 elite swimmers (40 male and 28 female). Twenty-four percent reported shoulder pain during competition and training at the time	III	

			of testing with a further 53% reporting a previous history of shoulder pain.		
McKenna et al (2018)	Case controls	Adult swimmers	26 symptomatic and 26 asymptomatic swimmers.	III	
Feijen et al (2021)	Prospective	Club to elite Swimmers	201 Club to elite level swimmers without shoulder pain.	I	
De Almeida et al (2015)	Cross sectional	National level swimmers	257 national level swimmers of which 21% reported current presence of shoulder pain.	III	
McLaine et al (2018)	Cross sectional	Age group club swimmers	85 club swimmers. A history of pain was reported in 27 swimmers, 58 reported no history of pain.	III	
McLaine et al (2019)	Cross sectional	Age group club swimmers	85 club swimmers. A history of pain was reported in 27 swimmers, 58 reported no history of pain.	III	
Wymore et al (2015)	Cross sectional	NCAA Collegiate swimmers	99 collegiate swimmers assessed using the KJOC Score	III	
Suzuki et al (2020)	Case control	Masters swimmers	20 masters swimmers with shoulder pain, 20 asymptomatic and 20 age-matched controls	III	
AGE					
Increased Risk					
Rodeo et al (2016)	Cross sectional	Elite swimmers	42 Elite swimmers (22 male and 20 female). A history of shoulder pain was reported by 29 of 42 (69%) athletes.	III	Low
McLaine et al (2018)	Cross sectional	Age group club swimmers	85 club swimmers. A history of pain was reported in 27 swimmers, 58 reported no history of pain.	III	
McLaine et al (2019)	Cross sectional	Age group club swimmers	85 club swimmers. A history of pain was reported in 27 swimmers, 58 reported no history of pain.	III	
No Association					
Griepp (1985)	Prospective	Competitive Swimmers (12-23 years)	168 Swimmers (54% of all females reported shoulder pain over the season, 71.9% of males reported shoulder pain over the season).	II	
Su et al (2004)	Case Control Pre and Post testing	Competitive Age Group Swimmers	20 Swimmers with diagnosed shoulder Impingement and 20 swimmers without.	I	
Walker et al (2009)	Prospective	Male and Female Competitive Swimmers	74 Swimmers (28/74 reported significant interfering pain and 17/74 reported significant interfering injury in the shoulder).	I	
Kruger et al (2012)	Cross Sectional	Male and Female Competitive Masters Swimmers	282 Masters Swimmers (62.4% of swimmers reported pain over the 3 year period. 28.7% reported recurring pain and 37.2% reported that pain was severe enough to cease training).	III	
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	

Chase et al (2013)	Prospective	Male and Female Competitive Swimmers	34 Swimmers (38% of swimmers were diagnosed with a shoulder injury during the season).	I	
Harrington et al (2014)	Cross Sectional	Female Collegiate Swimmers	37 Competitive Swimmers (12/37 dominant arm, 14/37 non-dominant arm).	III	
Sabzehparvar et al (2019)	Case control	Elite swimmers	12 Elite swimmers with shoulder pain and 12 pain free controls	III	
McKenna et al (2018)	Case controls	Adult swimmers	26 symptomatic and 26 asymptomatic swimmers.	III	
De Almeida et al (2015)	Cross sectional	National level swimmers	257 national level swimmers of which 21% reported current presence of shoulder pain.	III	
Wymore et al (2015)	Cross sectional	NCAA Collegiate swimmers	99 collegiate swimmers assessed using the KJOC Score	III	
Tessaro et al (2017)	Retrospective	Club swimmers	197 club swimmers. 51% of swimmers reported previous history of shoulder pain	II	
Suzuki et al (2020)	Case control	Masters swimmers	20 masters swimmers with shoulder pain, 20 asymptomatic and 20 age-matched controls	III	
SWIMMING TRAINING EQUIPMENT					
Increased Risk					
McMaster et al (1989)	Cross Sectional	Competitive Age Group Swimmers	473 Swimmers (27/40 reported shoulder pain).	III	
No Association					
Stocker et al (1995)	Cross Sectional	Competitive Collegiate and Masters Swimmers	532 Collegiate and 395 Masters level Swimmers (47% of college and 48% of masters swimmers reported pain lasting longer than 3 weeks in the shoulder).	III	Low
Puckree & Thomas (2006)	Cross Sectional	Competitive Swimmers	96 Swimmers (64% reported injury to the shoulder).	III	
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	
CROSS TRAINING AND STRETCHING					
Increased Risk					
Griepp (1985)	Prospective	Competitive Swimmers (12-23 years)	168 Swimmers (54% of all females reported shoulder pain over the season, 71.9% of males reported shoulder pain over the season).	II	
McMaster et al (1989)	Cross Sectional	Competitive Age Group Swimmers	473 Swimmers (27/40 reported shoulder pain).	III	Low
No Association					

Walker et al (2009)	Prospective	Male and Female Competitive Swimmers	74 Swimmers (28/74 reported significant interfering pain and 17/74 reported significant interfering injury in the shoulder).	I	
Kruger et al (2012)	Cross Sectional	Male and Female Competitive Masters Swimmers	282 Masters Swimmers (62.4% of swimmers reported pain over the 3 year period. 28.7% reported recurring pain and 37.2% reported that pain was severe enough to cease training).	III	
De Almeida et al (2015)	Cross sectional	National level swimmers	257 national level swimmers of which 21% reported current presence of shoulder pain.	III	
Tessaro et al (2017)	Retrospective	Club swimmers	197 club swimmers. 51% of swimmers reported previous history of shoulder pain	III	
STROKE DISTANCE AND STROKE SPECIALTY					
Increased Risk					
McMaster et al (1989)	Cross Sectional	Competitive Age Group Swimmers	473 Swimmers (27/40 reported shoulder pain).	III	Low
No Association					
Griepp (1985)	Prospective	Competitive Swimmers (12-23 years)	168 Swimmers (54% of all females reported shoulder pain over the season, 71.9% of males reported shoulder pain over the season).	II	
Walker et al (2009)	Prospective	Male and Female Competitive Swimmers	74 Swimmers (28/74 reported significant interfering pain and 17/74 reported significant interfering injury in the shoulder).	I	
Wolf et al (2009)	Retrospective	Collegiate Swimmers	94 Swimmers (number of injuries).	II	
Sein et al (2008)	Cross Sectional	Competitive Swimmers	80 Competitive Swimmers (73/80 report significant shoulder pain and 84% had a positive impingement sign).	III	
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	
Kruger et al (2012)	Cross Sectional	Male and Female Competitive Masters Swimmers	282 Masters Swimmers (62.4% of swimmers reported pain over the 3 year period. 28.7% reported recurring pain and 37.2% reported that pain was severe enough to cease training).	III	
Wymore et al (2012)	Cross Sectional	Elite Collegiate Swimmers	187 Swimmers (43% of butterfly swimmers reported no shoulder pain and 73% reported pain 2 days per week or less. 43% of breaststroke specialists, 42% of freestyle specialists and 41% of backstroke specialists reported no days of shoulder pain).	III	
McKenna et al (2018)	Case controls	Adult swimmers	26 symptomatic and 26 asymptomatic swimmers.	III	
De Almeida et al (2015)	Cross sectional	National level swimmers	257 national level swimmers of which 21% reported current presence of shoulder pain.	III	
BREATHING SIDE					

Increased Risk					
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	
No Association					
Stocker et al (1995)	Cross Sectional	Competitive Collegiate and Masters Swimmers	532 Collegiate and 395 Masters level Swimmers (47% of college and 48% of masters swimmers reported pain lasting longer than 3 weeks in the shoulder).	III	Low
Dieguez & Barden (2020)	Case Control	Competitive swimmers and uninjured controls	12 swimmers with unilateral shoulder pain and 12 pain free controls.	III	
INTERNAL/EXTERNAL ROTATION STRENGTH					
Increased Risk					
Swanik et al (2002)	Randomized Experimental Control Intervention	Competitive Swimmers	13 experimental and 13 controls. (injured)	I	Low
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	
Feijen et al (2021)	Prospective	Club to elite Swimmers	201 Club to elite level swimmers without shoulder pain.	I	
McLaine et al (2019)	Cross sectional	Age group club swimmers	85 club swimmers. A history of pain was reported in 27 swimmers, 58 reported no history of pain.	III	
Drigny et al (2020)	Prospective	Age group club swimmers	18 club swimmers. 39% reported previous history of shoulder pain.	II	
No Association					
Rupp et al (1995)	Case Control	Competitive Swimmers and Matched Controls	22 Cases and 22 Control (14/22 reported previous shoulder pain and 5/22 reported current shoulder pain).	III	
Bak & Magnusson (1997)	Case Control	Competitive Swimmers	7 Swimmers with unilateral shoulder pain and 8 swimmers with no pain.	III	
Harrington et al (2014)	Cross Sectional	Female Collegiate Swimmers	37 Competitive Swimmers (12/37 dominant arm, 14/37 non-dominant arm).	III	
Boettcher et al (2020)	Cross sectional	Elite swimmers	68 elite swimmers (40 male and 28 female). Twenty-four percent reported shoulder pain during competition and training at the time of testing with a further 53% reporting a previous history of shoulder pain.	III	
McKenna et al (2018)	Case controls	Adult swimmers	26 symptomatic and 26 asymptomatic swimmers.	III	

SCAPULAR KINEMATICS, STRENGTH AND DYSKINESIS					
Increased Risk					
Su et al (2004)	Case Control Pre and Post testing	Competitive Age Group Swimmers	20 Swimmers with diagnosed shoulder Impingement and 20 swimmers without.	I	Low
No Association					
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	
Harrington et al (2014)	Cross Sectional	Female Collegiate Swimmers	37 Competitive Swimmers (12/37 dominant arm, 14/37 non-dominant arm).	III	
Welbeck et al (2019)	Cross sectional	Division I Collegiate Swimmers	34 NCAA Division I swimmers (13 male and 21 female)	III	
McLaine et al (2018)	Cross sectional	Age group club swimmers	85 club swimmers. A history of pain was reported in 27 swimmers, 58 reported no history of pain.	III	
CORE STABILITY					
Increased Risk					
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	Low
No Association					
Harrington et al (2014)	Cross Sectional	Female Collegiate Swimmers	37 Competitive Swimmers (12/37 dominant arm, 14/37 non-dominant arm).	III	
GLENOHUMERAL TRANSLATION					
Increased Risk					
None					Low
No Association					
Borsa et al (2005)	Case Control	Collegiate Swimmers and matched Controls	42 Swimmers and 44 Controls (27/42).	III	
Lynch et al (2010)	Randomized Case Control Intervention	Collegiate Swimmers	14 Swimmers in Intervention group and 14 in Control (7/14 in intervention study reported decreased pain, 10/14 in control group reported increased pain).	I	

SWIMMING SCHOLARSHIP STATUS					
Increased Risk					
None					Low
No Association					
Wolf et al (2009)	Retrospective	Collegiate Swimmers	94 Swimmers (number of injuries).	II	
PECTORAL LENGTH					
Increased Risk					
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	Low
Harrington et al (2014)	Cross Sectional	Female Collegiate Swimmers	37 Competitive Swimmers (12/37 dominant arm, 14/37 non-dominant arm)	III	
No Association					
None					
TRICEP LENGTH					
Increased Risk					
None					Low
No Association					
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	
LATISSIMUS LENGTH					
Increased Risk					
Tate et al (2012)	Cross Sectional	Female Swimmers aged 8 to 77 years old	236 Competitive Swimmers (21.4% swimmers aged 8 to 11 years, 18.6% swimmers aged 12 to 14 years, 22.6% high school swimmers, and 19.4% masters swimmers).	III	Low
No Association					
None					
INADEQUATE TREATMENT					

Increased Risk					
Bansal et al (2007)	Cross Sectional	Elite Male Swimmers	161 National and International Level Swimmers (State level swimmers 7/61, national 18/91 and international 3/9 swimmers had impingement syndrome).	III	Low
No Association					
None					
SHOULDER FLEXIBILITY					
Increased Risk					
Griepp (1985)	Prospective	Competitive Swimmers (12-23 years)	168 Swimmers (54% of all females reported shoulder pain over the season, 71.9% of males reported shoulder pain over the season).	II	Low
Ozcaldiran (2002)	Case Control	Competitive Swimmers and Matched Controls	42 Swimmers and 31 Controls (16/42 reported shoulder pain).	III	
No Association					
Stocker et al (1995)	Cross Sectional	Competitive Collegiate and Masters Swimmers	532 Collegiate and 395 Masters level Swimmers (47% of college and 48% of masters swimmers reported pain lasting longer than 3 weeks in the shoulder).	III	
LOWER TRAPEZIUS THICKNESS					
Increased risk					Low
None					
No Association					
McKenna et al (2018)	Case controls	Adult swimmers	26 symptomatic and 26 asymptomatic swimmers.	III	
SERRATUS ANTERIOR THICKNESS					
Increased risk					Low
None					
No Association					
McKenna et al (2018)	Case controls	Adult swimmers	26 symptomatic and 26 asymptomatic swimmers.	III	
ARM SPAN					
Increased risk					Low
None					
No Association					
Feijen et al (2021)	Prospective	Club to elite Swimmers	201 Club to elite level swimmers without shoulder pain.	I	
ACUTE:CHRONIC WORKLOAD					
Increased risk					
Feijen et al (2021)	Prospective	Club to elite Swimmers	201 Club to elite level swimmers without shoulder pain.	I	Low

No Association					
None					
SUB-ACROMIAL BURSA THICKNESS					
Increased risk					
Couanis et al (2015)	Prospective	Open water swimmers	22 open water swimmers measured for sub-acromial bursa thickness.	II	Low
No Association					
Suzuki et al (2020)	Case control	Masters swimmers	20 masters swimmers with shoulder pain, 20 asymptomatic and 20 age-matched controls	III	

APPENDIX B3 – RCT CHAPTER 4 SUPPLEMENTARY TABLE AND FIGURES

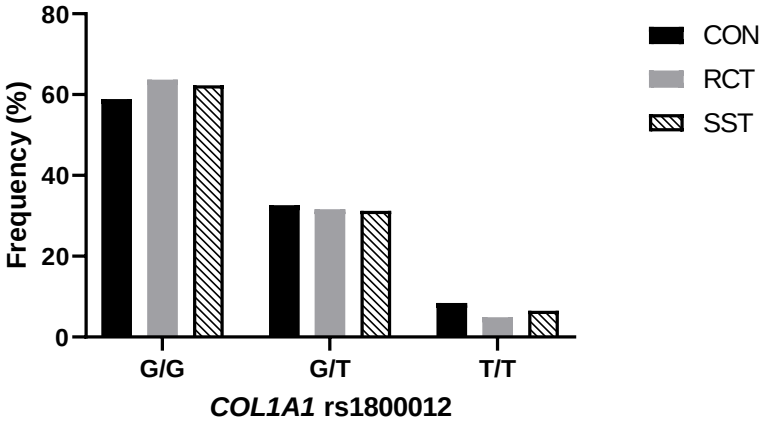


Figure B3.1. Genotypes frequencies are indicated for the *COL1A1* rs1800012 (G/T). Black bars indicate CON group and grey bars indicated RCT group and striped bars indicate supraspinatus tendinopathy sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

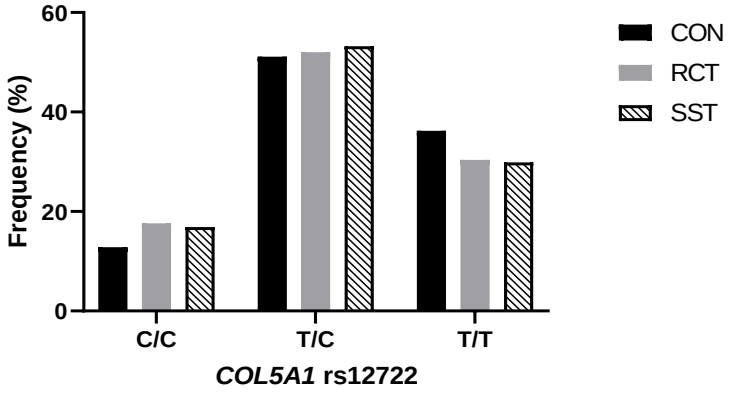


Figure B3.2. Genotypes frequencies are indicated for the *COL5A1* rs12722 (T/C). Black bars indicate CON group and grey bars indicated RCT group and striped bars indicate supraspinatus tendinopathy sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

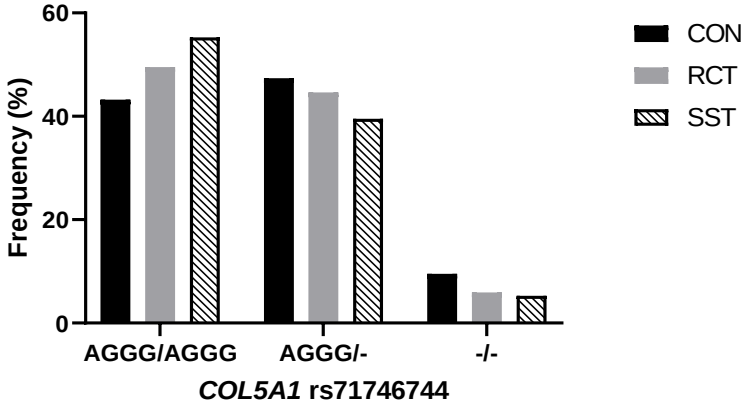


Figure B3.3. Genotypes frequencies are indicated for the *COL5A1* rs71746744 (AGGG/-). Black bars indicate CON group and grey bars indicated RCT group and striped bars indicate supraspinatus tendinopathy sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

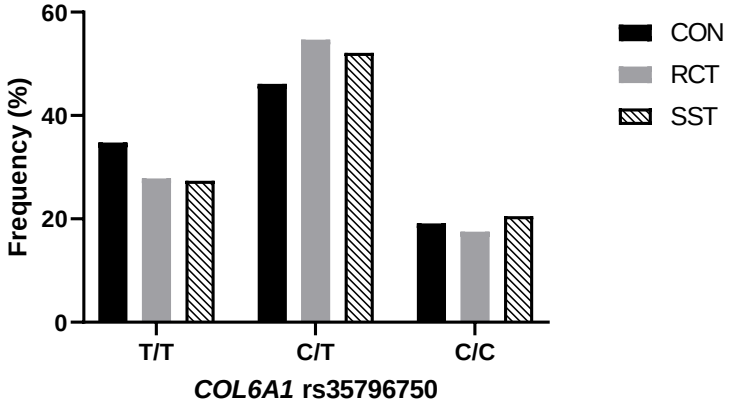


Figure B3.4. Genotypes frequencies are indicated for the *COL6A1* rs35796750 (C/T). Black bars indicate CON group and grey bars indicated RCT group and striped bars indicate supraspinatus tendinopathy sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

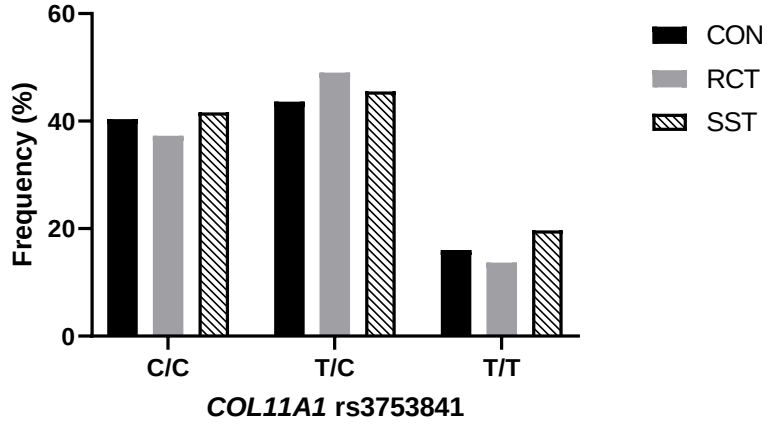


Figure B3.5. Genotypes frequencies are indicated for the *COL11A1* rs3753841 (T/C). Black bars indicate CON group and grey bars indicated RCT group and striped bars indicate supraspinatus tendinopathy sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

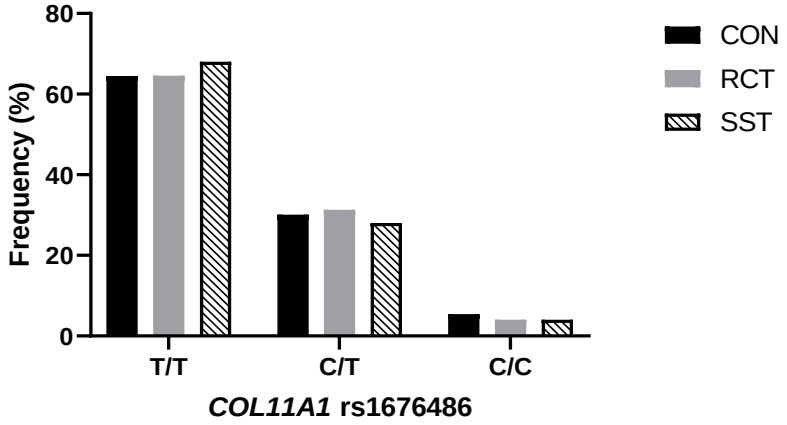


Figure B3.6. Genotypes frequencies are indicated for the *COL11A1* rs1676486 (C/T). Black bars indicate CON group and grey bars indicated RCT group and striped bars indicate supraspinatus tendinopathy sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

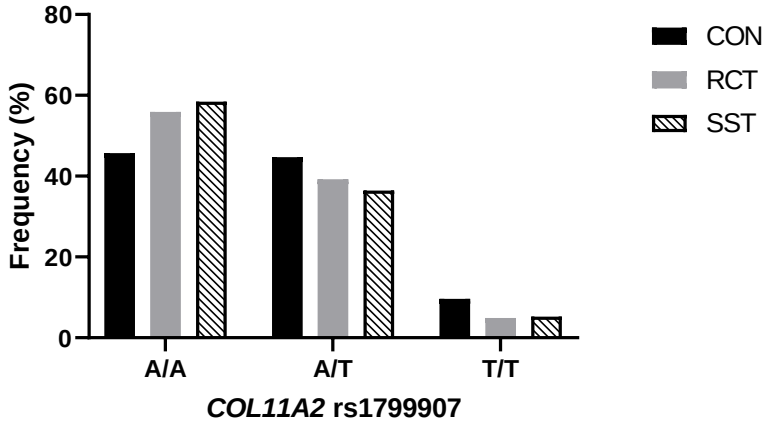


Figure B3.7. Genotypes frequencies are indicated for the *COL11A2* rs1799907 (A/T). Black bars indicate CON group and grey bars indicated RCT group and striped bars indicate supraspinatus tendinopathy sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

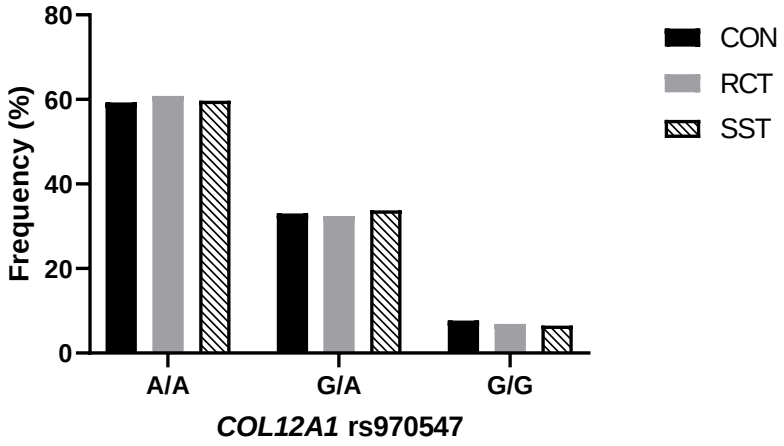


Figure B3.8. Genotypes frequencies are indicated for the *COL12A1* rs970547 (G/A). Black bars indicate CON group and grey bars indicated RCT group and striped bars indicate supraspinatus tendinopathy sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

Table B3.1: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL5A1* rs71746744 polymorphism in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-value
COL5A1	C-(-)	-0.159	54.9	55.2	54.8	0.987	-0.417	55.0	55.1	0.784
rs12722	T-(AGGG)	0.003	26.5	26.4	26.8	0.998	1.214	25.1	23.8	0.970
+	T-(-)	1.284	14.3	11.7	16.7	0.199	-0.769	15.3	19.5	0.198
COL5A1	C-(AGGG)	-1.459	4.3	6.8	1.8	0.145	-0.326	4.6	1.7	0.145
rs71746744										

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.2: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL11A2* rs1799907polymorphism in the control (CON) and Rotator Cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-value
COL11A1	C-T-T	-1.83	4.8	7.0	1.8	0.052	4.8	4.9	1.9	0.068
rs3753841	C-C-T	-1.39	4.1	5.9	2.1	0.164	4.1	4.9	3.6	0.318
+	T-C-T	-0.67	19.1	19.1	20.7	0.502	19.1	18.2	18.0	0.309
COL11A1	T-C-A	0.20	42.6	42.9	40.8	0.838	42.6	44.7	46.0	0.511
rs1676486	C-T-A	0.60	15.2	13.9	17.4	0.550	15.2	14.5	15.7	0.873
+	C-C-A	1.11	14.2	11.3	17.2	0.267	14.2	12.8	14.9	0.403
COL11A2										
rs1799907										

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.3: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL11A1* rs1676486 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-value
COL11A1	C-T	-0.19	20.0	20.9	19.2	0.853	-0.65	19.4	17.6	0.513
rs3753841	T-C	0.41	61.8	62.0	61.5	0.681	0.54	62.9	64.0	0.586
+	C-C	0.64	18.3	17.1	19.3	0.520	0.49	17.7	18.4	0.624
COL11A1	T-T	-	-	-	-	-	-	-	-	-
rs1676486										

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.4: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL11A2* rs1799907 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-value
COL11A1	C-T	-2.18	9.0	12.7	3.9	0.090	-1.94	9.8	5.3	0.052
rs3753841	T-T	-0.69	19.1	19.2	20.7	0.487	-1.03	18.3	18.1	0.303
+	T-A	0.20	42.9	43.0	41.0	0.843	0.65	44.9	46.1	0.518
COL11A2	C-A	0.84	29.1	25.0	34.4	0.402	0.30	27.0	30.4	0.766
rs1799907										

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.5: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486 and *COL11A2* rs1799907 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-value
COL11A1	T-A	-1.58	4.8	7.1	1.8	0.114	-1.79	4.9	1.8	0.074
rs1676486	C-A	-1.17	23.3	24.8	22.7	0.242	-1.34	23.2	21.6	0.182
+	T-T	0.70	15.2	13.3	17.8	0.482	0.35	14.5	16.2	0.724
COL11A2	C-T	0.99	56.7	54.8	57.7	0.322	1.19	57.5	60.4	0.235
rs1799907										

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.6: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL11A2* rs1799907 and *COL5A1* rs71746744 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-value
	C-T-T-(-)	-2.30	1.5	4.1	-	0.021	-2.23	2.1	-	0.026
COL11A1	C-C-T-(-)	-1.36	1.7	2.5	0.8	0.173	-0.90	2.2	1.8	0.367
rs3753841	C-T-T-AGGG	-1.15	3.1	1.3	2.1	0.249	-0.78	2.3	2.1	0.435
+	C-C-T-AGGG	-0.88	2.8	5.0	1.2	0.380	-0.78	3.3	1.4	0.436
COL11A1	T-C-T-(-)	-0.77	5.9	4.1	6.1	0.442	-0.75	5.0	4.7	0.453
rs1676486	T-C-A-(-)	-0.15	11.3	12.9	10.5	0.879	-0.14	11.8	11.1	0.889
+	T-C-T-AGGG	-0.14	13.1	15.1	14.5	0.887	-0.45	13.2	13.5	0.651
COL11A2	C-C-A-(-)	0.21	6.6	7.2	6.6	0.835	-0.34	6.0	5.2	0.731
rs1799907	C-T-A-(-)	0.29	3.7	2.4	4.3	0.771	-0.46	2.6	2.3	0.643
+	T-C-A-AGGG	0.45	31.5	29.8	30.4	0.654	1.03	32.8	34.7	0.055
COL5A1	C-T-A-AGGG	0.50	11.7	12.7	12.9	0.614	0.18	12.3	13.3	0.855
rs71746744		1.90	7.3	2.9	10.7	0.057	1.98	6.4	10.0	0.051
	C-C-A-AGGG									

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.7: Inferred pseudo-haplotype analysis of *COL5A1* rs71746744, *COL11A1* rs3753841 and *COL11A1* rs1676486 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-value
COL11A1 rs3753841	C-T-(-)	-0.81	5.2	6.2	4.2	0.416	-1.58	4.6	2.2	0.115
+	T-C-(-)	-0.58	17.2	17.8	16.8	0.562	-0.69	16.9	16.0	0.490
COL11A1 rs1676486	C-T-(AGGG)	-0.21	14.7	14.5	15.0	0.836	-0.47	14.7	15.5	0.636
+	C-C-(-)	-0.18	8.2	9.2	7.3	0.859	-.0.51	8.1	6.8	0.612
COL5A1 rs71746744	T-C-AGGG	0.40	44.5	44.1	44.7	0.689	0.92	46.0	47.9	0.360
	C-C-AGGG	1.11	10.1	8.2	12.0	0.265	1.07	9.8	11.6	0.284

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B1.8: Inferred pseudo-haplotype analysis of *COL5A1* rs71746744, *COL11A1* rs3753841 and *COL11A1* rs1799907 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-value
	C-A-(-)	-2.52	3.2	6.0	0.0	0.012	-2.12	4.1	0.4	***0.034***
COL11A1 rs3753841	C-A-AGGG	-1.36	5.9	6.6	4.4	0.175	-1.19	5.7	5.0	0.235
+	T-A-(-)	-0.85	5.9	4.8	6.8	0.396	-0.90	5.2	6.0	0.370
COL11A1 rs1799907	T-T-(-)	-0.21	11.7	13.5	10.2	0.835	-0.27	12.1	10.1	0.789
+	T-T-AGGG	-0.11	13.2	14.5	13.3	0.911	-0.37	13.2	12.1	0.715
COL5A1 rs71746744	T-A-AGGG	0.54	31.2	29.4	31.4	0.589	1.21	32.7	36.1	0.059
	C-T-(-)	0.57	9.8	8.8	11.2	0.566	-0.36	8.2	8.5	0.720
	C-T-AGGG	1.02	19.2	16.4	22.7	0.060	0.72	18.9	21.9	0.061

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.9: Inferred pseudo-haplotype analysis of *COL5A1* rs71746744, *COL11A1* rs1676486 and *COL11A1* rs1799907 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-value
	T-T-(-)	-2.07	1.9	4.0	0.0	0.038	-2.18	2.1	0.0	0.029
COL11A1 rs1676486	C-T-(-)	-0.91	6.9	6.5	6.9	0.361	-0.89	7.1	6.6	0.371
+	T-T-AGGG	-0.89	2.5	2.0	1.8	0.371	-0.88	2.4	1.8	0.029
COL11A1 rs1799907	C-T-AGGG	-0.66	16.7	19.5	15.8	0.512	-0.73	16.5	15.0	0.371
+	C-A-(-)	-0.05	18.5	20.5	17.2	0.959	-0.48	18.1	16.2	0.380
COL5A1 rs71746744	T-A-(-)	0.48	3.2%	2.1	4.2	0.628	-0.38	2.3	2.3	0.463
	T-A-AGGG	0.56	12.4	12.5	13.5	0.573	0.36	12.5	14.0	0.634
	C-A-AGGG	1.48	37.8	33.1	40.6	0.091	1.71	39.0	44.2	0.070

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B1.10: Inferred pseudo-haplotype analysis of *COL5A1* rs71746744 and *COL11A1* rs1676486 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-Value
	T-(-)	-0.60	17.6	18.5	16.9	0.550	-0.72	17.3	15.8	0.472
COL11A1 rs3753841	C-(-)	-0.50	13.0	14.7	11.3	0.615	-1.28	12.2	9.2	0.199
+	T-AGGG	0.48	44.3	43.7	44.8	0.632	1.03	45.8	48.5	0.716
COL5A1 rs71746744	C-AGGG	0.72	25.1	23.2	26.9	0.472	0.36	24.6	26.5	0.304

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.11: Inferred pseudo-haplotype analysis of *COL5A1* rs71746744 and *COL11A1* rs3753841 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-Value
	C-(-)	-0.68	25.6	27.5	23.8	0.498	-1.13	25.3	22.9	0.260
COL11A1 rs1676486	T-(-)	-0.60	5.0	5.6	4.4	0.546	-1.48	4.3	2.2	0.140
+	T-AGGG	-0.12	15.0	14.7	15.3	0.901	-0.39	15.0	15.9	0.698
COL5A1 rs71746744	C-AGGG	0.55	54.4	52.1	56.5	0.581	0.91	55.4	59.1	0.361

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.12: Inferred pseudo-haplotype analysis of *COL5A1* rs71746744 and *COL11A2* rs1799907 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-Value
COL11A2 rs1799907	T-(-)	-1.820	9.3	11.3	6.8	0.124	-1.806	9.3	6.5	0.132
+	T-AGGG	-0.866	18.7	20.6	17.7	0.386	-0.968	18.7	16.8	0.333
COL5A1 rs71746744	A-(-)	0.164	20.2	21.8	21.4	0.870	-0.536	20.2	18.5	0.591
	A-AGGG	2.236	51.7	46.2	54.1	0.078	2.048	51.7	58.2	0.072

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.13: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL6A1* rs35796750 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Model	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-Value
COL5A1 rs12722		C-T	-0.686	27.0	29.7	24.2	0.493	-0.405	27.1	23.4	0.685
+		T-C	-0.555	24.5	26	23	0.579	-0.973	23.6	20.4	0.331
COL6A1 rs35796750		C-C	-0.366	32.0	32.1	32.1	0.715	-0.400	32.3	33	0.689
		T-T	0.920	16.5	12.2	20.6	0.098	1.397	17.0	23.2	0.062

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.14: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL12A1* rs970547 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-value
	C-A	-0.915	14.2	16.5	12.3	0.360	-0.487	15.1	13.4	0.626
COL5A1 rs12722	C-G	-0.227	44.7	45.1	44.1	0.821	-0.128	44.3	43.1	0.438
+	T-A	0.686	9.4	7.6	10.8	0.493	0.446	8.7	10	0.655
COL12A1 rs970547	T-G	1.005	31.7	30.7	32.9	0.315	0.182	31.9	33.5	0.383

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.15: Inferred pseudo-haplotype analysis of *COL6A1* rs35796750 and *COL12A1* rs970547 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-value
	T-G	-1.208	10.8	7.6	7.3	0.227	0.446	8.7	10	0.655
COL6A1 rs35796750	C-A	0.281	43.7	45.1	39.5	0.426	-0.128	44.3	43.1	0.438
+	C-G	0.666	12.8	16.5	15.7	0.505	-0.487	15.1	13.4	0.626
COL12A1 rs970547	T-A	1.719	32.8	30.7	37.4	0.386	0.182	31.9	33.5	0.383

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Models: **Add** = **Add**itive; **Dom** = **Dom**inant; **Rec** = **Rec**essive.
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B3.16: Inferred pseudo-haplotype analysis of *COL5A1* rs12772, *COL6A1* rs35796750 and *COL12A1* rs970547 in the control (CON) and Rotator cuff tendinopathy (RCT) and Supraspinatus Tendinopathy (SST) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	RCT (103)	P-value	Hap Score	Hap Freq	SST (87)	P-value
COL5A1 rs12772	C-T-G	-1.622	7.0	9.8	4.3	0.105	-1.451	7.3	4.3	0.146
+	C-C-A	-0.611	24.8	24.9	25.3	0.541	-0.644	24.4	24.4	0.519
COL6A1 rs35796750	T-T-G	-0.276	3.8	5.1	3	0.783	-0.157	4.1	3.6	0.875
+	T-C-A	-0.020	18.9	24	14.3	0.052	-0.347	19.0	13.4	0.729
COL12A1 rs970547	C-C-G	0.180	7.2	7.2	7.8	0.857	0.384	7.1	9	0.700
	C-T-A	0.491	20.0	19.9	18.9	0.623	0.529	19.8	18.7	0.597
	T-C-G	1.429	5.6	2	7.9	0.153	1.173	4.6	6.5	0.241
	T-T-A	1.827	12.8	7	18.5	0.068	2.027	12.9	20.1	0.053

CON group, RCT group and SST sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterix (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in RCT.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

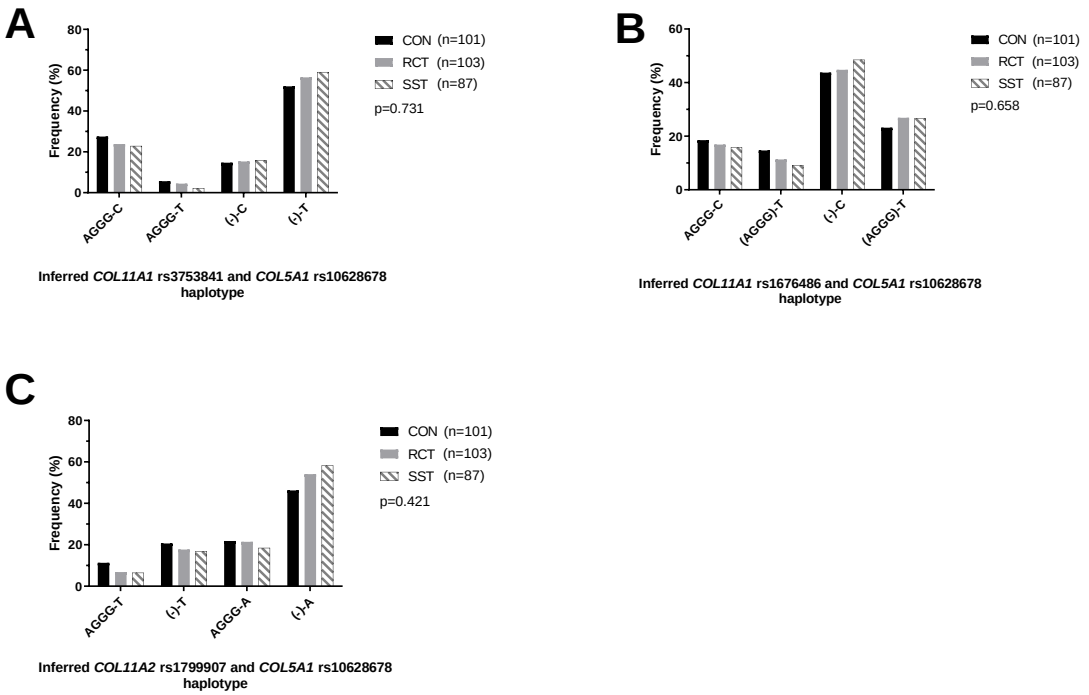


Figure B3.9 A) Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and *COL5A1* rs10628678 (AGGG/-). **B)** Inferred haplotype frequency distributions constructed from *COL11A1* rs1676486 (C/T) and *COL5A1* rs10628678 (AGGG/-) polymorphisms. **C)** Inferred haplotype frequency distributions constructed from *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) for control group (CON, black bars) and the rotator cuff tendinopathy (RCT, grey bars) groups and supraspinatus tendinopathy (SST, hatched bars). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

APPENDIX B4 – CHAPTER 5 SWEDISH ACL AND COMBINED SUPPLEMENTARY GRAPHS AND TABLES

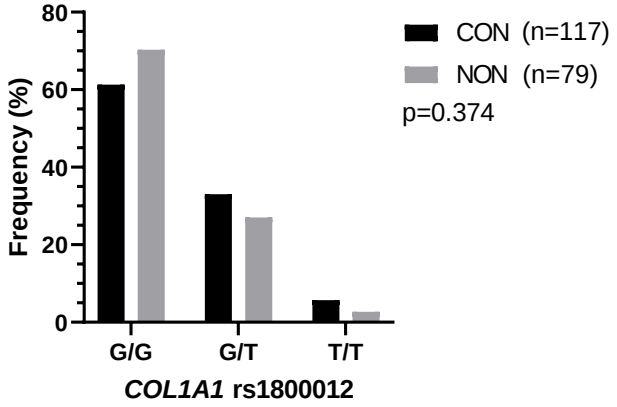


Figure B4.1. Genotypes frequencies are indicated for the *COL1A1* rs1800012 (G/T). Black bars indicate CON group and grey bars indicate NON. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

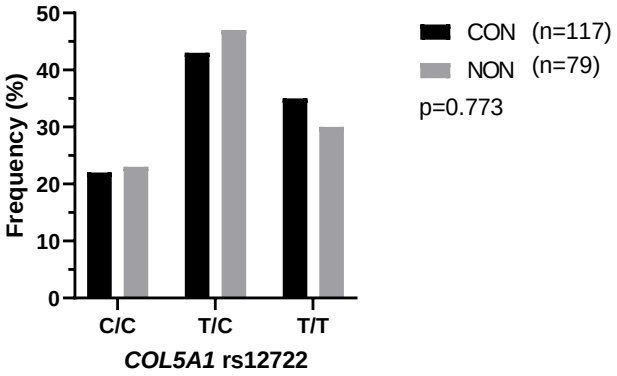


Figure B4.2. Genotypes frequencies are indicated for the *COL5A1* rs12722 (T/C). Black bars indicate CON group and grey bars indicate NON. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

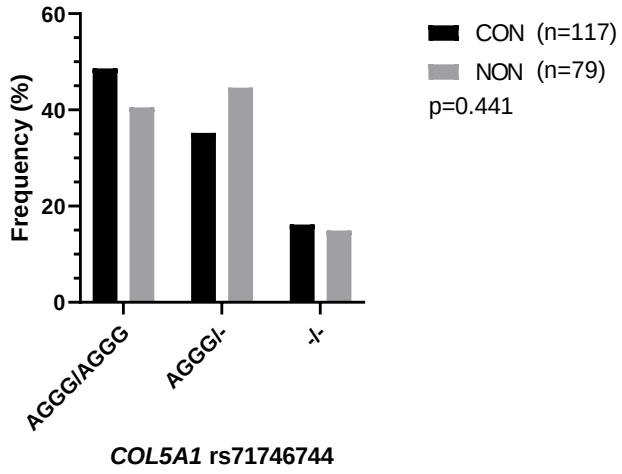


Figure B4.3. Genotypes frequencies are indicated for the *COL5A1* rs71746744 (AGGG/-). Black bars indicate CON group and grey bars indicate NON. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

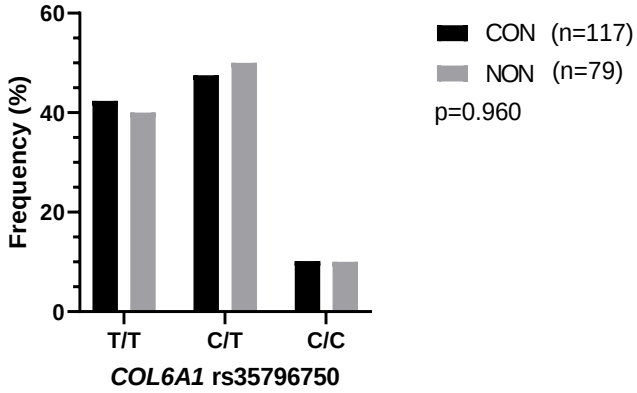


Figure B4.4. Genotypes frequencies are indicated for the *COL6A1* rs35796750 (C/T). Black bars indicate CON group and grey bars indicate NON. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

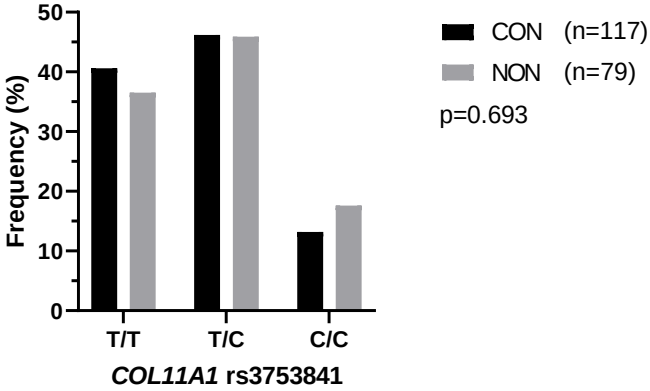


Figure B4.5. Genotypes frequencies are indicated for the *COL11A1* rs3753841 (T/C). Black bars indicate CON group and grey bars indicate NON. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

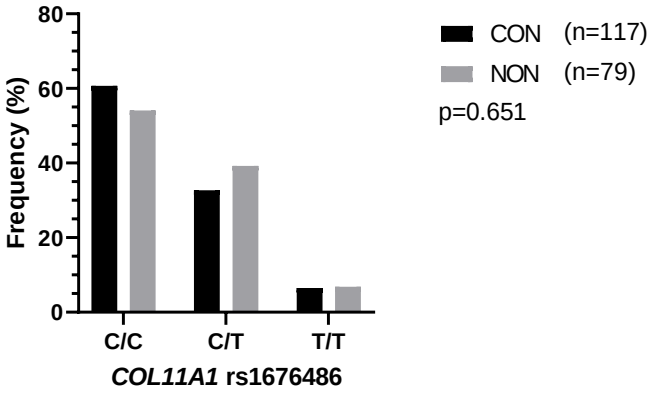


Figure B4.6. Genotypes frequencies are indicated for the *COL11A1* rs1676486 (C/T). Black bars indicate CON group and grey bars indicate NON. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

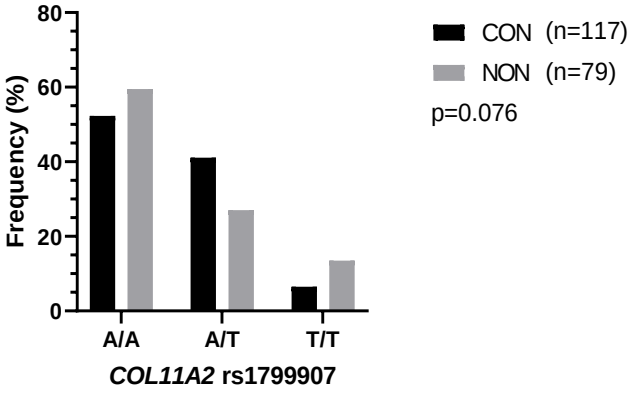


Figure B4.7. Genotypes frequencies are indicated for the *COL11A2* rs1799907 (C/T). Black bars indicate CON group and grey bars indicate NON. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

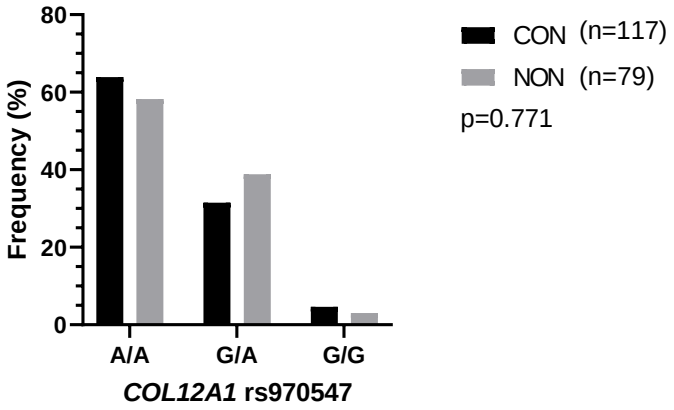


Figure B4.8. Genotypes frequencies are indicated for the *COL12A1* rs970547 (G/A). Black bars indicate CON group and grey bars indicate NON. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

Table B4.1: A sex-specific comparison of the genotype frequency distributions for the *COL1A1* rs1800012 (G/T) and *COL5A1* rs12722 (T/C), between ACL injury and control groups within the South African, Australian, Polish and Swedish cohorts

	COL1A1 rs1800012 genotype							*COL5A1* rs12722 genotype						
	n	G/G	G/T	T/T	p-value	Minor Allele	Allele p-value	n	C/C	T/C	T/T	p-value	Minor Allele	Allele p-value
All Participants														
Combined CON	572	67.0 (383)	29.0 (166)	4.0 (23)	**0.027***	18.5 (212)	0.661	728	20.3 (148)	46.7 (340)	33.0 (240)	0.060	44.2 (642)	0.325
Combined ACL	375	68.0 (255)	30.9 (116)	1.1 (4)		16.6 (124)		708	23.6 (167)	50.8 (360)	25.6 (181)		48.1 (695)	
SA CON	323	68.7 (222)	27.9 (90)	3.4 (11)	0.091	17.4 (112)	0.944	404	18.3 (74)	46.8 (189)	34.9 (141)	0.383	41.7 (337)	0.931
SA ACL	211	65.4 (138)	33.6 (71)	0.9 (2)		17.8 (75)		209	15.8 (33)	52.6 (110)	41.6 (66)		42.1 (176)	
Swe CON	106	61.3 (65)	33.0 (35)	5.7 (6)	0.556	22.2 (47)	0.160	101	24.8 (25)	36.6 (36)	38.6 (39)	0.291	43.1 (87)	0.688
Swe ACL	73	71.2 (52)	26.0 (19)	2.7 (2)		15.8 (23)		69	20.3 (14)	52.2 (36)	27.5 (19)		46.4 (64)	
Pol CON	143	67.1 (96)	28.7 (41)	4.2 (6)	0.138	18.5 (53)	0.643	143	21.0 (30)	52.4 (75)	26.6 (38)	0.343	47.2 (135)	0.367
POL ACL	91	71.4 (65)	28.6 (26)	0.0 (0)		14.3 (26)		91	16.5 (15)	48.4 (44)	35.2 (32)		40.7 (74)	
AUS CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	79	27.8 (22)	50.6 (40)	21.5 (17)	0.974	47.0 (74)	0.930
AUS ACL	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	354	29.1 (103)	49.4 (175)	21.5 (76)		46.0 (327)	
Male Participants														
Combined CON	335	66.5 (219)	31.0 (104)	3.6 (12)	0.299	19.1 (128)	0.668	507	20.5 (104)	48.3 (245)	32.2 (158)	0.238	45.1 (437)	0.513
Combined ACL	259	68.0 (176)	30.5 (79)	1.5 (4)		16.8 (87)		425	22.4 (95)	53.9 (229)	23.8 (101)		48.5 (432)	
SA CON	201	69.2 (139)	27.9 (56)	3.0 (6)	0.488	16.9 (68)	0.988	272	15.8 (43)	48.9 (133)	35.3 (96)	0.356	40.3 (219)	0.416
SA ACL	155	67.7 (105)	31.0 (48)	1.3 (2)		16.8 (52)		151	17.9 (27)	53.6 (81)	28.5 (43)		44.7 (135)	
Swe CON	35	42.9 (15)	48.6 (17)	8.6 (3)	0.067	32.9 (23)	0.364	34	35.3 (12)	35.3 (12)	29.4 (10)	0.489	52.9 (36)	0.474
Swe ACL	33	69.7 (23)	24.2 (8)	6.1 (5)		18.2 (12)		33	18.2 (6)	51.5 (12)	30.3 (10)		43.9 (29)	
Pol CON	99	65.7 (65)	31.3 (31)	3.0 (3)	0.335	18.7 (37)	0.807	99	23.2 (23)	52.5 (52)	24.2 (24)	0.328	49.5 (98)	0.332
POL ACL	71	67.6 (48)	32.4 (23)	0.0 (0)		16.2 (23)		71	16.9 (12)	49.3 (35)	33.8 (24)		41.5 (59)	
AUS CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	79	27.8 (22)	50.6 (40)	21.5 (17)	0.841	47.0 (74)	0.777
AUS ACL	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	195	28.2 (55)	53.3 (104)	18.5 (36)		45.0 (176)	
Female Participants														
Combined CON	230	68.3 (157)	27.0 (62)	4.8 (11)	**0.045***	18.3 (84)	0.760	221	19.9 (44)	43.0 (95)	37.1 (82)	0.055	47.9 (229)	0.078
Combined ACL	116	68.1 (79)	31.9 (37)	0.0 (0)		16.0 (37)		283	25.4 (72)	46.3 (131)	28.3 (80)		36.2 (71)	
SA CON	118	66.9 (79)	28.8 (34)	4.2 (5)	0.106	18.6 (44)	0.853	128	33.6 (43)	43.0 (55)	23.4 (30)	**0.002***	44.9 (115)	0.287
SA ACL	56	58.9 (33)	41.1 (23)	0.0 (0)		20.5 (23)		58	10.3 (6)	50.0 (29)	39.7 (23)		35.3 (41)	
Swe CON	68	69.1 (47)	26.5 (18)	4.4 (3)	0.431	17.6 (24)	0.669	67	19.4 (13)	37.3 (25)	43.3 (29)	0.172	38.1 (51)	0.316
Swe ACL	37	73.0 (27)	27.0 (10)	0.0 (0)		13.5 (35)		36	22.2 (8)	52.8 (19)	25.0 (9)		48.6 (35)	
Pol CON	44	70.5 (31)	22.7 (10)	6.8 (3)	0.340	18.2 (16)	0.656	44	15.9 (7)	52.3 (23)	31.8 (14)	0.810	42.0 (16)	0.803
POL ACL	20	85.0 (17)	15.0 (3)	0.0 (0)		7.5 (3)		20	15.0 (3)	45.0 (9)	40.0 (8)		37.5 (15)	
AUS CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
AUS ACL	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

Values are expressed as percentages with the number of participants in the parentheses. The total number (n) of participants genotyped in each individual or combined cohorts is also indicated. Bold typeset indicates significant p-values.
HWE, Hard-Weinberg Equilibrium; n.d., not determined.

Table B4.2: A sex-specific comparison of the genotype frequency distributions for the *COL5A1* rs10628678 (AGGG/-) and *COL6A1* rs35796750 (T/C), between ACL injury and control groups within the South African, Australian, Polish and Swedish cohorts

	COL5A1 rs10628678 genotype							*COL6A1* rs35796750 genotype						
	n	AGGG/AGGG	AGGG/-	-/-	p-value	Minor Allele	Allele p-value	n	T/T	T/C	C/C	p-value	Minor Allele	Allele p-value
All Participants														
Combined CON	544	52.2 (284)	37.7 (205)	10.1 (55)	0.074	29.0 (315)	0.254	234	36.8 (86)	46.6 (109)	16.7 (39)	0.394	40.0 (187)	0.832
Combined ACL	630	45.9 (289)	44.0 (277)	10.2 (64)		32.1 (405)		160	33.5 (56)	52.5 (84)	12.5 (20)		38.8 (124)	
SA CON	359	52.1 (187)	39.8 (143)	8.1 (29)	0.336	28.0 (201)	0.774	175	34.9 (61)	46.3 (81)	18.9 (33)	0.368	42.0 (147)	0.895
SA ACL	202	47.5 (96)	46.0 (93)	6.4 (13)		29.5 (119)		101	31.7 (32)	54.6 (55)	13.9 (14)		41.1 (83)	
Swe CON	105	48.6 (51)	35.2 (37)	16.2 (17)	0.501	33.8 (71)	0.510	59	42.4 (25)	47.5 (28)	10.2 (6)	0.981	33.9 (40)	0.857
Swe ACL	73	41.1 (30)	43.8 (32)	15.1 (11)		36.9 (54)		59	40.7 (24)	49.2 (29)	10.2 (6)		34.7 (41)	
POL CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
POL ACL	n.d.	n.d.	n.d.	n.d.				n.d.	n.d.	n.d.	n.d.			
AUS CON	80	57.5 (46)	31.2 (25)	11.2 (9)	0.139	27.0 (43)	0.187	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
AUS ACL	356	46.1 (164)	42.7 (152)	11.2 (40)		33.0 (232)		n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Male Participants														
Combined CON		52.4 (188)	39.0 (140)	8.6 (31)	0.652	28.1 (202)	0.844	125	36.8 (46)	44.0 (55)	19.2 (24)	0.220	41.2 (103)	0.670
Combined ACL		49.1 (185)	42.2 (159)	8.8 (33)		29.8 (225)		96	35.4 (34)	53.1 (51)	11.5 (11)		38.0 (73)	
SA CON	245	52.2 (128)	40.8 (100)	6.9 (17)	0.406	27.3 (134)	0.593	107	35.5 (38)	43.9 (47)	20.6 (22)	**0.000***	42.5 (91)	0.830
SA ACL	147	45.6 (67)	47.6 (70)	6.8 (10)		30.6 (90)		70	30.0 (21)	58.6 (41)	11.4 (8)		40.7 (57)	
Swe CON	34	41.2 (14)	44.1 (15)	14.7 (5)	0.698	36.8 (25)	0.651	18	44.4 (8)	44.4 (8)	11.1 (2)	0.922	33.3 (12)	0.890
Swe ACL	33	51.5 (17)	36.4 (12)	12.1 (4)		30.3 (20)		26	50.0 (13)	38.5 (10)	11.5 (3)		30.8 (16)	
POL CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
POL ACL	n.d.	n.d.	n.d.	n.d.				n.d.	n.d.	n.d.	n.d.			
AUS CON	80	57.5 (46)	31.2 (25)	11.2 (9)	0.465	27.0 (43)	0.605	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
AUS ACL	197	51.3 (115)	39.1 (77)	9.6 (19)		29.0 (115)		n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Female Participants														
Combined CON	177	52.0 (92)	36.2 (64)	11.9 (21)	0.259	29.9 (106)	0.655	107	36.4 (39)	50.5 (54)	13.1 (14)	0.958	38.3 (82)	0.864
Combined ACL	95	44.2 (42)	45.3 (43)	10.5 (10)		33.2 (63)		64	34.4 (22)	51.6 (33)	14.1 (9)		39.8 (51)	
SA CON	110	50.9 (56)	39.1 (43)	10.0 (11)	0.611	29.5 (65)	0.752	68	33.8 (23)	50.0 (34)	16.2 (11)	0.885	41.2 (56)	0.953
SA ACL	55	52.7 (29)	41.8 (23)	5.5 (3)		26.3 (29)		31	35.5 (11)	45.2 (14)	19.4 (6)		41.9 (26)	
Swe CON	67	53.7 (36)	31.3 (21)	14.9 (10)	0.107	30.6 (41)	0.269	39	41.0 (16)	51.3 (20)	7.7 (3)	0.636	33.3 (26)	0.627
Swe ACL	37	32.4 (12)	48.6 (18)	18.9 (7)		43.2 (32)		30	30.0 (9)	60.0 (18)	10.0 (3)		40.0 (24)	
POL CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
POL ACL	n.d.	n.d.	n.d.	n.d.				n.d.	n.d.	n.d.	n.d.			
AUS CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
AUS ACL	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

Values are expressed as percentages with the number of participants in the parentheses. The total number (n) of participants genotyped in each individual or combined cohorts is also indicated. Bold typeset indicates significant p-values. HWE, Hard-Weinberg Equilibrium; n.d., not determined.

Table B4.3: A sex-specific comparison of the genotype frequency distributions for the *COL11A1* rs3753841 (T/C) and *COL11A1* rs1676486 (C/T) between ACL injury and control groups within the South African, Australian, Polish and Swedish cohorts

COL11A1 rs3753841 genotype								*COL11A1* rs1676486 genotype						
	n	T/T	T/C	C/C	p-value	Minor Allele	Allele p-value	n	C/C	C/T	T/T	p-value	Minor Allele	Allele p-value
All Participants														
Combined CON	312	34.3 (107)	52.6 (164)	13.1 (41)	0.131	39.4 (246)	0.826	316	59.8 (189)	34.8 (110)	5.4 (17)	0.440	22.8 (144)	0.604
Combined ACL	275	37.1 (102)	45.1 (124)	17.8 (49)		40.4 (222)		276	64.9 (179)	30.1 (83)	5.1 (14)		20.1 (111)	
SA CON	206	31.1 (64)	55.8 (115)	13.1 (27)	0.088	41.0 (169)	0.897	209	59.3 (124)	35.9 (75)	4.8 (10)	0.142	22.7 (95)	0.457
SA ACL	202	37.1 (75)	45.1 (91)	17.8 (36)		40.3 (163)		203	68.5 (139)	27.1 (55)	4.4 (9)		18.0 (73)	
Swe CON	106	40.6 (43)	46.2 (49)	13.2 (14)	0.685	36.3 (77)	0.305	107	60.7 (65)	32.7 (35)	6.5 (7)	0.651	22.9 (49)	0.472
Swe ACL	73	37.0 (27)	45.2 (33)	17.8 (13)		40.4 (59)		73	54.8 (40)	38.4 (28)	6.8 (5)		26.0 (38)	
POL CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
POL ACL	n.d.	n.d.	n.d.	n.d.				n.d.	n.d.	n.d.	n.d.			
Male Participants														
Combined CON	159	32.1 (51)	54.1 (86)	13.8 (22)	0.316	40.9 (130)	0.987	159	58.5 (93)	35.2 (56)	6.3 (10)	0.331	23.9 (76)	0.384
Combined ACL	180	36.1 (65)	46.1 (83)	17.8 (32)		40.8 (147)		181	66.3 (120)	28.7 (52)	5.0 (9)		19.3 (70)	
SA CON	124	27.4 (34)	58.9 (73)	13.7 (17)	0.147	43.1 (107)	0.772	124	55.6 (69)	37.9 (47)	6.5 (8)	0.098	25.4 (63)	0.333
SA ACL	147	35.4 (52)	46.9 (69)	17.7 (26)		41.2 (121)		148	68.2 (101)	27.7 (41)	4.1 (6)		17.9 (53)	
Swe CON	35	48.6 (17)	37.1 (13)	14.3 (5)	0.740	32.9 (23)	0.640	35	68.6 (24)	25.7 (9)	5.7 (2)	0.630	18.6 (13)	0.646
Swe ACL	33	39.4 (13)	42.4 (14)	18.2 (6)		39.4 (26)		33	57.6 (19)	33.3 (11)	9.1 (3)		25.8 (17)	
POL CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
POL ACL	n.d.	n.d.	n.d.	n.d.				n.d.	n.d.	n.d.	n.d.			
Female Participants														
Combined CON	149	36.2 (54)	51.0 (76)	12.8 (19)	0.387	38.3 (114)	0.880	153	61.4 (94)	34.0 (52)	4.6 (7)	0.954	21.6 (66)	0.998
Combined ACL	95	38.9 (37)	43.2 (41)	17.9 (17)		39.5 (75)		95	62.1 (59)	32.6 (31)	5.3 (5)		21.6 (41)	
SA CON	85	36.6 (30)	51.2 (42)	12.2 (10)	0.381	37.8 (62)	0.967	85	64.7 (55)	32.9 (28)	2.4 (2)	0.445	18.8 (32)	0.957
SA ACL	55	41.8 (23)	40.0 (22)	18.2 (10)		38.2 (42)		55	69.1 (38)	25.5 (14)	5.5 (3)		18.2 (20)	
Swe CON	67	45.8 (24)	50.7 (34)	13.4 (9)	0.927	38.8 (52)	0.880	68	57.4 (39)	35.3 (24)	7.4 (5)	0.509	25.0 (34)	0.956
Swe ACL	37	35.1 (13)	48.6 (18)	16.2 (6)		40.5 (30)		37	54.1 (20)	43.3 (16)	2.7 (1)		24.3 (18)	
POL CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
POL ACL	n.d.	n.d.	n.d.	n.d.				n.d.	n.d.	n.d.	n.d.			

Values are expressed as percentages with the number of participants in the parentheses. The total number (n) of participants genotyped in each individual or combined cohorts is also indicated. Bold typeset indicates significant p-values.
HWE, Hard-Weinberg Equilibrium; n.d., not determined.

Table B4.4: A sex-specific comparison of the genotype frequency distributions for the *COL11A2* rs1799907 (A/T) and *COL12A1* rs970547 (G/A) between ACL injury and control groups within the South African, Australian, Polish and Swedish cohorts

	COL11A2 rs1799907 genotype							*COL12A1* rs970547 genotype						
	n	A/A	A/T	T/T	p-value	Minor Allele	Allele p-value	n	A/A	G/A	G/G	p-value	Minor Allele	Allele p-value
All Participants														
Combined CON	299	46.5 (139)	43.1 (129)	10.4 (31)	0.764	31.9 (191)	0.668	447	60.4 (270)	35.3 (158)	4.3 (19)	0.423	21.9 (196)	0.790
Combined ACL	262	43.5 (114)	45.0 (118)	11.5 (30)		34.0 (178)		327	63.7 (223)	31.3 (109)	5.1 (18)		20.7 (145)	
SA CON	192	43.2 (83)	44.3 (85)	12.5 (24)	0.333	34.6 (133)	0.744	203	58.1 (118)	37.9 (77)	3.9 (8)	0.133	22.9 (93)	0.824
SA ACL	189	37.6 (71)	51.9 (98)	10.6 (20)		36.5 (138)		193	63.7 (123)	29.5 (57)	6.7 (13)		21.5 (83)	
Swe CON	107	52.3 (56)	41.1 (44)	6.5 (7)	0.082	27.1 (58)	0.884	101	62.4 (63)	33.7 (34)	4.0 (4)	0.832	20.8 (42)	0.834
Swe ACL	73	58.9 (43)	27.4 (20)	13.7 (10)		27.8 (40)		66	59.1 (39)	37.9 (25)	3.0 (2)		22.0 (29)	
POL CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	143	62.2 (89)	32.9 (47)	4.9 (7)	0.701	21.3 (61)	0.714
POL ACL	n.d.	n.d.	n.d.	n.d.				91	67.0 (61)	29.7 (27)	3.3 (3)		18.1 (33)	
Male Participants														
Combined CON	150	47.3 (71)	42.0 (63)	10.7 (16)	0.606	31.7 (95)	0.712	253	64.4 (163)	32.0 (81)	3.6 (9)	0.656	19.6 (99)	0.739
Combined ACL	173	42.2 (73)	47.4 (82)	10.4 (18)		34.1 (118)		240	61.7 (143)	33.8 (81)	4.6 (11)		21.5 (103)	
SA CON	115	45.2 (52)	43.5 (50)	11.3 (13)	0.431	33.0 (76)	0.639	121	58.7 (71)	37.2 (45)	4.1 (5)	0.623	22.7 (55)	0.918
SA ACL	140	37.9 (53)	51.4 (72)	10.7 (15)		36.4 (102)		139	51.9 (86)	32.4 (45)	5.8 (8)		21.9 (61)	
Swe CON	35	54.3 (19)	37.1 (13)	8.6 (3)	0.836	27.1 (19)	0.847	33	75.8 (25)	24.2 (8)	0.0 (0)	0.275	12.1 (8)	0.581
Swe ACL	33	60.6 (20)	30.3 (10)	9.1 (3)		24.2 (16)		30	56.7 (17)	43.3 (13)	0.0 (0)		21.7 (13)	
POL CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	99	67.7 (67)	28.3 (28)	4.0 (4)	0.839	18.2 (36)	0.817
POL ACL	n.d.	n.d.	n.d.	n.d.				71	63.4 (45)	32.4 (23)	4.2 (3)		20.4 (29)	
Female Participants														
Combined CON	145	44.1 (64)	45.5 (66)	10.3 (15)	0.657	33.1 (96)	0.939	194	55.2 (107)	39.7 (77)	5.2 (10)	0.074	25.0 (97)	0.451
Combined ACL	89	46.1 (41)	40.4 (36)	13.5 (12)		33.7 (60)		110	68.2 (75)	25.5 (28)	6.4 (7)		19.1 (42)	
SA CON	77	40.3 (31)	45.5 (35)	14.3 (11)	0.655	37.0 (57)	0.977	82	57.3 (47)	39.0 (32)	3.7 (3)	0.073	23.7 (38)	0.770
SA ACL	49	36.7 (18)	53.1 (26)	10.2 (5)		36.7 (36)		54	68.5 (37)	22.2 (12)	9.3 (5)		20.4 (22)	
Swe CON	68	48.5 (33)	45.6 (31)	5.9 (4)	**0.046***	28.7 (39)	0.758	68	55.9 (38)	38.2 (26)	5.9 (4)	0.602	25.0 (34)	0.831
Swe ACL	37	54.1 (20)	27.0 (10)	18.9 (7)		32.4 (24)		36	61.1 (22)	33.3 (12)	5.6 (4)		22.2 (16)	
POL CON	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	44	50.0 (22)	43.2 (19)	6.8 (3)	0.064	28.4 (25)	0.443
POL ACL	n.d.	n.d.	n.d.	n.d.				20	80.0 (16)	20.0 (4)	0.0 (0)		10 (4)	

Values are expressed as percentages with the number of participants in the parentheses. The total number (n) of participants genotyped in each individual or combined cohorts is also indicated. Bold typeset indicates significant p-values.
HWE, Hard-Weinberg Equilibrium; n.d., not determined.

Appendix B4 – Swedish Cohort supplementary tables

Table B4.5: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL5A1* rs71746744 polymorphism in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-value
COL5A1 rs12722 + *COL5A1* rs71746744	T-(-)	-0.78	47.6	25.7	30.1	0.435
	T-(AGGG)	-0.52	27.	18.0	16.4	0.602
	C-(AGGG)	-0.43	17.4	48.4	46.5	0.668
	C-(-)	1.03	7.6	7.9	7.1	0.305

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.6: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL11A2* rs1799907polymorphism in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-value
COL11A1 rs3753841	C-C-A	-0.67	9.7	10.6	8.5	0.501
+	T-C-A	-0.61	45.5	47.2	43.5	0.541
COL11A1 rs1676486	C-T-T	-0.06	5.9	6.5	5.3	0.749
+	C-C-T	0.06	5.4	5.1	5.7	0.953
COL11A2 rs1799907	T-C-T	0.32	15.1	14.2	16.0	0.950
	C-T-A	1.31	17.4	14.5	21.1	0.191

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.7: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL11A1* rs1676486in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-value
COL11A1 rs3753841	C-C	-0.14	15.1	15.7	14.2	0.885
+	T-C	-0.11	60.6	61.4	59.5	0.911
COL11A1 rs1676486	C-T	1.06	23.3	21.1	26.4	0.290
	T-T	-	-	-	-	-

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.8: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL11A2* rs1799907 in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-value
COL11A1 rs3753841	T-A	-0.66	45.6	47.1	43.6	0.509
+	T-T	-0.50	16.3	16.5	15.8	0.615
COL11A2 rs1799907	C-A	0.44	27.3	25.8	29.3	0.657
	C-T	0.42	10.8	10.6	11.2	0.671

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.9: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486 and *COL11A2* rs1799907 in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-value
COL11A1 rs1676486	C-A	-1.24	55.5	58.1	52.1	0.216
+	C-T	-0.67	20.2	19.0	21.6	0.503
COL11A2 rs1799907	T-A	0.14	6.9	8.1	5.5	0.885
	T-T	1.16	17.4	14.8	20.9	0.247

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.10: Inferred pseudo-haplotype analysis of *COL5A1* rs71746744, *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL11A2* rs1799907in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-value
	T-C-A-AGGG	-0.98	29.8	33.3	25.1	0.236
COL11A1 rs3753841	C-C-A-AGGG	-1.18	6.5	7.8	4.5	0.327
+	T-C-T-AGGG	-0.14	9.2	7.5	11.0	0.620
COL11A1 rs1676486	C-T-T-AGGG	-0.45	3.0	4.6	1.3	0.650
+	C-C-A-(-)	0.13	3.2	2.9	4.0	0.887
COL11A2 rs1799907	C-T-A-(-)	-0.50	5.0	5.7	3.1	0.896
+	C-C-T-AGGG	0.28	2.9	2.5	2.9	0.882
COL5A1 rs71746744	T-C-T-(-)	0.15	5.9	6.7	5.2	0.776
	C-C-T-(-)	0.49	2.4	2.5	2.7	0.625
	T-C-A-(-)	0.73	15.7	14.0	18.2	0.494
	C-T-T-(-)	0.68	2.9	2.1	4.0	0.468
	C-T-A-AGGG	1.29	12.4	8.7	18.0	0.198

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON group.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.11: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL5A1* rs71746744 in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-value
COL11A1 rs3753841	T-C-AGGG	-0.75	39.0	40.7	36.3	0.451
+	C-C-AGGG	-0.56	9.4	10.3	7.8	0.579
COL11A1 rs1676486	C-C-(-)	0.22	5.7	5.4	6.4	0.824
+	C-T-(-)	0.55	7.9	7.8	7.5	0.580
COL5A1 rs71746744	C-T-AGGG	0.76	15.4	13.4	18.8	0.446
	T-C-(-)	0.90	21.6	20.7	23.2	0.366

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.12: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1799907 and *COL5A1* rs71746744 in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-value
COL11A1 rs3753841	T-A-AGGG	-1.27	30.0	33.1	25.0	0.205
+	T-T-AGGG	-0.97	10.1	9.4	10.7	0.332
COL11A1 rs1799907	C-T-AGGG	0.00	5.7	7.0	4.8	0.998
+	T-T-(-)	0.11	6.2	7.1	5.0	0.913
COL5A1 rs71746744	C-A-AGGG	0.34	18.9	16.5	22.4	0.736
	C-A-(-)	0.45	8.4	9.2	6.9	0.655
	T-A-(-)	0.72	15.6	14.0	18.8	0.469
	C-T-(-)	1.01	5.1	3.5	6.5	0.314

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.13: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486, *COL11A1* rs1799907 and *COL5A1* rs71746744 in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-value
COL11A1 rs1676486	C-A-AGGG	-1.29	36.4	41.1	29.8	0.197
+	C-T-AGGG	-0.89	11.8	9.8	13.7	0.051
COL11A1 rs1799907	T-T-(-)	-0.68	4.1	6.1	1.3	0.496
+	T-A-(-)	0.13	4.9	5.6	2.9	0.893
COL5A1 rs71746744	C-T-(-)	0.17	8.4	9.1	7.9	0.866
	C-A-(-)	0.58	19.1	17.0	22.2	0.559
	T-T-(-)	0.73	2.8	2.1	4.1	0.464
	T-A-AGGG	1.12	12.5	9.1	18.0	0.263

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.14: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486 and *COL5A1* rs71746744 in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-Value
COL11A1 rs3753841	T-AGGG	-1.08	40.2	42.7	36.4	0.279
+	C-AGGG	0.32	24.6	23.5	26.5	0.748
COL5A1 rs71746744	T-(-)	0.82	21.8	20.9	23.1	0.412
	C-(-)	1.01	13.4	12.9	14.1	0.315

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.15: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL5A1* rs71746744 in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-Value
COL11A1 rs1676486	C-AGGG	-0.64	48.3	50.9	44.1	0.525
+	T-(-)	0.54	7.8	7.6	7.6	0.587
COL5A1 rs71746744	T-AGGG	0.62	16.5	15.3	18.7	0.536
	C-(-)	0.79	27.4	26.2	29.5	0.429

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.16: Inferred pseudo-haplotype analysis of *COL11A2* rs1799907 and *COL5A1* rs71746744 in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-Value
COL11A2 rs1799907	T-AGGG	-1.10	15.9	16.4	15.2	0.270
+	A-AGGG	-1.04	48.9	49.8	47.7	0.300
COL5A1 rs71746744	T-(-)	0.17	11.2	10.7	11.9	0.869
	A-(-)	0.73	24.0	23.1	25.3	0.467

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.18: Inferred pseudo-haplotype analysis of *COL5A1* rs12772, *COL6A1* rs35796750 and *COL12A1* rs970547 in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-value
COL5A1 rs12772	T-T-G	0.46	12.6	13.3	10.3	0.645
+	C-T-A	0.11	2.7	2.9	0.0	0.913
COL6A1 rs35796750	T-C-G	0.12	21.2	20.3	22.8	0.901
+	C-C-G	-0.18	31.3	32.5	28.6	0.855
COL12A1 rs970547	C-C-A	0.46	8.1	8.4	8.9	0.986
	T-T-A	0.79	5.5	4.5	8.6	0.432
	C-T-G	0.82	13.5	13.1	16.2	0.412
	T-C-A	-0.02	5.1	5.0	4.6	0.646

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.19: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL12A1* rs970547 in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-value
COL5A1 rs12722	C-A	0.19	11.0	11.3	10.3	0.847
+	T-G	0.90	34.0	33.6	34.5	0.367
COL12A1 rs970547	C-G	0.41	44.6	45.6	43.3	0.678
	T-A	0.88	10.4	9.5	11.9	0.379

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.20: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL6A1* rs35796750 in the Swedish control (CON) and Non-Contact ACL injury (NON).

	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-Value
COL5A1 rs12722	T-T	0.36	17.5	17.9	16.4	0.350
+	T-C	0.37	26.3	25.2	37.2	0.971
COL6A1 rs35796750	C-C	0.46	40.5	40.9	27.9	0.706
	C-T	0.66	15.7	16.0	18.4	0.350

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B4.21: Inferred pseudo-haplotype analysis of *COL6A1* rs35796750 and *COL12A1* rs970547 in the Swedish control (CON) and Non-Contact ACL injury (NON).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (101)	NON (87)	P-value
COL6A1 rs35796750	C-G	-0.47	52.5	53.3	51.6	0.640
+	T-G	0.56	26.0	25.9	26.0	0.573
COL12A1 rs970547	T-A	0.60	8.5	8.0	9.2	0.550
	C-A	0.16	13.0	12.9	13.2	0.875

CON and NON groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Models: **Add** = **Add**itive; **Dom** = **Dom**inant; **Rec** = **Rec**essive.
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in NON.
Hap Frequency = indicates frequency of inferred pseudo-haplotype and is represented by a %.

APPENDIX B5 – COMBINED ACL INJURIES SUPPLEMENTARY TABLES

Table B5.1: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL5A1* rs71746744 polymorphism in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL5A1 rs12722	T-(-)	-0.96	27.8	27.1	28.7	0.222	0.384	0.778
+	T-AGGG	0.10	15.5	15.5	15.6	0.274	0.717	0.942
COL5A1	C-AGGG	-1.47	53.9	54.7	52.6	0.496	0.529	0.631
rs71746744	C-(-)	1.50	2.9	2.7	3.1	0.079	0.433	0.849

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.2: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL11A2* rs1799907polymorphism in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-T-T	-0.84	0.4	0.6	0.0	0.330	0.331	-
rs3753841	T-C-A	-0.84	40.4	41.8	38.8	0.883	0.819	0.245
+	C-T-A	-0.47	14.2	16.2	13.2	0.128	0.277	0.060
COL11A1	T-T-A	0.09	0.3	0.3	0.6	0.962	0.962	-
rs1676486	T-C-T	0.31	18.7	19.0	19.1	0.795	0.619	0.716
+	C-C-A	0.42	12.5	12.2	13.0	0.462	0.653	0.245
COL11A2	C-T-T	1.23	7.0	5.0	8.1	0.943	0.842	0.626
rs1799907	C-C-T	1.40	6.5	4.9	7.3	0.204	0.298	-

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.3: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL11A1* rs1676486in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-C	-0.71	59.3	59.4	59.2	0.956	0.352	0.426
rs3753841	C-T	-0.38	20.9	22.0	19.7	0.260	0.168	0.964
+	T-T	-0.24	0.6	0.8	0.4	0.573	0.686	-
COL11A1	C-C	0.96	19.0	17.8	20.7	0.220	0.472	0.070
rs1676486								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.4: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL11A2* rs1799907 in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-A	-1.47	26.6	27.7	25.8	0.556	0.794	0.369
rs3753841	T-A	-0.97	40.5	40.4	40.2	0.845	0.717	0.625
+	T-T	0.25	19.6	20.2	19.4	0.986	0.789	0.635
COL11A2	C-T	1.60	13.3	11.7	14.6	0.274	0.214	0.967
rs1799907								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.5: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486 and *COL11A2* rs1799907 in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-T	-1.74	14.2	15.6	12.8	0.132	0.287	0.056
rs1676486	T-A	-0.12	7.3	7.2	7.4	0.898	0.925	0.408
+	C-A	-0.10	52.9	52.5	53.3	0.749	0.926	0.681
COL11A2	C-T	1.70	25.6	24.7	26.6	0.397	0.475	0.487
rs1799907								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.6: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 and *COL5A1* rs71746744, in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
	T-T-T-AGGG	-1.14	0.3	0.4	0.0	0.448	0.448	-
COL11A1	T-C-A-AGGG	-1.07	28.1	28.3	27.9	0.653	0.683	0.738
rs3753841	C-T-A-(-)	-1.01	2.8	3.8	1.4	0.252	0.259	-
+	T-C-T-AGGG	-0.42	12.1	13.1	11.5	0.646	0.757	0.554
COL11A1	C-C-A-AGGG	-0.31	9.4	9.6	8.8	0.654	0.858	-
rs1676486	C-T-A-AGGG	-0.29	11.3	11.7	11.2	0.172	0.277	0.166
+	T-T-A-AGGG	-0.15	0.2	0.2	0.4	0.760	0.759	-
COL11A2	T-C-A-(-)	-0.03	12.1	12.1	11.9	0.683	0.864	0.434
rs1799907	C-T-T-AGGG	0.57	4.4	5.0	3.4	0.525	0.651	-
+	C-C-T-AGGG	1.22	4.0	2.4	5.2	0.167	0.206	-
COL5A1	C-C-T-(-)	1.25	2.5	3.3	2.0	0.598	0.584	-
rs71746744	T-C-T-(-)	1.39	7.0	5.9	8.0	0.240	0.214	-
	C-C-A-(-)	1.49	3.2	2.5	4.6	0.868	0.365	0.315
	C-T-T-(-)	1.58	2.4	1.5	3.8	0.231	0.222	-

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.7: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL5A1* rs71746744 in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-C-AGGG	-1.06	40.2	41.0	39.1	0.501	0.523	0.632
rs3753841	T-T-AGGG	-0.63	0.5	0.6	0.4	0.778	0.968	-
+	C-T-AGGG	-0.33	15.7	16.8	14.3	0.177	0.152	0.668
COL11A1	C-T-(-)	0.30	5.3	5.3	5.3	0.951	0.993	-
rs1676486	C-C-AGGG	0.30	13.4	12.3	14.7	0.307	0.547	0.108
+	T-C-(-)	0.82	19.1	18.3	20.2	0.332	0.464	0.314
COL5A1	C-C-(-)	1.77	5.7	5.5	6.0	0.382	0.345	-
rs71746744								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.8: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1799907 and *COL5A1* rs71746744 in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-A-AGGG	-2.28	20.8	21.4	20.3	0.504	0.936	0.117
rs3753841	T-A-AGGG	-1.13	28.2	28.1	27.9	0.608	0.540	0.870
+	T-T-AGGG	0.17	12.5	13.6	11.5	0.519	0.610	0.520
COL11A1	T-A-(-)	0.40	12.2	12.1	12.3	0.678	0.812	0.496
rs1799907	C-T-(-)	1.78	4.8	4.1	5.7	0.199	0.179	-
+	C-A-(-)	-0.21	6.0	6.5	5.5	0.831	0.930	-
COL5A1	C-T-AGGG	0.60	8.3	7.5	8.8	0.546	0.439	0.799
rs71746744	T-T-(-)	1.02	7.1	6.6	8.0	0.307	0.273	-

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.9: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486, *COL11A1* rs1799907 and *COL5A1* rs71746744 in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-A-AGGG	-1.59	11.5	11.6	11.6	0.206	0.335	0.160
rs1676486	C-A-AGGG	-1.31	37.5	37.8	36.7	0.853	0.771	0.497
+	C-A-(-)	0.18	15.4	14.8	16.5	0.480	0.393	0.912
COL11A1	C-T-(-)	0.42	9.4	8.9	9.9	0.229	0.294	0.313
rs1799907	C-T-AGGG	0.77	16.1	15.6	16.8	0.799	0.931	0.614
+	T-A-(-)	-1.28	2.9	3.9	1.3	0.199	0.205	-
COL5A1	T-T-AGGG	-0.83	4.8	5.7	3.5	0.409	0.576	-
rs71746744	T-T-(-)	1.07	2.4	1.6	3.8	0.285	0.274	-

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.10: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL5A1* rs71746744 in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-AGGG	-1.14	40.8	41.7	39.5	0.412	0.328	0.717
rs3753841	C-(-)	-1.12	10.8	10.5	11.3	0.543	0.438	0.683
+	C-AGGG	-0.63	29.1	29.0	29.1	0.817	0.783	0.956
COL5A1	T-(-)	0.44	19.3	18.8	20.2	0.388	0.497	0.402
rs71746744								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.11: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486 and *COL5A1* rs71746744 in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-AGGG	-0.34	16.3	17.4	14.9	0.174	0.191	0.438
rs1676486	C-AGGG	-0.29	53.6	53.3	53.6	0.988	0.805	0.821
+	T-(-)	0.39	5.3	5.5	5.2	0.830	0.870	-
COL5A1	C-(-)	1.67	24.8	23.8	26.3	0.222	0.172	0.740
rs71746744								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.12: Inferred pseudo-haplotype analysis of *COL11A2* rs1799907 and *COL5A1* rs71746744 in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL11A2	A-AGGG	-1.44	48.9	49.4	48.1	0.360	0.671	0.280
rs1799907	A-(-)	-0.68	18.3	18.8	17.9	0.829	0.565	0.458
+	T-AGGG	0.54	20.1	21.3	20.4	0.890	0.992	0.695
COL5A1	T-(-)	0.94	11.8	10.5	13.6	0.155	0.190	0.356
rs71746744								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.13: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL6A1* rs35796750 in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL5A1 rs12772 + *COL6A1* rs35796750 + *COL12A1* rs970547	C-C-A	-0.98	6.0	7.2	4.3	0.325	0.311	-
	T-T-G	-0.48	14.1	15.0	12.5	0.628	0.950	0.067
	T-C-A	-0.36	6.9	7.5	5.3	0.715	0.597	-
	C-T-A	-0.20	5.8	5.5	6.3	0.842	0.700	-
	T-C-G	-0.18	19.3	18.8	20.0	0.855	0.960	0.452
	C-T-G	-0.10	17.2	18.1	15.5	0.923	0.849	0.433
	T-T-A	0.77	2.6	1.8	4.7	0.439	0.632	-
	C-C-G	1.10	28.1	26.0	31.3	0.271	0.308	0.403

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.14: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL12A1* rs970547 in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL5A1 rs12722 + *COL12A1* rs970547	C-G	-0.98	45.3	44.2	46.9	0.422	0.506	0.502
	T-G	-0.53	33.3	33.9	32.4	0.667	0.942	0.434
	C-A	-0.43	11.8	12.7	10.1	0.924	0.969	0.616
	T-A	0.51	9.6	9.3	10.6	0.449	0.395	0.936

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.15: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL6A1* rs35796750 in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL5A1 rs12722 + *COL6A1* rs35796750	C-T	-1.95	22.7	23.5	21.4	0.891	0.633	0.168
	T-T	-1.32	16.7	16.5	17.3	0.922	0.784	0.180
	C-C	-1.19	34.3	33.4	36.0	0.578	0.343	0.838
	T-C	-0.95	26.2	26.6	25.2	0.685	0.944	0.236

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B5.16: Inferred pseudo-haplotype analysis of *COL6A1* rs35796750 and *COL12A1* rs970547 in the control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (662)	ACL (346)	P-value[a]	P-value[b]	P-value[c]
COL6A1	C-A	-0.97	13.0	14.8	9.7	0.330	0.171	0.272
rs35796750	T-G	-0.39	31.1	33.0	28.0	0.696	0.671	0.125
+	T-A	0.28	8.4	7.1	11.0	0.744	0.899	***0.030***
COL12A1	C-G	0.89	47.5	45.1	51.3	0.371	0.409	0.483
rs970547								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

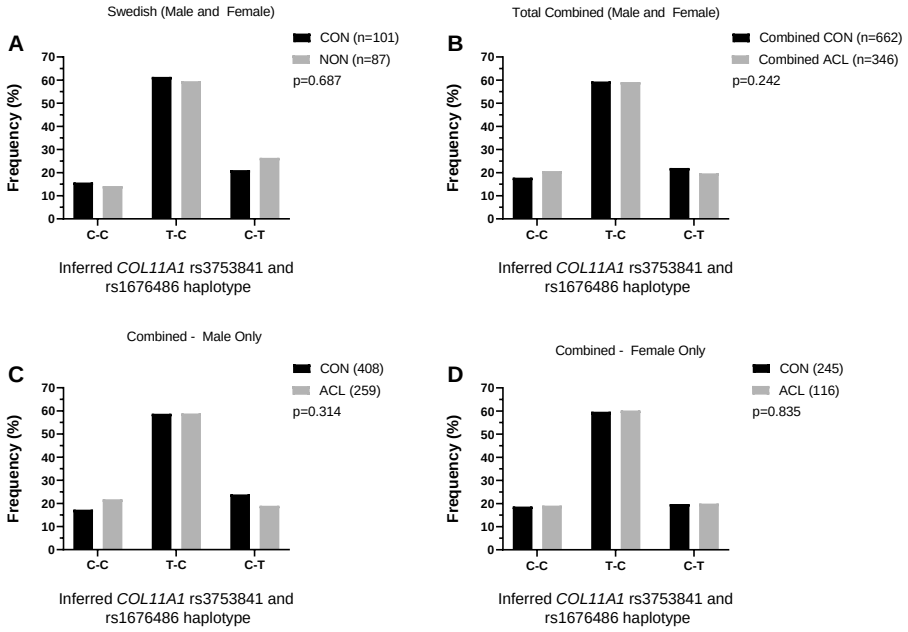


Figure B5.9 Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and rs1676486 (C/T) polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

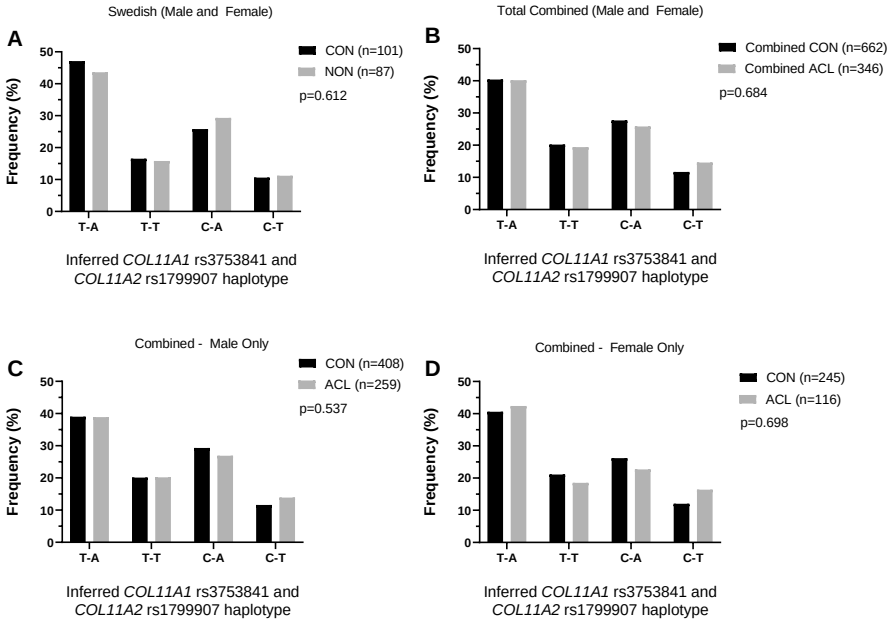


Figure B5.10 Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and *COL11A2* rs1799907 (A/T) polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

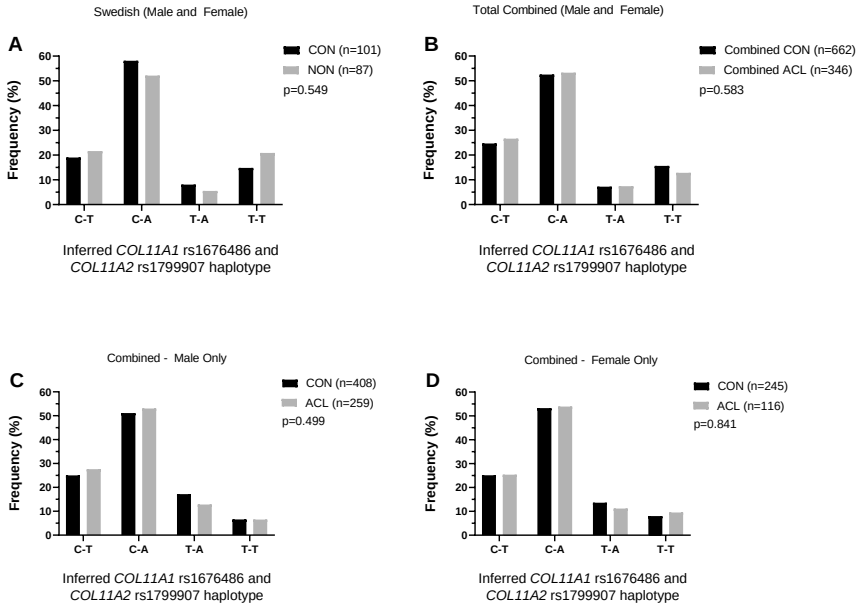


Figure B5.11 Inferred haplotype frequency distributions constructed from *COL11A1* rs1676484 (T/C) and *COL11A2* rs1799907 (A/T) polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend.

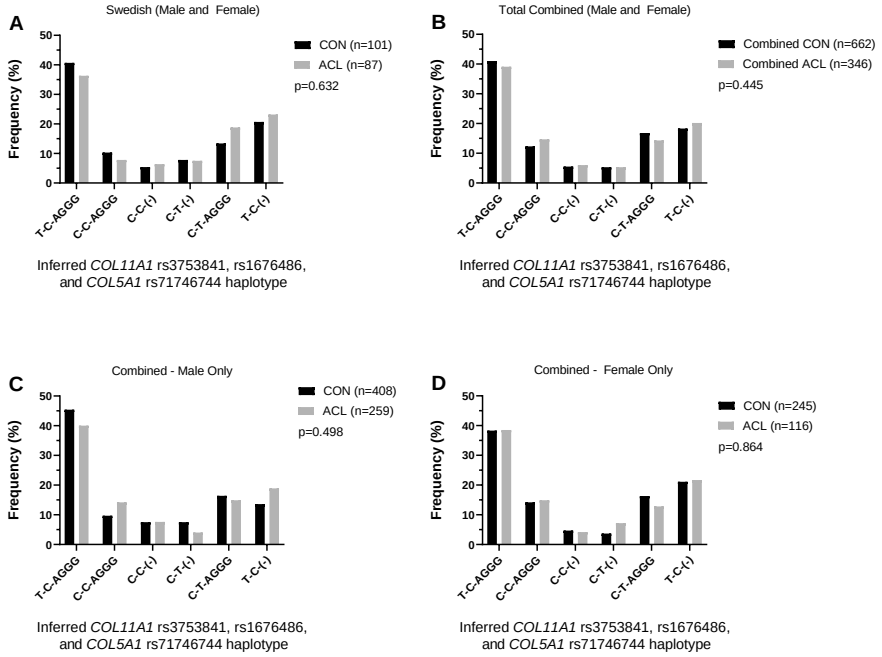


Figure B5.12 Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C), rs1676486 (C/T) and *COL5A1* rs10628678 (AGGG/-) polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

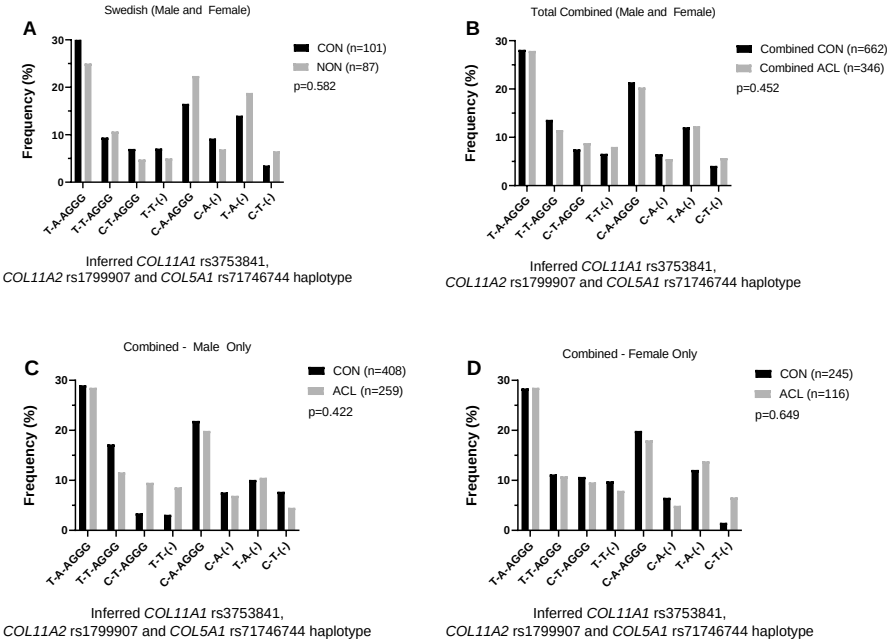


Figure B5.13 Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C), *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

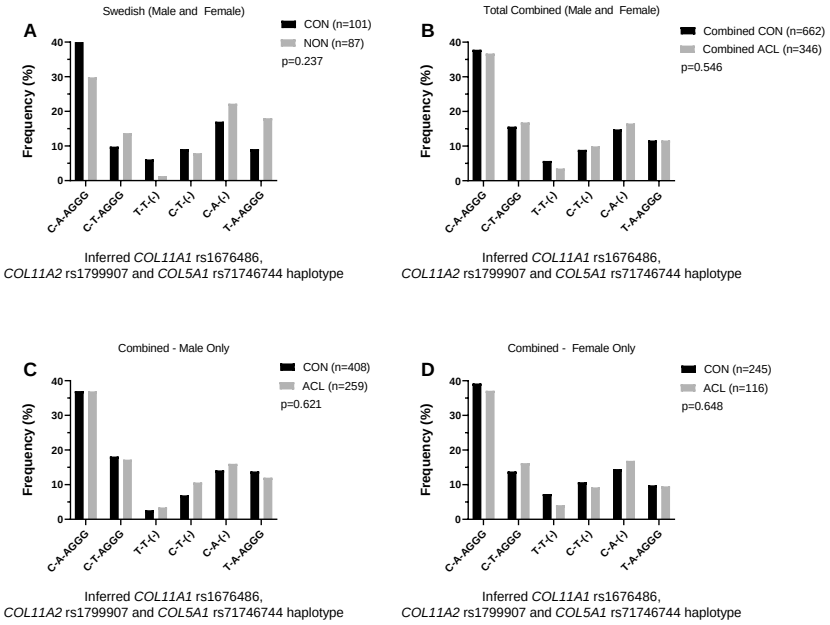


Figure B5.14 Inferred haplotype frequency distributions constructed from *COL11A1* rs1676486 (C/T), *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

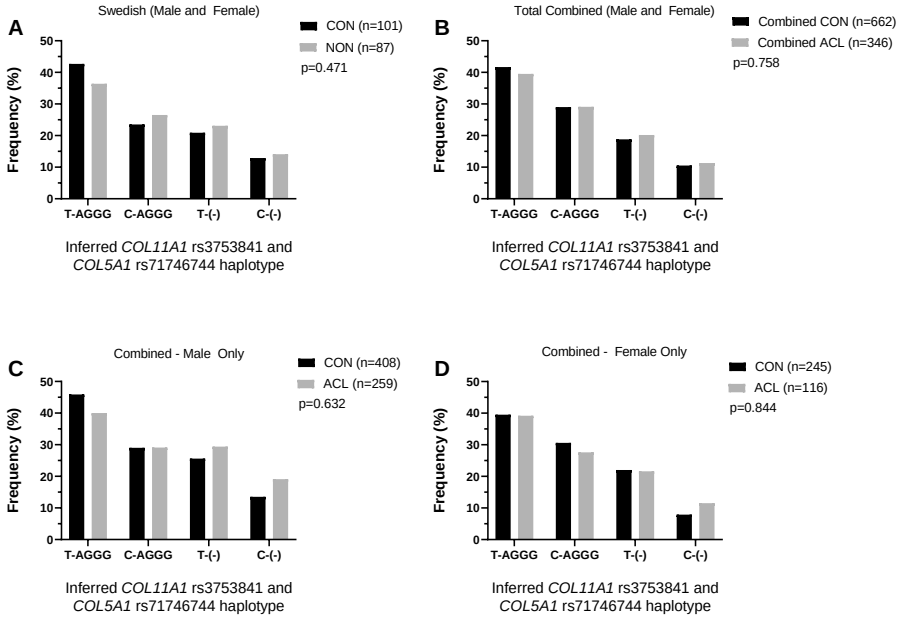


Figure B5.15. Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and *COL5A1* rs10628678 (AGGG/-) polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

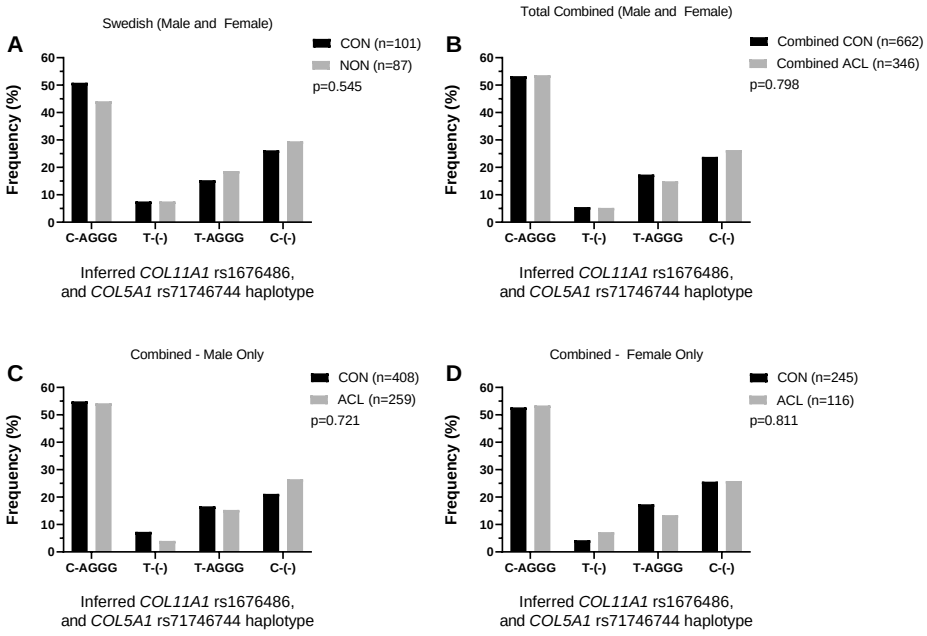


Figure B5.16 Inferred haplotype frequency distributions constructed from *COL11A1* rs1676486 (C/T) and *COL5A1* rs10628678 (AGGG/-) polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

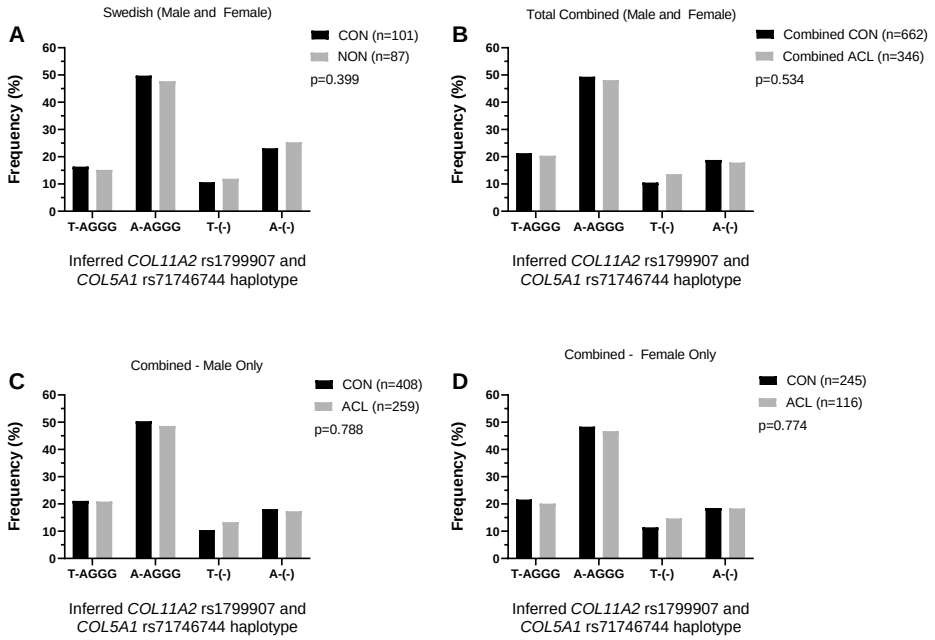


Figure B5.17 Inferred haplotype frequency distributions constructed from *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) polymorphisms of the Swedish (Male and Female) cohort (A), combined ACL (B), male only (C) and female only (D), for CON (black bar) and ACL groups (grey bar). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

APPENDIX B6 – MALE ONLY COMBINED ACL COHORT INJURIES SUPPLEMENTARY TABLES

Table B6.1: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL5A1* rs71746744 polymorphism in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL5A1 rs12722	T-(-)	-0.55	28.0	28.0	27.9	0.582	0.914	0.257
+	C-AGGG	-0.28	54.3	54.7	53.7	0.779	0.861	0.784
COL5A1	T-AGGG	0.52	15.7	15.5	16.0	0.604	0.365	0.411
rs71746744	C-(-)	1.50	2.0	1.8	2.4	0.134	0.276	-

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.2: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL11A2* rs1799907 polymorphism in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-T-T	-0.97	6.5	6.7	6.3	0.330	0.432	-
rs3753841	C-T-A	-0.91	14.8	17.3	12.7	0.361	0.553	0.219
+	T-C-T	-0.65	19.8	20.0	19.9	0.514	0.375	0.835
COL11A1	C-C-T	0.13	6.5	5.0	7.6	0.894	0.971	-
rs1676486	C-C-A	0.86	13.2	12.3	14.2	0.388	0.513	0.304
+	T-C-A	0.90	39.1	38.8	39.0	0.370	0.357	0.600
COL11A2								
rs1799907								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.3: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL11A1* rs1676486 in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-T	-1.20	21.3	23.9	19.0	0.228	0.179	0.770
rs3753841	T-C	0.35	58.8	58.8	58.9	0.724	0.800	0.473
+	C-C	0.75	19.7	17.3	21.8	0.456	0.900	0.052
COL11A1	T-T						-	-
rs1676486		-	-	-	-	-		

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.4: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL11A2* rs1799907 in the Male Only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-T	-0.58	20.0	20.1	20.2	0.562	0.436	0.864
rs3753841	C-T	-0.55	12.9	11.6	13.9	0.585	0.783	0.271
+	C-A	-0.01	27.9	29.3	26.9	0.992	0.615	0.378
COL11A2	T-A	0.81	39.1	39.0	38.9	0.417	0.408	0.629
rs1799907								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.5: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486 and *COL11A2* rs1799907 in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-T	-0.85	6.6	6.6	6.5	0.395	0.518	-
rs1676486	T-A	-0.84	14.9	17.2	12.8	0.402	0.617	0.216
+	C-T	-0.44	26.4	25.0	27.6	0.660	0.463	0.781
COL11A2	C-A	1.33	52.1	51.1	53.0	0.185	0.382	0.190
rs1799907								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.6: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 and *COL5A1* rs71746744, in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-T-A-(-)	-1.27	2.3	4.0	1.2	0.205	0.208	-
rs3753841	T-C-T-AGGG	-1.05	13.9	17.0	11.1	0.294	0.281	0.713
+	C-T-T-(-)	-0.97	3.2	3.5	2.6	0.333	0.342	-
COL11A1	C-C-T-(-)	-0.71	2.9	4.1	1.8	0.475	0.484	-
rs1676486	C-T-A-AGGG	-0.58	12.8	13.7	11.3	0.564	0.730	0.382
+	C-T-T-AGGG	-0.51	3.0	2.6	3.9	0.613	0.808	-
COL11A2	T-C-T-(-)	0.43	6.0	3.1	8.7	0.665	0.813	-
rs1799907	C-C-A-(-)	0.53	4.4	3.5	6.3	0.598	0.739	-
+	T-C-A-(-)	0.57	10.5	10.2	9.9	0.565	0.729	0.227
COL5A1	T-C-A-AGGG	0.63	28.5	28.8	29.3	0.530	0.623	0.412
rs71746744	C-C-A-AGGG	0.71	8.8	8.5	8.0	0.475	0.560	-
	C-C-T-AGGG	0.82	3.5	1.0	5.7	0.411	0.477	-

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).

[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.7: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL5A1* rs71746744 in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-T-(-)	-1.35	5.5	7.5	4.0	0.176	0.189	-
rs3753841	C-T-AGGG	-0.73	15.9	16.4	14.9	0.465	0.439	0.792
+	T-C-AGGG	-0.10	42.3	45.4	40.0	0.923	0.992	0.876
COL11A1	C-C-(-)	-0.07	7.3	7.5	7.6	0.942	0.991	-
rs1676486	T-C-(-)	0.63	16.6	13.6	18.9	0.529	0.879	0.160
+	C-C-AGGG	1.02	12.3	9.7	14.2	0.309	0.634	0.066
COL5A1								
rs71746744								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.8: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1799907 and *COL5A1* rs71746744 in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-T-(-)	-1.15	6.2	7.7	4.5	0.250	0.270	-
rs3753841	T-T-AGGG	-0.89	14.2	17.2	11.6	0.373	0.380	0.678
+	C-A-(-)	-0.50	6.4	7.6	6.9	0.619	0.722	-
COL11A1	C-A-AGGG	0.24	21.7	21.9	19.9	0.807	0.365	0.220
rs1799907	C-T-AGGG	0.31	6.5	3.4	9.5	0.756	0.652	-
+	T-A-AGGG	0.39	28.3	29.0	28.5	0.696	0.982	0.431
COL5A1	T-T-(-)	0.43	5.9	3.1	8.6	0.668	0.737	-
rs71746744							0.675	0.246
	T-A-(-)	0.72	10.8	10.1	10.5	0.474		

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.9: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486, *COL11A1* rs1799907 and *COL5A1* rs71746744 in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-A-(-)	-1.17	2.4	3.8	1.0	0.243	0.247	-
rs1676486	T-T-(-)	-0.92	3.1	3.7	3.0	0.356	0.365	-
+	T-A-AGGG	-0.51	12.8	13.8	12.0	0.609	0.573	0.381
COL11A1	C-T-AGGG	-0.43	17.5	18.1	17.2	0.668	0.791	0.910
rs1799907	T-T-AGGG	-0.41	3.2	2.6	3.4	0.684	0.900	-
+	C-T-(-)	-0.04	8.9	6.9	10.6	0.969	0.936	-
COL5A1	C-A-(-)	0.69	14.9	14.1	16.0	0.493	0.544	0.608
rs71746744	C-A-AGGG	0.88	37.2	37.0	36.9	0.378	0.493	0.420

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.10: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL5A1* rs71746744 in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-(-)	-0.98	12.6	15.0	11.4	0.329	0.503	-
rs3753841	T-AGGG	-0.15	42.5	45.9	40.0	0.881	0.836	0.970
+	T-(-)	0.28	28.2	25.6	29.4	0.779	0.825	0.801
COL5A1	C-AGGG	0.66	16.7	13.5	19.1	0.507	0.841	0.165
rs71746744								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.11: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486 and *COL5A1* rs71746744 in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-(-)	-1.27	5.5	7.3	4.0	0.204	0.219	-
rs1676486	T-AGGG	-0.63	16.0	16.6	15.3	0.530	0.522	0.778
+	C-(-)	0.45	23.8	21.2	26.5	0.650	0.766	0.581
COL5A1	C-AGGG	0.55	54.7	54.9	54.2	0.585	0.484	0.791
rs71746744								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.12: Inferred pseudo-haplotype analysis of *COL11A2* rs1799907 and *COL5A1* rs71746744 in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL11A2	T-AGGG	-0.61	21.0	21.1	20.9	0.544	0.602	0.635
rs1799907	T-(-)	-0.46	11.8	10.4	13.3	0.646	0.715	-
+	A-(-)	0.24	17.5	18.1	17.3	0.810	0.822	0.883
COL5A1	A-AGGG	0.56	49.7	50.4	48.6	0.578	0.606	0.676
rs71746744								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.13: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL6A1* rs35796750 in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL5A1 rs12772 + *COL6A1* rs35796750 + *COL12A1* rs970547	C-T-G	-1.59	16.2	19.4	11.5	0.113	0.147	0.259
	T-C-A	-0.72	5.4	7.1	3.4	0.473	0.593	-
	T-T-G	-0.53	13.0	15.0	11.6	0.597	0.711	0.422
	C-C-A	-0.25	4.5	5.2	3.2	0.801	0.777	-
	T-C-G	-0.17	21.2	19.9	21.8	0.867	0.962	0.512
	C-T-A	0.76	6.5	5.6	8.0	0.449	0.617	-
	C-C-G	1.09	29.2	26.3	33.5	0.275	0.485	0.192
	T-T-A	1.66	4.1	1.7	7.0	0.097	0.097	-

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.14: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL12A1* rs970547 in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL5A1 rs12722	T-G	-0.44	34.2	34.9	33.4	0.663	0.576	0.956
+	C-G	0.00	45.4	45.6	45.2	0.998	0.650	0.636
COL12A1	C-A	0.31	11.0	10.8	11.1	0.759	0.831	0.696
rs970547	T-A	0.45	9.4	8.7	10.3	0.656	0.581	-

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.15: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL6A1* rs35796750 in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL5A1 rs12722	C-T	-0.96	22.5	24.9	18.7	0.336	0.480	0.299
+	T-C	-0.43	26.3	27.3	24.6	0.664	0.982	0.215
COL6A1	T-T	0.26	17.3	16.3	19.2	0.797	0.678	0.666
rs35796750	C-C	1.01	33.9	31.5	37.6	0.311	0.481	0.233

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B6.16: Inferred pseudo-haplotype analysis of *COL6A1* rs35796750 and *COL12A1* rs970547 in the Male only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (408)	ACL (259)	P-value[a]	P-value[b]	P-value[c]
COL6A1	T-G	-1.58	28.9	33.7	23.0	0.115	0.583	***0.010***
rs35796750	C-A	-0.68	9.8	12.0	6.6	0.493	0.562	0.523
+	C-G	0.89	50.6	46.7	55.4	0.373	0.260	0.665
COL12A1	T-A	1.44	10.7	7.6	15.0	0.150	0.282	0.068
rs970547								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

APPENDIX B7 – FEMALE ONLY COMBINED ACL COHORT INJURIES SUPPLEMENTARY TABLES

Table B7.1: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL5A1* rs71746744 polymorphism in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL5A1 rs12722	T-AGGG	-1.51	15.2	16.6	12.1	0.131	0.201	0.196
+	C-AGGG	0.16	54.1	54.1	54.5	0.870	0.307	0.525
COL5A1	T-(-)	0.56	26.4	25.8	27.7	0.579	0.259	0.548
rs71746744	C-(-)	1.24	4.3	3.4	5.7	0.213	0.423	-

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.2: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL11A2* rs1799907 polymorphism in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-T-A	-0.40	12.0	13.2	10.5	0.686	0.825	-
rs3753841	C-C-A	-0.27	12.3	12.5	12.3	0.788	0.740	-
+	T-C-T	0.10	18.6	19.2	18.6	0.921	0.762	0.715
COL11A1	C-T-T	0.12	7.9	6.6	9.5	0.905	0.787	-
rs1676486	C-C-T	0.18	6.6	6.2	6.8	0.854	0.939	-
+	T-C-A	0.28	41.3	40.5	41.6	0.781	0.740	0.395
COL11A2								
rs1799907								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.3: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL11A1* rs1676486 in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-T	-0.39	1.3	1.7	0.7	0.695	0.825	-
rs3753841	C-T	-0.21	19.9	19.8	20.0	0.836	0.624	0.516
+	C-C	-0.11	18.9	18.7	19.1	0.913	0.925	0.930
COL11A1	T-C	0.35	59.9	59.7	60.2	0.729	0.684	0.428
rs1676486								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.4: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL11A2* rs1799907 in the Female Only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-A	-0.55	24.5	26.2	22.7	0.583	0.504	0.993
rs3753841	T-T	-0.21	19.7	21.1	18.5	0.830	0.971	0.635
+	C-T	0.37	14.1	12.0	16.4	0.714	0.722	0.859
COL11A2	T-A	0.37	41.7	40.6	42.4	0.710	0.684	0.276
rs1799907								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.5: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486 and *COL11A2* rs1799907 in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-A	-0.34	12.6	13.6	11.2	0.733	0.879	-
rs1676486	T-T	-0.12	8.6	8.0	9.5	0.907	0.919	-
+	C-A	0.12	53.6	53.3	53.9	0.902	0.533	0.489
COL11A2	C-T	0.16	25.2	25.1	25.4	0.874	0.896	0.546
rs1799907								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.6: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486, *COL11A2* rs1799907 and *COL5A1* rs71746744, in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-T-A-AGGG	-0.75	8.8	9.8	6.4	0.453	0.563	-
rs3753841	C-C-A-(-)	-0.68	1.9	2.5	0.0	0.495	0.496	-
+	C-T-T-AGGG	-0.56	6.2	6.2	6.1	0.578	0.594	-
COL11A1	T-C-A-AGGG	-0.21	28.3	28.5	26.5	0.830	0.944	0.732
rs1676486	C-C-A-AGGG	-0.13	10.4	10.2	11.9	0.894	0.674	-
+	T-C-T-AGGG	0.03	10.2	10.1	11.0	0.977	0.996	0.947
COL11A2	C-C-T-(-)	0.10	2.6	2.2	3.1	0.917	0.915	-
rs1799907	C-C-T-AGGG	0.16	4.0	3.9	4.2	0.874	0.861	-
+	T-C-T-(-)	0.38	8.2	8.6	7.8	0.702	0.631	-
COL5A1	T-C-A-(-)	0.63	13.0	12.2	14.8	0.527	0.575	0.628
rs71746744	C-T-A-(-)	0.65	3.2	3.0	4.7	0.514	0.511	-
	C-T-T-(-)	1.07	1.8	0.8	2.8	0.286	0.285	-

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).

[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.7: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL5A1* rs71746744 in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-T-AGGG	-0.85	15.1	16.3	12.8	0.395	0.321	0.960
rs3753841	T-T-AGGG	-0.14	1.0	1.3	0.6	0.890	0.884	-
+	C-C-AGGG	-0.13	14.4	14.2	14.9	0.896	0.804	0.760
COL11A1	C-C-(-)	-0.11	4.6	4.7	4.2	0.911	0.914	-
rs1676486	T-C-AGGG	-0.09	38.4	38.3	38.5	0.926	0.926	0.952
+	T-C-(-)	0.62	21.3	21.1	21.7	0.535	0.451	0.970
COL5A1	C-T-(-)	1.13	4.9	3.7	7.2	0.259	0.256	-
rs71746744								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.8: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1799907 and *COL5A1* rs71746744 in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-A-AGGG	-0.65	19.4	19.9	18.0	0.516	0.632	0.486
rs3753841	C-T-AGGG	-0.37	10.1	10.7	9.6	0.714	0.763	-
+	T-T-AGGG	-0.29	11.1	11.2	10.8	0.773	0.701	0.958
COL11A1	C-A-(-)	0.02	5.2	6.5	4.9	0.985	0.916	-
rs1799907	T-A-AGGG	0.05	28.4	28.4	28.5	0.964	0.992	0.914
+	T-T-(-)	0.14	8.6	9.8	7.9	0.890	0.810	-
COL5A1	T-A-(-)	0.51	13.2	12.1	13.8	0.611	0.646	0.711
rs71746744	C-T-(-)	1.32	4.0	1.5	6.6	0.186	0.186	-

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.9: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486, *COL11A1* rs1799907 and *COL5A1* rs71746744 in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-T-AGGG	-0.80	6.6	7.3	4.1	0.422	0.460	-
rs1676486	T-A-AGGG	-0.42	9.4	9.8	9.5	0.671	0.820	-
+	C-A-AGGG	-0.29	38.6	39.2	37.1	0.773	0.854	0.758
COL11A1	C-T-AGGG	0.05	14.4	13.8	16.2	0.957	0.725	0.979
rs1799907	T-A-(-)	0.18	3.2	3.5	1.7	0.854	0.852	-
+	C-T-(-)	0.37	10.6	10.7	9.2	0.709	0.846	-
COL5A1	C-A-(-)	0.48	15.1	14.5	16.9	0.631	0.568	0.385
rs71746744	T-T-(-)	1.02	2.2	1.2	5.3	0.306	0.305	-

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.10: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL5A1* rs71746744 in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL11A1	C-AGGG	-0.74	29.5	30.6	27.6	0.462	0.366	0.869
rs3753841	T-AGGG	-0.13	39.4	39.5	39.2	0.895	0.721	0.859
+	T-(-)	0.46	21.8	22.0	21.6	0.649	0.521	0.891
COL5A1	C-(-)	0.86	9.2	7.9	11.5	0.392	0.456	-
rs71746744								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.11: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486 and *COL5A1* rs71746744 in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL11A1	T-AGGG	-0.82	16.0	17.4	13.4	0.414	0.405	0.744
rs1676486	C-AGGG	-0.20	52.9	52.7	53.4	0.839	0.790	0.928
+	C-(-)	0.55	25.8	25.6	25.9	0.582	0.377	0.740
COL5A1	T-(-)	0.83	5.3	4.3	7.2	0.405	0.402	-
rs71746744								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.12: Inferred pseudo-haplotype analysis of *COL11A2* rs1799907 and *COL5A1* rs71746744 in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL11A2	T-AGGG	-0.48	21.1	21.6	20.1	0.632	0.428	0.716
rs1799907	A-AGGG	-0.42	47.8	48.4	46.7	0.678	0.683	0.769
+	A-(-)	0.43	18.5	18.5	18.4	0.669	0.545	0.893
COL5A1	T-(-)	0.75	12.6	11.4	14.7	0.454	0.683	0.232
rs71746744								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.13: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL6A1* rs35796750 in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL5A1 rs12772 + *COL6A1* rs35796750 + *COL12A1* rs970547	T-T-A	-0.76	1.8	2.1	3.0	0.445	0.433	-
	T-T-G	-0.51	14.2	14.8	10.9	0.607	0.818	-
	C-T-A	-0.51	4.1	4.8	0.0	0.610	0.564	-
	T-C-A	-0.38	8.0	8.0	6.9	0.702	0.432	-
	T-C-G	-0.14	17.6	17.6	19.0	0.892	0.949	0.816
	C-C-A	-0.13	8.9	10.1	9.2	0.897	0.752	-
	C-C-G	0.51	26.4	25.8	24.8	0.610	0.410	0.842
	C-T-G	0.73	19.0	16.7	26.2	0.463	0.372	0.979

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.14: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL12A1* rs970547 in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL5A1 rs12722	T-A	-0.51	9.9	10.1	9.4	0.613	0.455	-
+	T-G	-0.40	31.7	32.4	30.5	0.688	0.879	0.279
COL12A1	C-A	-0.29	12.9	14.9	9.7	0.776	0.564	-
rs970547	C-G	0.84	45.4	42.6	50.4	0.402	0.255	0.809

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.

Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.15: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL6A1* rs35796750 in the Feamle only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL5A1 rs12722	T-T	-0.57	15.9	16.6	14.3	0.567	0.690	0.410
+	T-C	-0.42	25.7	25.9	25.6	0.673	0.914	0.447
COL6A1	C-T	0.30	35.0	35.9	33.4	0.764	0.283	0.416
rs35796750	C-C	0.62	23.4	21.7	26.7	0.537	0.303	0.596

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B7.16: Inferred pseudo-haplotype analysis of *COL6A1* rs35796750 and *COL12A1* rs970547 in the Female only control (CON) and ACL-Contact ACL injury (ACL).

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (245)	ACL (116)	P-value[a]	P-value[b]	P-value[c]
COL6A1	T-A	-0.83	5.6	6.5	4.0	0.407	0.371	-
rs35796750	C-A	-0.38	17.2	18.5	15.0	0.705	0.229	***0.026***
+	T-G	0.25	33.4	31.7	36.3	0.805	0.731	0.999
COL12A1	C-G	0.35	43.8	43.3	44.6	0.730	0.934	0.486
rs970547								

CON and ACL groups are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
[a]; dominant model, [b]; recessive model, [c]; additive model
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in ACL.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Appendix B8: Mixed ancestry ACL Chapter 6 Supplementary Table and Figures

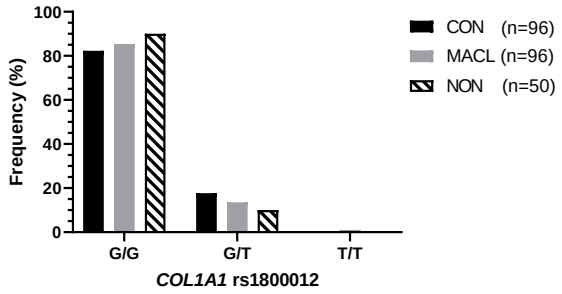


Figure B8.1 Genotypes frequencies (%) are indicated for the COL1A1 rs1800012 (G/T) polymorphism. Black bars indicate CON group, grey bars indicate MACL group and striped bars indicate the non-contact ACL injury (NON) sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

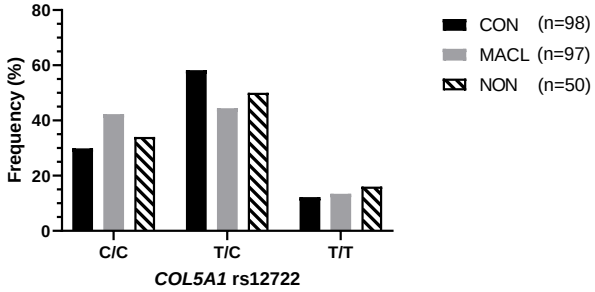


Figure B8.2 Genotypes frequencies (%) are indicated for the COL5A1 rs12722 (T/C) polymorphism. Black bars indicate CON group, grey bars indicated MACL group and striped bars indicate the non-contact ACL injury (NON) sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

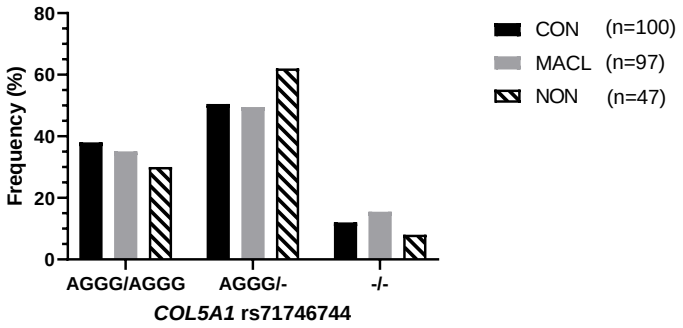


Figure B8.3 Genotypes frequencies (%) are indicated for the COL5A1 rs10628678 (AGGG/-) polymorphism. Black bars indicate CON group, grey bars indicated ACL group and striped bars indicate NON sub-group. Significance is indicated with an *. Significance is accepted at $p<0.05$.

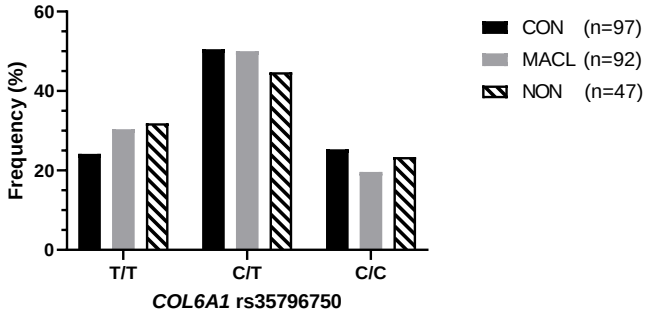


Figure B8.4 Genotypes frequencies (%) are indicated for the COL6A1 rs35796750 (C/T) polymorphism. Black bars indicate CON group, grey bars indicated MACL group and striped bars indicate the non-contact ACL injury (NON) sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

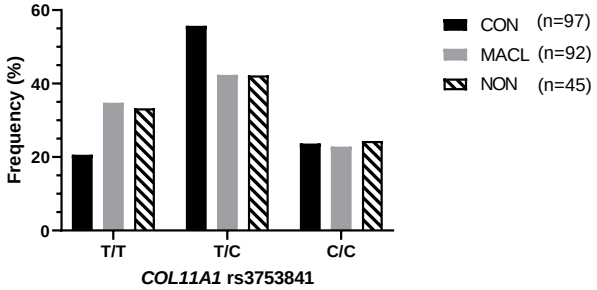


Figure B8.5 Genotypes frequencies (%) are indicated for the COL11A1 rs3753841 (T/C) polymorphism. Black bars indicate CON group and grey bars indicated MACL group and striped bars indicate the non-contact ACL injury (NON) sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

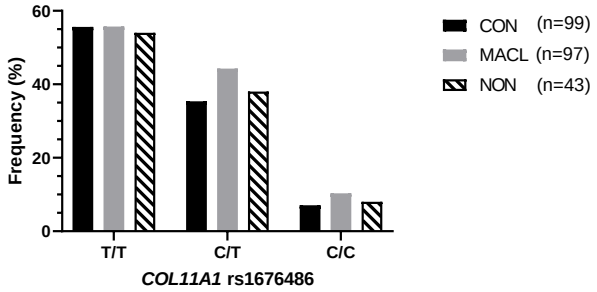


Figure B8.6 Genotypes frequencies (%) are indicated for the COL11A1 rs1676486 (C/T) polymorphism. Black bars indicate CON group and grey bars indicated MACL group and striped bars indicate the non-contact ACL injury (NON) sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

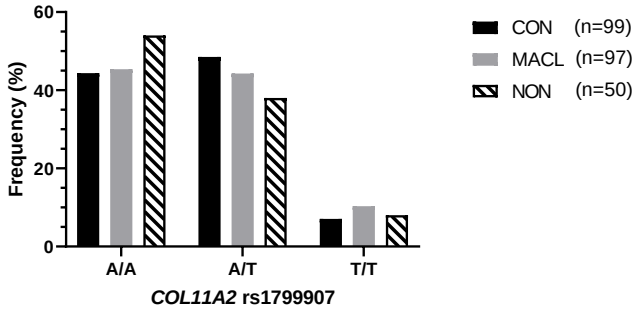


Figure B8.7 Genotypes frequencies (%) are indicated for the COL11A2 rs1799907 (A/T) polymorphism. Black bars indicate CON group and grey bars indicated MACL group and striped bars indicate the non-contact ACL injury (NON) sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

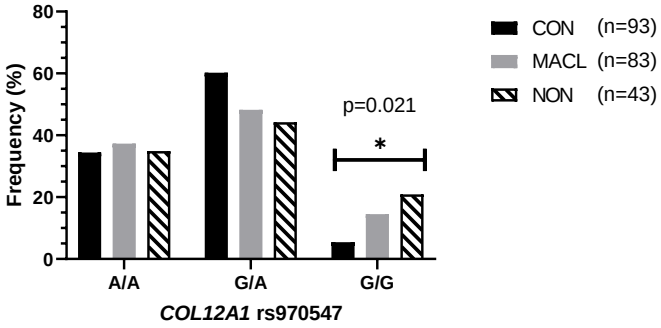


Figure B8.8 Genotypes frequencies (%) are indicated for the COL12A1 rs970547 (G/A) polymorphism. Black bars indicate CON group and grey bars indicated MACL group and striped bars indicate the non-contact ACL injury (NON) sub-group. Significance is accepted at $p<0.05$ and is indicated with an asterisk (*).

Table B8.1: Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL5A1* rs10628678 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-value
COL5A1 rs12722 + *COL5A1* rs10628678	C-AGGG	-2.09	32.4	35.0	29.7	0.231	-0.78	35.0	35.0	0.438
	C-(-)	-0.47	6.0	6.0	5.9	0.635	-0.02	6.1	6.0	0.983
	T-(AGGG)	0.83	29.1	28.0	30.1	0.407	0.11	27.4	26.	0.912
	T-(-)	0.97	32.6	31.0	34.3	0.334	1.06	31.6	33.0	0.288

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B8.2: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL11A2* rs1799907 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-value
COL11A1 rs3753841	T-T-A	-1.24	3.3	4.9	2.3	0.215	-1.42	3.0	0.0	0.157
+	T-T-T	-1.09	2.5	3.8	1.2	0.276	-1.28	2.5	1.2	0.200
COL11A1 rs1676486	T-C-T	-0.71	14.6	16.0	13.5	0.475	-0.841	16.3	15.2	0.400
+	T-C-A	-0.24	27.4	26.5	27.2	0.811	0.760	27.5	29.7	0.448
COL11A2 rs1799907	C-T-A	-0.12	11.5	10.7	11.7	0.902	0.980	12.7	15.2	0.327
	C-C-A	0.55	26.0	26.6	26.3	0.581	0.517	26.9	28.1	0.605
	C-C-T	0.66	5.0	4.1	5.6	0.509	-1.461	2.4	0.0	0.144
	C-T-T	0.85	9.8	7.3	12.1	0.397	0.320	8.6	10.6	0.749

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %. 0.929

Table B8.3: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL11A1* rs1676486 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-value
COL11A1 rs3753841	T-C	-1.25	42.0	42.4	40.8	0.213	-0.40	43.8	44.7	0.511
+	T-T	-1.10	5.7	8.9	3.4	0.270	-1.49	5.7	1.3	0.136
COL11A1 rs1676486	C-T	0.53	21.3	17.8	23.9	0.596	1.18	21.2	25.7	0.237
	C-C	0.67	30.9	30.9	31.9	0.502	0.40	29.6	28.3	0.689

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %. 0.518

Table B8.4: Inferred pseudo-haplotype analysis of *COL11A1* rs3753841 and *COL11A2* rs1799907 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-value
COL11A1 rs3753841	T-T	-1.23	31.7	32.7	30.1	0.219	-0.04	32.0	30.6	0.739
+	T-A	-0.97	16.1	18.8	14.0	0.333	-1.10	17.6	14.9	0.270
COL11A2 rs1799907	C-T	0.10	36.4	36.0	37.5	0.924	0.48	38.1	42.4	0.233
	C-A	0.76	15.8	12.5	18.5	0.449	-0.51	12.3	12.1	0.613

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %. 0.479

Table B8.5: Inferred pseudo-haplotype analysis of *COL11A1* rs1676486 and *COL11A2* rs1799907 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-value
COL11A1 rs1676486	T-T	-0.45	19.9	20.4	19.3	0.655	-1.18	19.1	16.5	0.201
+	C-A	-0.37	15.0	15.9	14.2	0.711	0.56	16.1	16.5	0.575
COL11A2 rs1799907	T-A	-0.36	53.1	52.8	53.4	0.722	0.09	54.0	56.5	0.930
	C-T	0.01	12.0	10.9	13.1	0.990	-0.43	10.7	10.5	0.670

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %. 0.545

Table B8.6: Inferred pseudo-haplotype analysis of *COL5A1* rs10628678, *COL11A1* rs3753841, *COL11A1* rs1676486 and *COL11A2* rs1799907 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-value
COL11A1 rs3753841	T-T-T-AGGG	-1.01	1.7	3.0	0.0	0.314	-0.80	2.1	0.0	0.425
+	T-C-A-AGGG	-1.00	17.2	21.2	12.7	0.098	-0.22	18.9	11.5	0.105
COL11A1 rs1676486	T-T-A-AGGG	-0.82	2.3	2.6	2.7	0.413	-	-	-	-
+	T-C-T-(-)	-0.44	8.9	10.7	6.4	0.659	-0.59	10.2	8.2	0.554
COL11A2 rs1799907	T-C-T-AGGG	-0.38	5.9	4.5	7.0	0.702	-0.51	5.9	8.0	0.608
+	C-T-A-(-)	0.06	5.9	6.0	5.8	0.954	0.74	6.9	8.7	0.456
COL5A1 rs10628678	C-C-T-AGGG	0.18	2.6	3.0	2.6	0.861	-	-	-	-
	C-C-A-(-)	0.33	6.8	7.5	5.5	0.739	0.39	7.5	4.7	0.696
	C-C-A-AGGG	0.39	19.5	18.9	20.8	0.698	0.29	19.5	23.3	0.774
	C-T-A-AGGG	0.40	5.3	4.5	5.0	0.691	0.84	5.7	5.6	0.400
	C-T-T-AGGG	0.61	6.9	5.2	9.0	0.545	0.61	7.3	10.9	0.542
	C-C-T-(-)	0.71	2.2	1.5	3.4	0.480	-	-	-	-
	C-T-T-(-)	0.96	3.0	2.2	3.3	0.337	-	-	-	-
	T-C-A-(-)	1.41	9.9	5.9	14.3	0.059	1.94	8.9	17.4	0.052

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %. 0.245

Table B8.7: Inferred pseudo-haplotype analysis of *COL5A1* rs10628678, *COL11A1* rs3753841 and *COL11A1* rs1676486 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-value
COL11A1 rs3753841	T-C-AGGG	-1.55	23.6	26.4	20.8	0.122	-0.97	25.3	22.9	0.333
+	T-T-(-)	-1.08	2.1	3.3	1.0	0.281	-1.44	2.1	0.0	0.150
COL11A1 rs1676486	T-T-AGGG	-1.06	3.7	6.1	2.4	0.287	-1.08	3.7	1.3	0.281
+	C-C-AGGG	0.35	21.3	21.9	21.5	0.728	0.00	20.3	15.0	0.998
COL5A1 rs10628678	T-C-(-)	0.54	18.5	15.5	20.3	0.592	0.82	18.6	21.8	0.413
	C-T-AGGG	0.61	12.8	8.6	15.1	0.544	1.03	13.1	21.2	0.303
	C-T-(-)	0.71	8.5	8.8	8.8	0.476	0.31	8.0	4.5	0.759
	C-C-(-)	0.96	9.6	9.4	10.2	0.338	0.39	9.0	12.7	0.696

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %. 0.961

Table B8.8: Inferred pseudo-haplotype analysis of *COL5A1* rs10628678, *COL11A1* rs3753841 and *COL11A1* rs1799907 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MAC L (92)	P-value	Hap Score	Hap Freq	NON (49)	P-value
COL11A1 rs3753841	T-A-AGGG	-1.35	19.8	24.1	15.8	0.176	-0.53	21.4	15.5	0.599
+	T-T-(-)	-0.72	8.8	11.4	6.1	0.472	-0.92	10.2	8.1	0.358
COL11A1 rs1799907	T-T-AGGG	-0.63	7.6	7.4	7.7	0.530	-0.54	7.8	8.0	0.589
+	T-A-(-)	0.39	11.7	8.5	14.6	0.693	0.91	10.2	14.8	0.363
COL5A1 rs10628678	C-A-AGGG	0.45	24.3	23.3	25.2	0.653	0.72	24.4	26.6	0.473
	C-T-AGGG	0.63	9.7	8.2	11.2	0.529	-0.23	8.7	10.9	0.819
	C-T-(-)	0.65	5.8	4.3	7.5	0.515	-0.96	3.2	0.0	0.338
	C-A-(-)	0.92	12.3	12.8	12.0	0.358	1.34	14.1	16.1	0.182

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %. 0.554

Table B8.9: Inferred pseudo-haplotype analysis of *COL5A1* rs10628678, *COL11A1* rs1676486 and *COL11A1* rs1799907 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MAC L (92)	P-value	Hap Score	Hap Freq	NON (49)	P-value
COL11A1 rs1676486	C-A-AGGG	-0.69	36.4	39.8	32.4	0.493	0.17	38.1	32.3	0.862
+	T-A-(-)	-0.55	7.0	8.3	5.1	0.584	-0.02	8.3	5.9	0.986
COL11A1 rs1799907	C-T-AGGG	-0.44	8.6	7.8	9.7	0.658	-1.04	7.5	7.6	0.297
+	C-T-(-)	-0.13	11.1	12.2	9.6	0.896	-0.90	11.1	8.4	0.368
COL5A1 rs10628678	T-T-AGGG	0.09	8.5	8.4	8.7	0.929	0.09	9.3	11.0	0.926
	T-A-AGGG	0.14	7.9	7.1	9.0	0.885	0.54	7.4	10.1	0.591
	T-T-(-)	0.45	3.7	3.0	4.5	0.650	-1.06	1.9	0.0	0.290
	C-A-(-)	1.08	16.8	13.5	21.0	0.280	1.43	16.4	24.7	0.152

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %. 0.861

Table B8.10: Inferred pseudo-haplotype analysis of *COL5A1* rs10628678 and *COL11A1* rs1676486 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-Value
COL5A1 rs10628678	T-AGGG	-2.02	27.5	31.8	23.5	0.736	-1.54	29.2	23.8	0.124
+	T-(-)	-0.22	20.5	19.8	20.7	0.822	0.04	20.5	22.2	0.970
COL11A1 rs3753841	C-AGGG	0.39	33.9	31.2	36.3	0.698	0.28	33.1	37.2	0.776
	C-(-)	1.53	18.1	17.2	19.5	0.126	1.12	17.2	16.8	0.264

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %. 0.825

Table B8.11: Inferred pseudo-haplotype analysis of *COL5A1* rs10628678 and *COL11A1* rs3753841 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-Value
COL5A1 rs10628678	C-AGGG	-0.82	44.9	48.1	41.9	0.410	-0.13	45.5	38.2	0.897
+	T-AGGG	-0.04	16.5	14.9	17.9	0.967	0.49	16.8	22.8	0.626
COL11A1 rs1676486	T-(-)	0.04	10.5	11.9	9.4	0.966	-0.57	10.1	4.2	0.570
	C-(-)	1.05	28.1	25.1	30.8	0.292	1.13	27.6	34.8	0.259

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %. 0.469

Table B8.12: Inferred pseudo-haplotype analysis of *COL5A1* rs10628678 and *COL11A2* rs1799907 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-Value
COL5A1 rs10628678	A-AGGG	-0.58	43.9	46.9	41.0	0.560	0.69	45.4	42.0	0.491
+	T-(-)	-0.36	14.4	15.3	13.7	0.717	-1.28	12.9	8.0	0.200
COL11A2 rs1799907	T-AGGG	0.01	17.5	16.1	18.7	0.992	-0.50	17.0	19.0	0.616
	A-(-)	0.66	24.2	21.7	26.5	0.508	1.36	24.8	31.0	0.173

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %. 0.527

Table B8.13 Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL6A1* rs35796750 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-Value
COL5A1 rs12722	C-T	-0.81	17.2	21.7	13.4	0.416	-0.60	19.3	16.4	0.547
+	T-T	-0.53	30.5	28.9	31.1	0.598	-0.01	29.6	29.4	0.994
COL6A1 rs35796750	T-C	0.17	31.1	29.8	33.3	0.869	0.48	29.2	29.6	0.630
	C-C	1.52	21.3	19.7	22.2	0.128	0.21	21.9	24.6	0.833

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %. 0.527

Table B8.14 Inferred pseudo-haplotype analysis of *COL5A1* rs12722 and *COL12A1* rs970547 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-value
COL5A1 rs12722	C-G	-1.63	24.2	23.8	23.1	0.103	0.03	15.7	15.5	0.973
+	C-A	-0.98	14.3	17.6	12.5	0.325	-0.56	25.5	25.5	0.579
COL12A1 rs970547	T-G	-0.44	38.9	40.6	38.5	0.661	-1.35	36.6	31.7	0.176
	T-A	1.02	22.7	17.9	26.0	0.090	0.77	22.2	27.3	**0.039***

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B8.15 Inferred pseudo-haplotype analysis of *COL6A1* rs35796750 and *COL12A1* rs970547 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-value
COL6A1 rs35796750	T-G	-1.75	29.1	33.4	24.0	0.080	-1.93	29.9	22.1	0.053
+	C-G	-0.40	34.0	31.1	37.2	**0.029***	-1.21	32.3	34.4	**0.027***
COL12A1 rs970547	C-A	0.37	18.4	18.4	18.2	0.712	0.59	18.8	19.9	0.555
	T-A	0.37	18.6	17.1	20.6	0.708	0.93	19.1	23.6	0.350

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

Table B8.16 Inferred pseudo-haplotype analysis of *COL5A1* rs12772, *COL6A1* rs35796750 and *COL12A1* rs970547 polymorphism in the control (CON) and Mixed Ancestry ACL Injury (MACL) and Non-Contact ACL Injury (NON) sub-group.

Gene-Gene Interactions	Inferred Allele Combinations	Hap Score	Hap Freq	CON (100)	MACL (92)	P-value	Hap Score	Hap Freq	NON (49)	P-value
COL5A1 rs12772	C-T-G	-2.09	6.4	9.7	2.9	0.051	-1.96	6.4	0.0	0.056
+	C-C-G	-0.84	17.7	16.5	18.0	0.400	-0.17	19.1	21.3	0.869
COL6A1	C-T-A	-0.54	11.1	11.8	11.3	0.586	0.65	12.9	16.9	0.515
rs35796750	T-T-G	-0.50	21.6	21.7	21.2	0.616	-0.83	21.7	21.8	0.409
+	C-C-A	-0.31	3.3	3.6	3.3	0.760	-0.30	3.1	2.8	0.764
COL12A1 rs970547	T-T-A	0.62	8.6	7.3	9.3	0.538	0.39	8.1	7.4	0.699
	T-C-G	0.79	17.1	16.7	18.3	0.427	-0.56	14.8	12.1	0.578
	T-C-A	1.34	14.1	12.7	15.7	0.179	1.12	14.0	17.8	0.265

CON group, MACL group and NON sub-group are represented as a frequency (%). Significant p-values ($p<0.05$) are indicated by an italicised bold type-face and an asterisk (*).
Hap.Score = Haplotype Score indicates di**rec**tion of haplotype, i.e. a negative score indicates over-representation in CON and positive score indicates over-representation in MACL, NON.
Hap Freq = Hap Frequency indicates frequency of inferred pseudo-haplotype and is represented by a %.

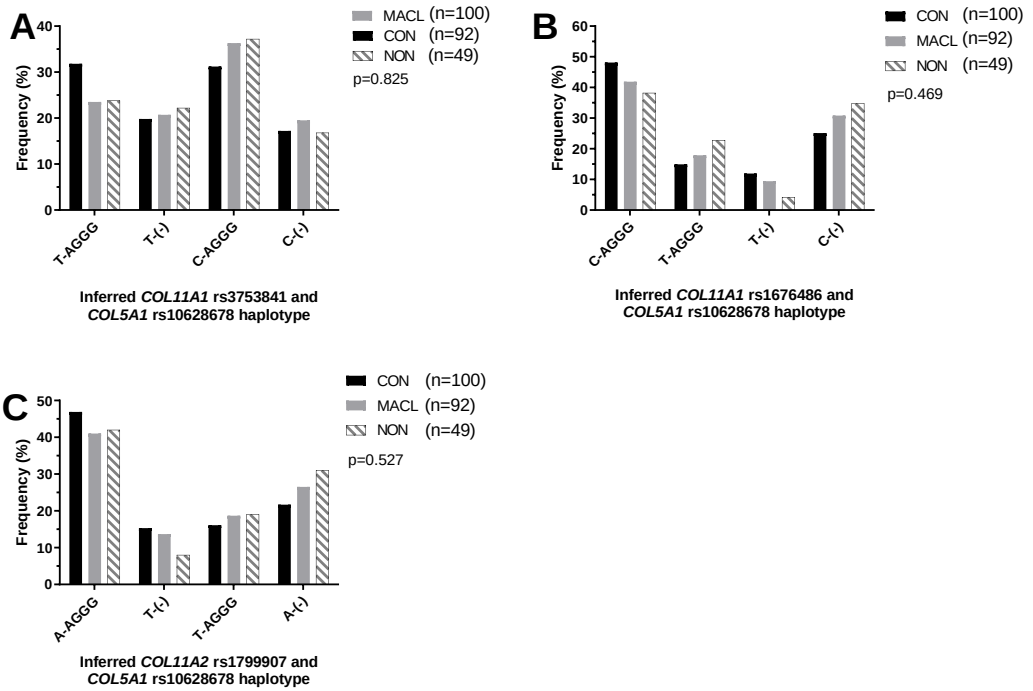


Figure B8.9 A) Inferred haplotype frequency distributions constructed from *COL11A1* rs3753841 (T/C) and *COL5A1* rs10628678 (AGGG/-). **B)** Inferred haplotype frequency distributions constructed from *COL11A1* rs1676486 (C/T) and *COL5A1* rs10628678 (AGGG/-) polymorphisms. **C)** Inferred haplotype frequency distributions constructed from *COL11A2* rs1799907 (A/T) and *COL5A1* rs10628678 (AGGG/-) for control group (CON, black bars) and the anterior cruciate ligament injury (MACL, grey bars) groups and non-contact injury subgroup (NON, hatched bars). The number (n) of subjects in each group is in parenthesis. Global p-value is indicated below the legend. AGGG: insertion allele, (-): deletion allele of *COL5A1* rs10628678 polymorphism.

End.