Association of Type XI Collagen Genes with Chronic Achilles Tendinopathy in Independent Populations from South Africa and Australia

Melanie Hay¹, Jon Patricios²,³, Robert Collins³,⁴, Andy Branfield⁴, Jill Cook⁵, Christopher J Handley⁶, Alison V September¹, Michael Posthumus¹, Malcolm Collins⁷,¹

¹MRC/UCT Research Unit for Exercise Science and Sports Medicine, Department of Human Biology, University of Cape Town, Cape Town, South Africa. ²Morningside Sports Medicine, Johannesburg, South Africa. ³Section of Sports Medicine, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa. ⁴The Centre for Sports Medicine & Orthopaedics, Johannesburg, South Africa. ⁵Department of Physiotherapy, Monash University, Melbourne, Victoria, Australia. ⁶School of Public Health and Human Biosciences, La Trobe University, Melbourne, Victoria, Australia. ⁷South African Medical Research Council, Cape Town, South Africa

Address for correspondence:
Prof M Collins
UCT/MRC Research Unit for Exercise Science and Sports Medicine
P.O. Box 115, Newlands 7725, South Africa
Tel: +27-21-650 4574
Fax: +27-21-686 7530
Email: malcolm.collins@uct.ac.za

Word Count: 2705
Running title: Type XI collagen genes and Achilles tendinopathy.
ABSTRACT

**Background**: Type XI collagen, which is expressed in developing tendons and is encoded by the *COL11A1*, *COL11A2* and *COL2A1* genes, shares structural and functional homology with type V collagen, which plays an important role in collagen fibril assembly. We investigated the association of these three polymorphisms with Achilles tendinopathy (AT) and whether these polymorphisms interact with *COL5A1* to modulate the risk of AT. **Methods**: 184 participants diagnosed with chronic AT (TEN) and 338 appropriately matched asymptomatic controls (CON) were genotyped for the three polymorphisms. **Results**: Although there were no independent associations with AT, the TCT pseudohaplotyte constructed from rs3753841 (T/C), rs1676486 (C/T) and rs1799907 (T/A) was significantly over-represented (p=0.006) in the TEN (25.9%) compared to the CON (17.1%) group. The TCT(AGGG) pseudohaplotypes constructed using these type XI collagen polymorphisms and the functional *COL5A1* rs71746744 (-/AGGG) polymorphism was also significantly over-represented (p<0.001) in the TEN (25.2%) compared to the CON (9.1%) group. **Discussion**: Genes encoding structural and functionally related type XI (*COL11A1* and *COL11A2*) and type V (*COL5A1*) collagens interact with one another to collectively modulate the risk for Achilles tendinopathy. Although there are no immediate clinical applications, the results of this study provide additional evidence that inter-individual variations in collagen fibril assembly might be an important molecular mechanism in the etiology of chronic Achilles tendinopathy. **Key words**: *COL11A1*, *COL11A2*, *COL5A1*, type V collagen, tendon, injury
INTRODUCTION

Injury to the mid-substance of the Achilles tendon is a multi-factorial condition, resulting from the poorly understood interactions of several extrinsic and intrinsic risk factors. 1 The genetic profile of an individual is an important intrinsic risk factor. 2 Although DNA sequence variants within several genes have been independently associated with risk of chronic Achilles tendinopathy 2 3, and one of the most extensively studied genes to date is COL5A1. 4 5 6 7 COL5A1 encodes the α1 chain of type V collagen, which plays an important role in collagen fibril nucleation and the regulation of fibril diameter (fibrillogenesis). 8 Several variants within the 3′-untranslated region (3′-UTR) of COL5A1 have been associated with Achilles tendinopathy. 5 7 These variants are believed to be functional, altering COL5A1 mRNA stability within the cytoplasm of the tenocyte. 6 Increased mRNA stability has been associated with tendinopathy, which is believed to result in increased α1(V) chain and type V collagen production, decreasing the fibril diameter and packing density, and potentially altering the mechanical properties of the tendon. 9

Type XI collagen, a heterotrimer consisting of α1(XI), α2(XI) and α3(XI) chains encoded by the COL11A1 (chromosome 1p21), COL11A2 (chromosome 6p21.3) and COL2A1 (chromosome 12q13.11) genes respectively 10, shares structural and functional homology with type V collagen. 11 Although predominately expressed in cartilage 11, type XI collagen is also expressed in developing tendons. 12

Mutations in all three genes have been implicated in various inherited Mendelian connective tissue disorders 13 and polymorphisms within these genes have also
been associated with multifactorial musculoskeletal injuries and connective tissue disorders. 14, 15, 16-21

Single nucleotide polymorphisms (SNPs), namely COL11A1 rs3753841 (T/C), COL11A1 rs1676486 (C/T) and/or COL11A2 rs1799907 (T/A), have been associated with lumbar disc herniation (LDH) 18, limbus vertebra in gymnasts 21, lumbar spine stenosis 17, ossification of the posterior longitudinal ligament of the spine (OPLL) 14-16, and rheumatoid arthritis. 19

Therefore, we investigated the association of the COL11A1 polymorphisms rs3753841 and rs1676486, as well as the COL11A2 rs1799907 polymorphism with Achilles Tendinopathy in a South African and an Australian population. A secondary aim of this study was to investigate whether the COL11A1 and COL11A2 polymorphisms interact with the previously investigated rs71746744 (-/AGGGG) polymorphism within the COL5A1 3'-UTR 7 to modulate the risk of Achilles Tendinopathy.

METHODS

Participants and DNA extraction
A total of 267 unrelated, physically active participants from South Africa, consisting of 161 asymptomatic controls (SA CON) and 106 individuals diagnosed with chronic Achilles tendinopathy (SA TEN) were recruited as previously described. 4, 22 The SA TEN group were recruited from Sports Medicine clinical practices in Cape Town and Johannesburg, South Africa. The diagnosis was confirmed by a Sports Physician,
using the previously described inclusion and exclusion criteria. The SA CON participants reported no history of tendon pathology and were recruited from various recreational sporting clubs. A total of 255 participants from Australia, consisting of 177 asymptomatic controls (AUS CON) and 78 patients with chronic Achilles tendinopathy (AUS TEN) were recruited as previously described. The diagnosis of the tendinopathy participants, who were recruited from the Musculoskeletal Research Centre at La Trobe University in Melbourne Australia, was confirmed using the same clinical criteria by a Sports Physiotherapist.

All participants were of self-reported Caucasian European ancestry. They gave written informed consent and completed questionnaires concerning their medical history and involvement in physical activity. This study was approved by the Human Research Ethics Committee at the University of Cape Town, South Africa, and the Human Ethics Committees of La Trobe and Deakin Universities in Melbourne, Australia. Total DNA was extracted from approximately 4.5 ml of venous blood as previously described.

**COL11A1 and COL11A2 Genotyping**

The COL11A1 rs3753841 (T/C) and rs1676486 (C/T) non-synonymous SNPs, as well as the COL11A2 rs1799907 (IVS6-4, T/A) SNP, were genotyped using fluorescence-based Taqman® PCR assays (Applied Biosystems, Foster City, CA, USA). Inventoried allele specific probes and flanking primer sets were used along with a pre-made PCR mastermix containing ampliTaq® DNA polymerase Gold (Applied Biosystems, Foster City, CA, USA) in a final reaction volume of 8 µl. The PCR reactions were conducted on an Applied Biosystems StepOnePlus™ Real-
Time PCR system (Applied Biosystems), using the Applied Biosystems Step-OnePlus™ Real-Time PCR software Version 2.2.2 (Applied Biosystems) following the manufacturers recommended cycling conditions.

**Statistical Analysis**

Quanto version 1.2.4, was used to determine the statistical power for a given sample size and minor allele frequency. Assuming a minor allele frequency of 0.36, 0.20 and 0.26 for rs3753841, rs1676486 and rs1799907 respectively, a sample size of approximately 180 cases and 360 controls, would be adequate to detect a genetic effect size of at least 2.25, 1.85 and 1.95 respectively, at a power of 80% and a significance level of 5%.

Data were analysed using STATISTICA Version 10.0 (Stat-Soft, Tulsa, OK, USA) and GraphPad Prism version 5.0d for Mac OS X (GraphPad Software, San Diego, CA, USA) programs. A one-way analysis of variance was used to determine any significant differences between the characteristics of the TEN and CON groups within the AUS, SA and combined cohorts. A χ2-analysis or Fisher’s exact test was used to analyse any differences in the genotype frequencies and other categorical data between the groups. Hardy-Weinberg equilibrium (HWE) was established using the program Genepop web version 3.4 (http://genepop.curtin.edu.au/). Linkage disequilibrium (LD) between the polymorphisms within *COL11A1* was calculated using CubeX: cubic exact solution (http://www.oege.org/software/cubex/). The Chaplin case-control haplotype inference package was used to infer haplotypes and pseudohaplotypes. Statistically significant differences were accepted when p<0.05.
RESULTS

Participant Characteristics

The SA CON and SA TEN groups were matched for height, sex and country of birth (Table 1). The SA TEN group was however significantly older (p=0.007) and heavier (weight: p=0.002; BMI: p=0.003) than the SA CON group. The mean difference between age of injury and age of recruitment of the SA TEN group was 7.1 ± 8.9 years. There was therefore also a significant difference between the age at recruitment of the SA TEN (48.2 ± 11.6 years, n=99) and the SA CON (36.4 ± 10.8 years, n=154, p<0.001) groups. When adjusted for age of recruitment, weight and BMI were no longer significantly different.

The AUS CON and AUS TEN groups were similarly matched for age and country of birth (Table 1). There was significantly more males in the AUS CON (40.3%) compared to the AUS TEN (71.8%, p<0.001) groups. After adjusting for sex, the groups were matched for height. The mean difference between age of injury and age of recruitment of the AUS TEN group was 9.0 ± 10.1 years. There was therefore a significant difference between the age at recruitment of the AUS TEN (49.6 ± 12.8 years, n=78) and the AUS CON (39.4 ± 12.3 years, n=174, p<0.001) groups. After adjusting for both age at recruitment and sex, the AUS CON and AUS TEN groups were also matched for weight and BMI (Table 1).
**Table 1:** Characteristics of the South African (SA) and Australian (AUS) chronic Achilles tendinopathy (TEN) groups and their respective control (CON) groups.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>SA CON</th>
<th>SA TEN</th>
<th>p value</th>
<th>AUS CON</th>
<th>AUS TEN</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td>36.4 ± 10.8 (154)</td>
<td>40.9 ± 14.8 (92)</td>
<td>0.007</td>
<td>39.4 ± 12.3 (174)</td>
<td>40.7 ± 14.5 (77)</td>
<td>0.487</td>
</tr>
<tr>
<td><strong>Sex (%male )</strong></td>
<td>63.8 (160)</td>
<td>67.6 (105)</td>
<td>0.518</td>
<td>40.3 (176)</td>
<td>71.8 (78)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Height (cm)</strong></td>
<td>175 ± 9 (155)</td>
<td>176 ± 9 (92)</td>
<td>0.345</td>
<td>171 ± 9 (175)</td>
<td>174 ± 10 (75)</td>
<td>0.059</td>
</tr>
<tr>
<td><strong>Weight (kg)</strong></td>
<td>72.2 ± 11.9 (159)</td>
<td>77.2 ± 13.4 (97)</td>
<td>0.002</td>
<td>72.7 ± 14.4 (176)</td>
<td>79.7 ± 13.4 (78)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>BMI (kg/cm²)</strong></td>
<td>23.6 ± 2.8 (151)</td>
<td>24.8 ± 3.3 (81)</td>
<td>0.003</td>
<td>24.7 ± 3.9 (175)</td>
<td>26.2 ± 3.5 (75)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Participants were included in the analysis if they were genotyped at least once for rs3753841, rs1676486 or rs1799907. Variables are expressed as mean ± standard deviation, except for sex and country of birth, which are represented as a percentage (%). The number of participants for which data was available is in parenthesis (n). Age of CON groups = age of recruitment; age of TEN groups = age of initial onset of symptoms. Weight, height and BMI values are those reported at recruitment. BMI, body mass index; n.d., not determined.

- Covaried for age of recruitment.
- Covaried for sex.
- Covaried for age of recruitment and sex.
When the SA and AUS groups were combined, there were significantly more males in the TEN group (69.4%) compared to the CON group (51.5%, p<0.001). The TEN group (age of initial injury: 40.8 ± 14.7 years, weight: 78.3 ± 13.4 kg) was also significantly older and heavier than the CON group (age of recruitment: 38.0 ± 11.6 years, p=0.022; weight: 72.5 ± 13.2 kg, adjusted p=0.049). The groups were however matched for height (TEN: 175 ± 10 cm, CON: 173 ± 9 cm, adjusted p=0.419) and BMI (TEN: 25.5 ± 3.5 kg/cm², CON: 24.1 ± 3.5 kg/cm², adjusted p=0.214).

There were no significant COL11A1 or COL11A2 genotype effects on any of the combined SA and AUS participant characteristics (data not shown).

**COL11A1 and COL11A2 genotype and allele frequency distributions**

There were no significant differences in any of the genotype or allele frequency distributions between the SA CON and SA TEN or the AUS CON and AUS TEN groups for COL11A1 rs3753841 (T/C), COL11A1 rs1676486 (C/T) and COL11A2 rs1799907 (T/A) (Supplementary Table 1). In addition, there were no significant differences in the genotype distribution of these three polymorphisms between the participants from South Africa and Australia (Supplementary Table 1). For this reason, the two populations were combined. As illustrated in figure 1, there were no significant differences (p>0.100) in the genotype frequency distributions between CON and TEN groups of the combined Australian and South African cohorts when any of the three polymorphisms were analysed. There was also no significant difference in the COL11A1 rs3753841 (p=0.121, Figure 1C) and rs1676486 (p=0.221, Figure 1D) allele frequency distributions between combined CON and TEN
Figure 1

Genotype and allele frequency distributions of the combined South African and Australian CON (black bars) and TEN (clear bars) groups for COL11A1 rs3753841 (T/C), COL11A1 rs1676486 (C/T) and COL11A2 rs1799907 (T/A). (A) rs3753841 genotype distribution, CON HWE=0.795, TEN HWE=0.052. (B) rs1676486 genotype distribution, CON HWE=0.082, TEN HWE=0.602. (C) rs1799907 genotype frequency distribution CON HWE=1.000, TEN HWE=0.100. (D) rs3753841 allele frequency distribution. (E) rs1676486 allele frequency distribution. (F) rs1799907 allele frequency distribution. The number of genotyped participants (n) for each polymorphism is indicated in parenthesis in graphs A to C. The number of alleles (n, 2 for each participant) for each polymorphism is indicated parenthesis in graphs D to F. The p value is also indicated in each graph.
groups. However a tendency for the minor T allele of COL11A2 rs1799907 (p=0.095) to be over-represented in the combined TEN group was noted (Figure 1F). Although there was a tendency for some to deviate, all the groups were in HWE (Table 1 and Figure 1).

**COL11A1 and COL11A2 interactions**

![Figure 2](image)

The frequency of seven of the possible eight inferred pseudohaplotypes with a frequency >0.4% constructed from COL11A1 rs3753841 (T/C), COL11A1 rs1676486 (C/T) and COL11A2 rs1799907 (T/A) in the combined Australian and South African Achilles tendinopathy (TEN, clear bars) and control (CON, solid bars) groups. The p value of the significantly different TCT pseudohaplotype is indicated. The total number (n) of pseudohaplotypes within the CON and TEN groups is indicated in parenthesis in the graph.

Polymorphisms COL11A1 rs3753841 and rs1676486 were in high linkage disequilibrium for the CON (SA D'=0.862; AUS D'=0.911) and TEN (SA D'=1.000; AUS D'=0.898) groups, as well as the combined group (D'=0.932). Since the three polymorphisms investigated in this study are located within type XI collagen genes,
inferred pseudohaplotypes were constructed. Seven of the possible eight pseudohaplotypes, were inferred with a frequency greater than 0.4% from the two COL11A1 and one COL11A2 polymorphisms (Figure 2). The TCT pseudohaplotype was significantly over-represented (p=0.006) in the TEN group (25.9%, n=47) compared to the CON group (17.1%, n=57) when the combined populations were investigated (Figure 2). Similar pseudohaplotype distributions were obtained when the SA and AUS cohorts were analysed separately, with the TCT pseudohaplotype being significantly over-represented in the TEN group of both the SA (CON 19.6% vs TEN 29.7%, p=0.017) and AUS (CON 12.9% vs TEN 21.3%, p=0.041) cohorts (Supplementary Figure 1).

**Interactions between COL11A1, COL11A2 and COL5A1**

Type V and type XI collagen both regulate fibrillogenesis, therefore inferred pseudohaplotypes were constructed between the COL11A1 and COL11A2 polymorphisms investigated in this study, and the previously investigated COL5A1 rs71746744 (-/AGGG) polymorphism. Eleven of the possible 16 pseudohaplotypes were inferred with a frequency greater than 2% from the COL11A1, COL11A2 and COL5A1 polymorphisms (Figure 3). The TCT(AGGG) pseudohaplotype was significantly over-represented (p=0.003) in the TEN group (25.2%, n=46) compared to the CON group (9.1%, n=30) (Figure 3). Similar pseudohaplotype distributions were obtained when the SA and AUS cohorts were analysed separately, with the TCT(AGGG) pseudohaplotype being significantly over-represented in the TEN group of both the SA (CON 10.0% vs TEN 29.4%, p=0.003) and AUS (CON 7.2% vs TEN 20.3%, p=0.008) cohorts (Supplementary Figure 2).
Figure 3

The frequency of eleven of the possible 16 inferred pseudohaplotypes with a frequency >2% constructed from \( \text{COL11A1} \) rs3753841 (T/C), \( \text{COL11A1} \) rs1676486 (C/T), \( \text{COL11A2} \) rs1799907 (T/A) and \( \text{COL5A1} \) rs71746744 (-/AGGG) in the combined Australian and South African Achilles tendinopathy (TEN, clear bars) and control (CON, solid bars) groups. The p value of the significantly different TCT(AGGG) pseudohaplotype is indicated. The total number (n) of pseudohaplotypes within the CON and TEN groups is indicated in parenthesis in the graph.

DISCUSSION

Polymorphisms in several genes, including \( \text{COL5A1} \), have been identified as intrinsic risk factors for Achilles tendinopathy. \(^2\) \( \text{COL5A1} \) encodes the \( \alpha 1 \) chain of type V collagen, which plays an important role in collagen fibril assembly; the basic building block of tendons. \(^8\) Type XI collagen, which is encoded by the \( \text{COL11A1}, \text{COL11A2} \) and \( \text{COL2A1} \) genes \(^10\), shares structural and functional homology with type V collagen. \(^11\)
None of the three investigated polymorphisms within the *COL11A1* (rs3753841 and rs1676486) and *COL11A2* (rs1799907) genes were independently associated with chronic Achilles tendinopathy in the Australian, South African or combined cohorts. The genotype and allele frequencies calculated for these three SNPs were similar to the reported frequencies presented in the NCBI SNP database (http://www.ncbi.nlm.nih.gov/snp/).

The main finding of this study was the association of the TCT inferred pseudohaplotype, constructed from the three polymorphisms within *COL11A1* (rs3753841 T/C and rs1676486 C/T) and *COL11A2* (rs17999079 T/A), with increased risk of chronic Achilles tendinopathy. The non-synonymous rs3753841 polymorphism within exon 52 of *COL11A1*, results in a predicted deleterious substitution of a leucine for a proline at amino acid position 1323 of the α1(XI) chain. The rs1676486 polymorphism within exon 62, also results in an amino acid substitution, proline to serine at position 1535, which could potentially cause a conformational change in type XI collagen. In addition, the T allele of rs1676486 also appears to be associated with increased mRNA degradation.

Furthermore, the T allele of *COL11A2* rs1799907 (IVS6-4, T/A), which was also implicated in the pseudohaplotypes, produces a distinct isoform of the α2(XI) chain in which several amino acids in the amino terminal acidic domain are deleted. This acidic domain provides potential sites for the interaction of type XI collagen with other molecules and may prevent further deposition of collagen molecules in the fibril. It is therefore reasonable to hypothesise that the risk associated
pseudohaplotype has a biological consequence. Although expressed in the
developing tendon, further research is however required to determine the functional
significance of the type XI collagen gene polymorphisms in tendinopathy.

Intriguingly, while the minor C allele of \textit{COL11A1} rs3753841 and the minor T allele of
\textit{COL11A1} of rs1676486 was over-represented in patients with lumber disc
hermiation, the current study however implicated the alternate alleles in the risk-
associated pseudohaplotype. Similarly, the functional rare TT genotype of the
\textit{COL1A1} Sp1 binding site polymorphism was reported to be associated with risk for
several multifactorial disorders, including lumber disc disease \cite{28-31}, but was
associated with protection in anterior cruciate ligament ruptures in three independent
populations. \cite{32-34} These findings suggest that both the alternative alleles within a
gene can be associated with increased risk of different multifactorial disorders.

An additional finding of this study was the additive contribution of the AGGG allele of
the \textit{COL5A1} -/AGGG polymorphism together with the TCT inferred pseudohaplotype
of the \textit{COL11A1} and \textit{COL11A2} genes increased risk of chronic Achilles
tendinopathy. This gene-gene association is perhaps not surprising taking into
account the mounting evidence that the structural and functional homology between
the α1(XI) and α1(V) chains encoded by \textit{COL11A1} and \textit{COL5A1}, respectively,
facilitate an interchangeability between the two polypeptides resulting in heterotrimer
[α1(XI)2α2(V)] formation. \cite{35 36 37 38} It is possible that the reported interaction in this
study between the \textit{COL5A1} and \textit{COL11A1} polymorphisms indicate the role of a
minor fibrillar collagen consisting of type V/XI α-chains in tendinopathy. Alternatively
the traditional type XI collagen heterotimer may functionally replace or compensate
for type V collagen in tendons. The possible role of type XI collagen in pathology of
the mature tendon needs to be investigated.

Although we are mindful of the theoretical biological effects of the amino acid
substitutions implicated by the risk associated pseudohaplotype within the α1(XI) and
α2(XI) chains, it is nevertheless tempting to speculate that the association between
the COL11A1, COL11A2 and COL5A1 genes and increased risk of chronic Achilles
tendinoathy is due to their collective effect on (i) mRNA stability, (ii) potential effects
on types V and XI collagen production, regulation of collagen fibril diameter, and (iv)
altered biomechanical properties of the collagen fibril that is reflected at the tissue
level (Figure 4). It is therefore reasonable to conceive that individuals who have
these functional alleles as reflected from the pseudohaplotype could possibly be
producing a functionally altered types XI and type V collagen which collectively are
responsible for the altered biomechanical property of the tendon collagen fibrils. As
previously mentioned although type XI collagen is classified as a cartilage protein, it
is produced in the developing tendon. To our knowledge there is no evidence
that the protein is produced in mature tendons. It is therefore possible that the
proposed biological consequences of the reported association in this study could be
due to altered protein profiling: (i) the expression of type XI collagen in the mature,
diseased and/or healing tendon, (ii) its expression only in the fibrocartilaginous
regions of the mature tendon and (iii) the result of its expression and function during
tendon development. Irrespective of the mechanism(s), further research is required
to replicate these findings in larger independent populations and to explore the
functional mechanisms underlying the complex genetic associations with the type XI
Figure 4

A hypothetical schematic diagram illustrating the proposed mechanism of how polymorphisms within COL5A1 and COL11A1 potentially affect fibrilogenesis. Although there is no evidence that type XI collagen is produced in the mature, diseased and/or healing tendon, it is produced and functional during tendon development. \(^{12}\) (A) The COL5A1 rs71746744 (-/AGGG) and COL11A1 rs1676486 (C/T) polymorphisms are part of an inferred pseudohaplotype that is associated with chronic Achilles tendinopathy. The COL5A1 rs71746744 - allele \(^{6,7}\) and the COL11A1 rs1676486 T allele \(^{18}\) are both believed to associated with increased mRNA degradation. Increased mRNA degradation is indicated in the left panel, while decreased mRNA degradation is indicated in the right panel). (B) The altered mRNA stability associated with these polymorphisms is believed to result in altered α1(V) and α1(XI) chain and types V and XI collagen production (decreased in left panel and increased in the right panel). (C) Types V and XI collagen regulates collagen fibril assembly and diameter (fibrillogenesis) and thus the mechanical properties of tendons. There is an inverse relationship between the types V and XI collagen content of the fibril and its diameter. Thinner more densely packed collagen fibrils are produced due to the increased production of types V and XI collagen (Right Panel). It has previously been proposed that the thinner fibrils are associated with chronic Achilles tendinopathy. \(^{9}\)
collagen encoding genes and their interactions with the type V collagen encoding gene.

In conclusion, the functional variants within the type XI collagen genes investigated in this study were not independently associated with chronic Achilles tendinopathy. This study does however provide evidence suggesting that the genes that encode for the structurally and functionally related type XI (COL11A1 and COL11A2) and type V (COL5A1) collagens interact with one another to collectively modulate the risk for Achilles tendinopathy. Although expressed in the developing tendon, the role of type XI collagen in the pathology of the mature tendon requires future investigation.

"What are the new findings"
- Functional polymorphisms within the type XI collagen genes (COL11A1 and COL11A2) were not independently associated with chronic Achilles tendinopathy.
- These functional type XI collagen polymorphisms interact with a functional polymorphism within the type V collagen (COL5A1) gene to modulate the risk for Achilles tendinopathy in two independent populations.
- Since types XI and V collagen are structurally and functionally related, we propose that these polymorphisms interact to regulate collagen fibril assembly.

"How might it impact on clinical practice in the near future"
- Although there are no immediate clinical applications, the results of this study provide additional evidence that inter-individual variations in collagen fibril, the basic building blocks of tendons, assembly might be an important molecular mechanism in the etiology of chronic Achilles tendinopathy.
- Genetic risk factors could one day be included into multifactorial models to determine an individuals risk for Achilles tendinopathy.
• Genetic risk factors could however never to be used in isolation to diagnose or predict these injuries.

ACKNOWLEDGEMENTS

This study was supported in part by funds from the National Research Foundation of South Africa, University of Cape Town, and the South African Medical Research Council. The authors would like to thank Dr. G.G. Mokone for recruiting the SA participants, Prof. M.P. Schwellnus for the clinical diagnoses of the South African chronic Achilles tendinopathy participants, as well as N. Kariem for his laboratory assistance. M.P. was supported by the Thembakazi Trust. Authors M.C. and A.S. have filed patents on the application of specific sequence variations (not included in this manuscript) related to risk assessment of AT and anterior cruciate ligament injuries.

REFERENCES


19 Lee H-S, Lee AT, Criswell LA, Seldin MF, Amos CI, Carulli JP, et al. Several regions in the major histocompatibility complex confer risk for anti-CCP-antibody


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


35 Brown KE, Lawrence R, Sonenshein GE. Concerted modulation of alpha 1(XI)


Supplementary Table 1. Genotype and allele frequency distribution of the functional single nucleotide polymorphisms (SNPs) rs3753841 and rs1676486 within \textit{COL11A1} and rs1799907 within \textit{COL11A2}, of the South African (SA) and Australian (AUS) chronic Achilles tendinopathy (TEN) groups and their respective control (CON) groups.

<table>
<thead>
<tr>
<th>\textit{COL11A1}</th>
<th>rs3753841 genotype</th>
<th>rs3753841 allele</th>
<th>\textit{COL11A2}</th>
<th>rs1799907 genotype</th>
<th>rs1799907 allele</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>TT</td>
<td>TC</td>
<td>CC</td>
<td>P value</td>
</tr>
<tr>
<td>\textbf{South Africa}</td>
<td>247</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SA CON</td>
<td>150</td>
<td>37.3 (56)</td>
<td>46.7 (70)</td>
<td>16.0 (24)</td>
<td>0.861</td>
</tr>
<tr>
<td>SA TEN</td>
<td>97</td>
<td>46.4 (45)</td>
<td>40.2 (39)</td>
<td>13.4 (13)</td>
<td>0.367</td>
</tr>
<tr>
<td>\textbf{Australia}</td>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUS CON</td>
<td>104</td>
<td>29.8 (31)</td>
<td>49.0 (51)</td>
<td>21.2 (22)</td>
<td>1.000</td>
</tr>
<tr>
<td>AUS TEN</td>
<td>56</td>
<td>41.1 (23)</td>
<td>35.7 (20)</td>
<td>23.2 (13)</td>
<td>0.235</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\textbf{COL11A1}</th>
<th>rs1676486 genotype</th>
<th>rs1676486 allele</th>
<th>\textbf{COL11A2}</th>
<th>rs1799907 genotype</th>
<th>rs1799907 allele</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>CC</td>
<td>CT</td>
<td>TT</td>
<td>P value</td>
</tr>
<tr>
<td>\textbf{South Africa}</td>
<td>230</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SA CON</td>
<td>139</td>
<td>64.0 (89)</td>
<td>29.5 (41)</td>
<td>6.5 (9)</td>
<td>0.195</td>
</tr>
<tr>
<td>SA TEN</td>
<td>91</td>
<td>68.1 (62)</td>
<td>27.5 (25)</td>
<td>4.4 (4)</td>
<td>0.571a</td>
</tr>
<tr>
<td>\textbf{Australia}</td>
<td>205</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUS CON</td>
<td>138</td>
<td>62.3 (86)</td>
<td>31.2 (43)</td>
<td>6.5 (9)</td>
<td>0.318</td>
</tr>
<tr>
<td>AUS TEN</td>
<td>67</td>
<td>67.2 (45)</td>
<td>29.9 (20)</td>
<td>3.0 (2)</td>
<td>0.538a</td>
</tr>
</tbody>
</table>

Values are percentages, with sample number of participants (n) displayed in parenthesis. HWE is the p-value for the exact tests for Hardy-Weinberg Equilibrium. \(^a\) CC vs CT + TT.

SA CON vs AUS CON: rs3753841 p=0.370, rs1676486 p=0.953 and rs1799907 p=0.500.

SA TEN vs AUS TEN: rs3753841 p=0.298, rs1676486 p=01.000 and rs1799907 p=0.885.
Supplementary Figure 1

The frequency of the eight inferred pseudohaplotypes constructed from **COL11A1** rs3753841 (C/T), **COL11A1** rs1676486 (T/C), **COL11A2** rs1799907 (T/A) and **COL5A1** rs71746744 (-/AGGG) in the (A) South African (SA) and (B) Australian (AUS) chronic Achilles tendinopathy (TEN, clear bars) and control (CON, solid bars) groups. The p value of the significantly different TCT pseudohaplotype is indicated in each graph. The total number (n) of inferred pseudohaplotypes within the CON and TEN groups is indicated in parenthesis in the graph.
Supplementary Figure 2

The frequency of eleven of the possible 16 inferred pseudohaplotypes with a frequency >2% constructed from COL11A1 rs3753841 (C/T), COL11A1 rs1676486 (T/C), COL11A2 rs1799907 (T/A) and COL5A1 rs7174644 (-/AGGG) in the (A) South African (SA) and (B) Australian (AUS) chronic Achilles tendinopathy (TEN, clear bars) and control (CON, solid bars) groups. The p value of the significantly different TCT(AGGG) pseudohaplotype is indicated in each graph. The total number (n) of inferred pseudohaplotypes within the CON and TEN groups is indicated in parenthesis in the graphs.