Variants within the COMP and THBS2 genes are not associated with Achilles tendinopathy in a case-control study of South African and Australian populations

Colleen J. Saunders, Lize Van Der Merwe, Jill Cook, Christopher J. Handley, Malcolm Collins and Alison V. September

Abstract
Cartilage oligomeric matrix protein is a structural protein of the extracellular matrix, while thrombospondin-2 is a matricellular protein involved in cell–matrix interactions. Recent studies have shown that genetic variation is a significant risk factor for Achilles tendinopathy, and the genes encoding cartilage oligomeric matrix protein (COMP) and thrombospondin-2 (THBS2) were identified as good candidate genes for association with Achilles tendinopathy. This study aimed to test the association of sequence variants within these candidate genes with the risk of Achilles tendinopathy in participants from South Africa (SA) and Australia (AUS). Three-hundred and forty (133 SA; 207 AUS) control participants with no history of Achilles tendinopathy and 178 (94 SA; 84 AUS) participants clinically diagnosed with Achilles tendinopathy were genotyped for five single nucleotide polymorphisms within the COMP and THBS2 genes in this case-control study. There was no difference in genotype distributions between control and tendinopathy groups for either the THBS2 variants rs9505888, rs6422747 and rs9283850, or the COMP variants rs730079 and rs28494505 in the SA and AUS populations. As the selection of COMP and THBS2 as candidate genes was hypothesis driven, based on biological function, the possibility that other variants within these genes are associated with Achilles tendinopathy cannot be excluded.

Introduction
The Achilles tendon is exposed to high mechanical forces during exercise and is often a site of overuse injuries in athletes and the general population (de Jonge et al., 2011; Jarvinen et al., 2001). Achilles tendinopathy is a clinical syndrome comprising of swelling in and around the tendon, tenderness to palpation and pain-induced impairment of physical performance (Jarvinen et al., 2001; Maffulli, Khan, & Puddu, 1998; Rees, Dent, & Caterson, 2009; Xu & Murrell, 2008). Achilles tendinopathy encompasses a range of more specific diagnoses, which includes, amongst others, tendinosis, a condition characterised by chronic degeneration of the tendon without histological signs of inflammation (Jarvinen et al., 1997; Khan, Cook, Kannus, Maffulli, & Bonar, 2002). The cumulative lifetime incidence of Achilles tendinopathy in former elite athletes has been estimated at 23.9%, and is as high as 52% in middle and long distance runners (Kujala, Sarna, & Kaprio, 2005). Although the pathology of Achilles tendinopathy has been well characterised with clearly documented clinical
and imaging diagnostic criteria (Del Buono, Chan, & Maffulli, 2013; Jarvinen et al., 1997; Kader, Saxena, Movin, & Maffulli, 2002), the exact aetiology of the condition remains undefined. Several extrinsic and intrinsic risk factors for developing Achilles tendinopathy have been proposed, including a significant heritable component (Collins & Raleigh, 2009; Kannus, 1997; Kannus & Natri, 1997; September, Mokone, Schwellnus, & Collins, 2006). To date, polymorphisms in several genes coding for structural and associated proteins of the extracellular matrix have been investigated as candidate genes for the risk of developing Achilles tendinopathy. In particular, sequence variants within the COL5A1, TNC, GDF-5, MMP3, CASP8 and COL27A1 genes, as well as the interleukin genes IL-1β, IL-1RN and IL-6, have been shown to influence an individual’s susceptibility to Achilles tendinopathy (Mokone et al., 2005; Mokone, Schwellnus, Noakes, & Collins, 2006; Nell et al., 2012; Posthumus et al., 2010; Raleigh et al., 2009; Saunders et al., 2013; September et al., 2009; September et al., 2011).

In this study we investigated two further candidate genes, THBS2 and COMP, which code for thrombospondin-2 and cartilage oligomeric matrix protein (thrombospondin-5) respectively. All thrombospondins contain a highly conserved signature protein domain at the carboxy terminal and a number of disease associated genetic variants have been mapped to this domain (Carlson, Lawler, & Mosher, 2008). Thrombospondin-2, in particular, is involved in the healing response in connective tissues and plays a significant role in cell–matrix interactions (Adams & Lawler, 2004; Bornstein, Agah, & Kyriakides, 2004; Bornstein, Armstrong, Hankenson, Kyriakides, & Yang, 2000). Two single nucleotide polymorphisms (SNPs) within the THBS2 gene (chr6q27), rs9283850 (Intron 9) and rs6422747 (Intron 14), have previously been implicated in a haplotype associated with lumbar disc herniation in a Japanese population (Hirose et al., 2008). Cartilage oligomeric matrix protein interacts with type I and type III collagens and plays an important role in matrix assembly and the repair of injured tissues (Hecht, Hayes, Haynes, & Cole, 2005; Posey, Yang, Veerisetty, Sharan, & Hecht, 2008; Rosenberg, Olsson, Morgelin, & Heinegard, 1998; Sodersten, Hultenby, Heinegard, Johnston, & Ekman, 2013). Genetic variants clustered in the conserved calcium binding domains and carboxy terminal of the cartilage oligomeric matrix protein have been repeatedly associated with skeletal dysplasias (Adams & Lawler, 2004; Hecht et al., 2005; Ikegawa et al., 1998) and other cartilage disorders including osteoarthritis of the knee (Valdes et al., 2007). The C-allele of the rs730079 (-1417C>G) SNP within the 5’-UTR of the COMP gene (chr19p13.11) was previously associated with osteoarthritis in a Caucasian male population (Valdes et al., 2007), while the rs28494505 (Intron 18) Tag SNP is located within the highly conserved C-terminal domain of the Cartilage oligomeric matrix protein. These observations, together with the biological roles of these proteins, suggest that both COMP and THBS2 are good candidate genes for association with risk of developing Achilles tendinopathy. The aim of this study was,
therefore, to test the association of sequence variants within the COMP, rs730079 (C>G) and rs28494505 (A>G), and THBS2, rs9283850 (A>G), rs6422747 (A>G) and rs9505888 (A>G), candidate genes with the risk of Achilles tendinopathy in participants from South Africa and Australia.

Methods
As previously described, 133 South African asymptomatic control participants (SA CON) and 94 South African participants clinically diagnosed with AT (SA TEN), as well as 207 Australian asymptomatic control participants (AUS CON) and 84 Australian participants clinically diagnosed with AT (AUS TEN) were recruited for this study (Mokone et al., 2005; September et al., 2009). In particular, tendinopathy participants were diagnosed with chronic, degenerative tendinosis without signs of inflammation. The presence of gradual progressive pain over the posterior lower leg for more than six months, together with at least one of the following clinical criteria, was used to diagnose chronic Achilles tendinopathy: (i) early morning pain; (ii) early morning stiffness; (iii) history of swelling over the Achilles tendon area; (iv) tenderness to palpation; (v) palpable nodular thickening over the affected Achilles; and/or (vi) a positive shift test during plantar-/dorsi-flexion (Kader et al., 2002; Schepsis, Jones, & Haas, 2002). The diagnosis of each tendinopathy participant was reviewed by the same clinician in SA and AUS, and was confirmed using soft tissue ultrasound in a sub-group of SA participants and all the AUS participants. Control participants were physically active and asymptomatic with no history of Achilles tendon injury. Individuals with a self-reported history of treatment with either fluoroquinolone antibiotics or local corticosteroid injections, as well as those who had previously suffered, or were currently suffering from any connective tissue disorders or other systemic diseases associated with Achilles tendinopathy were excluded from the study. All participants were of self-reported Caucasian ancestry, gave written informed consent in accordance with the Declaration of Helsinki and completed a questionnaire providing personal particulars and medical history. This study was approved by the Human Research Ethics Committee of the Faculty of Health Sciences, University of Cape Town and the Human Ethics Committee of La Trobe and Monash Universities, Melbourne, Australia.

The three THBS2 SNPs rs9283850 (A>G), rs6422747 (A>G) and rs9505888 (A>G) and the two COMP SNPs, rs730079 (-1417C>G) and rs28494505 (A>G), were selected based on their linkage, heterozygosity score (>35%), minor allele frequency (>20%) and previous association with other clinical phenotypes (Figures 1 and 2) (Hirose et al., 2008; Valdes et al., 2007). The rs9505888 SNP was selected based on its linkage to rs6422747. In addition, all three THBS2 SNPs are identified as Tag SNPs (http://gvs.gs.washington.edu/GVS137) and therefore provide moderate coverage of this 38.3kb gene. The rs28494505 variant was selected based on its location within the highly conserved C-terminal domain of the cartilage oligomeric matrix protein and its identification as a TagSNP.
DNA for all South African (227) and Australian (291) participants was isolated from venous blood as described by Lahiri and Nurnberger (1991) and modified by Mokone et al. (2005). Variants rs9505888, rs6422747 and rs28494505 were genotyped using standard polymerase chain reaction (PCR) and restriction fragment length polymorphism techniques (PCR conditions detailed in supplementary material). PCR reactions were performed on a thermal cycler (Hybaid: PCR Express, Middlesex, UK) with the resulting amplicon incubated with the BstUI, MfeI and MspI nucleases respectively. Restriction fragments were resolved, together with a 100bp size standard, by 6% polyacrylamide gel electrophoresis and visualised under UV light after staining with SYBR Gold nucleic acid gel stain (Invitrogen). Variants rs9283850 and rs730079, as well as a subset of samples for rs28494505, were genotyped using TaqMan SNP Genotyping Assays (Applied Biosystems) that were amplified and distinguished using the StepOnePlus Real-Time PCR System (Applied Biosystems). For quality control purposes, a number of positive and DNA-free controls were included in all PCR reactions, and a small subset of samples was analysed for repeatability within each variant.

The programming environment R (R Development Core Team, 2010), and R package, genetics (Warnes, Gorjanc, Leisch, & others, 2011), was used for all statistical analyses. Age, sex, country grouping and country of birth were considered potential confounders, and so were corrected for in the logistic regression models of genotype and allele association with
Achilles tendinopathy, as well as in the linear models of the association of genotype and allele frequencies with physical characteristics of the participants.

**Results**

Basic characteristics of the study groups were presented previously and are included in (Mokone et al., 2005; September et al., 2009). Tendinopathy participants were significantly heavier than the control participants at the time of recruitment (P<0.001); however, this difference was much less pronounced after adjusting for country grouping, sex and age (P=0.044). Genotype and minor allele frequency distributions for each of the polymorphisms, together with Hardy-Weinberg Equilibrium (HWE) p-values, are shown in Table II. The frequency distributions of the SNPs tested in this study were found to differ significantly between the South African and Australian participant groups at the rs28494505 (P<0.001) and rs6422747 loci (P=0.024) (Table II).

Table I

Figure 2. (A) Chromosomal location of the COMP gene. (B) Exonic structure of COMP* Significant association with osteoarthritis (Valdes et al., 2007) **Significant association with PSACH and MED (Delot, King, Briggs, Wilcox, & Cohn, 1999) Variants in bold are included in the present study. (C) COMP protein domain architecture. UTR: Untranslated region; EGF: Epidermal growth factor; THBS-N: Thrombospondin N-terminal domain. (Compiled from www.ensembl.org and Carlson et al., 2008.)
The two populations were therefore summarised separately. All genotype distributions were in HWE except at the rs28494505 locus that deviated significantly in the AUS control group ($P=0.043$). It was, however, noted that a lower genotype call rate was observed for this variant in the AUS samples. There were no significant genotype effects on age, weight, height, BMI or sex after adjusting for possible confounders (Table III). After adjusting for the potential confounder’s age, sex, country grouping and country of birth, there were no significant differences in either the genotype or allele frequency distributions between the control and tendinopathy groups.

<table>
<thead>
<tr>
<th></th>
<th>CON AUS ($n = 207$)</th>
<th>CON SA ($n = 133$)</th>
<th>TEN AUS ($n = 84$)</th>
<th>TEN SA ($n = 94$)</th>
<th>$P$-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height (cm)</td>
<td>171 ± 9</td>
<td>175 ± 9</td>
<td>174 ± 10</td>
<td>176 ± 9</td>
<td>&gt;0.001</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>73 ± 14</td>
<td>71 ± 12</td>
<td>80 ± 14</td>
<td>78 ± 13</td>
<td>0.138</td>
</tr>
<tr>
<td>BMI (kg · m$^{-2}$)</td>
<td>24.7 ± 3.9</td>
<td>23.3 ± 2.7</td>
<td>26.5 ± 3.8</td>
<td>24.8 ± 3.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age (years)</td>
<td>39 ± 12</td>
<td>37 ± 11</td>
<td>40 ± 14</td>
<td>39 ± 15</td>
<td>0.253</td>
</tr>
<tr>
<td>Sex (Male count (%))</td>
<td>83 (40)</td>
<td>84 (63)</td>
<td>61 (73)</td>
<td>67 (72)</td>
<td>0.001</td>
</tr>
<tr>
<td>Born ‘here’ (count (%))</td>
<td>171 (85)</td>
<td>107 (82)</td>
<td>65 (78)</td>
<td>74 (82)</td>
<td>0.873</td>
</tr>
</tbody>
</table>

Values are mean ± SD unless specified. Age is participant’s age at recruitment (CON) or age at injury (TEN). Born ‘here’ is the number and proportion of participants born in Australasia (AUS) or Southern Africa (SA). $P$-values are for the difference between countries and between diagnostic groups respectively, adjusted for each other.
The primary finding of this study is that there was no association of the selected genetic variants within the COMP and THBS2 candidate genes with Achilles tendinopathy in the two populations studied. This study was only sufficiently powered to detect large effects on risk of developing Achilles tendinopathy (odds ratio (OR) >2) (Gauderman, 2002), and was therefore unlikely to detect any risk variant with a small effect. However, it should be noted that there was no tendency towards association with Achilles tendinopathy for any of the variants investigated. Although we cannot exclude the possibility that other variants within these genes are associated with Achilles tendinopathy, all the SNPs investigated in this study were identified as Tag SNPs and it is, therefore, unlikely that any SNPs tightly linked to these would show a significant effect on risk of Achilles tendinopathy. Both THBS2 (38,3kb) and COMP (8,5kb) are small genes and the Tag SNPs selected span 11,6kb and 9,9kb respectively, therefore giving moderate coverage of both genes. However, it still remains possible that other rare or common sequence variants within COMP and THBS2 may be associated with the risk of Achilles tendinopathy.
Previous work in this area has shown that genetic variants in genes encoding structural components of the collagen fibril often have large effects on risk of developing Achilles tendinopathy (OR>2) (Mokone et al., 2005; Mokone et al., 2006; Raleigh et al., 2009). Furthermore, rare mutations within genes encoding these structural components cause a range of severe disorders of connective tissue including pseudoachondroplasia, multiple epiphyseal dysplasia, osteogenesis imperfecta and Ehlers-Danlos syndrome (Hecht et al., 2005; Ikegawa et al., 1998; Malfait et al., 2005; Michalickova, Susic, Willing, Wenstrup, & Cole, 1998; Pollitt et al., 2006; Smith, Schwarze, Goldstein, & Byers, 1997). Based on these observations it has been suggested that there is limited redundancy within the biological mechanisms leading to collagen fibrillogenesis, and that common variants in the associated genes may have large effects on the risk of developing mild connective tissue pathology, such as Achilles tendinopathy (Ribbans & Collins, 2013).

The thrombospondins are calcium binding glycoproteins that can be divided into two subfamilies with distinct biological roles (Adams & Lawler, 2004; Bornstein et al., 2004). However, the carboxy terminal signature domain of all thrombospondins, which contains the EGF-like repeats, type III calcium binding repeats and a carboxy terminal lectin-like globule, is highly conserved between the five members of this protein family (Adams & Lawler, 2004; Carlson et al., 2008). Thrombospondin-1 and thrombospondin-2 are matricellular proteins that form homotrimers and modulate cell functions and cell–matrix interactions (Bornstein et al., 2004), while thrombospondin-3,-4 and -5 are structural proteins of the extracellular matrix and form homopentamers (Adams & Lawler, 2004; Bornstein et al., 2000; Bornstein et al., 2004). Investigation of thrombospondin-2 null mice has implicated this protein in the regulation of collagen fibrillogenesis and fibril diameter, possibly through its regulation of ambient matrix metalloproteinase-2 levels (Bornstein et al., 2000; Bornstein et al., 2004; Kyriakides et al., 1998). Cartilage oligomeric matrix protein is expressed in bovine and equine tendons, particularly in growing tendons and in response to mechanical load (DiCesare, Hauser, Lehman, Pasumarti, & Paulsson, 1994; Smith, Birch, Goodman, Heinegard, & Goodship, 2002; Smith, Zunino, Webbon, & Heinegard, 1997). Recent investigation has identified a mechano-responsive element in the 3kb proximal promoter of the human COMP gene (Amanatullah et al., 2012). In numerous studies of equine tendons, Smith et al. (Smith, Birch, et al., 2002; Smith, Gerard, et al., 2002; Smith, Zunino et al., 1997) have shown that cartilage oligomeric matrix protein levels are correlated to ultimate tensile stress and stiffness of tendon tissue and suggest that the appropriate expression of cartilage oligomeric matrix protein is necessary for the formation of a functional ECM.
and the structural integrity of tendons. Additionally, serum cartilage oligomeric matrix protein has been inversely associated with joint hypermobility (Chen et al., 2008), a risk factor for Achilles tendinopathy (Collins, Mokone, September, van der Merwe, & Schwellnus, 2009). There is also evidence to suggest that variations within the signature domain of the COMP gene alter protein structure and folding, and disrupt calcium binding (Chen, Deere, Hecht, & Lawler, 2000; Hecht et al., 2005; Hou, Putkey, & Hecht, 2000; Kvansakul, Adams, & Hohenester, 2004; Maddox, Mokashi, Keene, & Bachinger, 2000). It is therefore possible that sequence variants that have even a minor effect on transcription, translation or protein structure and function will result in at least a mild inherited clinical dysplasia. Taking into account the high sequence conservation of both the COMP and THBS2 genes between species, it remains possible that rare sequence variants (minor allele frequency 0.1–3.0%) within these genes might be associated with a large effect on the risk of Achilles tendinopathy (Bodmer & Bonilla, 2008; Ovcharenko, Nobrega, Loots, & Stubbs, 2004). Such rare variants often have a larger effect on risk as well as greater penetrance, making them good targets for prophylactic treatments (Bodmer & Bonilla, 2008). For example, the rare TT genotype (<5% in Caucasian populations) of the functional Sp1-binding site polymorphism (rs1800012; G>T) within intron 1 of the COL1A1 gene is associated with reduced risk of acute soft tissue ruptures (Collins, Posthumus, & Schwellnus, 2009; Khoschnau et al., 2008; Posthumus et al., 2009). Associations of common multifactorial diseases with rare variants are unlikely to be found by either genome-wide association studies or case-control association studies such as the current study, and further work would therefore be necessary to test this hypothesis (Bodmer & Bonilla, 2008).

**Strengths and limitations**

A limitation to this study is that the tendinopathy participants were significantly heavier than the control participants at the time of recruitment. Although the tendinopathy participant’s weight at the time of injury was not determined, several participants reported an increase in weight after injury due to inactivity. In addition, this difference may be explained by the higher proportion of males in the tendinopathy groups as the difference was much less pronounced after adjusting for country, sex and age. Another limitation is that, due to a low sample size, this study was only sufficiently powered to detect large effects on risk of AT (OR>2.0). The strength of this study lays in the hypothesis-driven selection of the candidate genes and variants investigated, and the stringent selection of homogenous tendinopathy participants.

**Conclusions**

In conclusion, this study did not identify an association between the common sequence variants investigated and increased risk of Achilles tendinopathy.

**Acknowledgements**

This research was supported by funds received from the National Research Foundation (NRF), the Medical Research Council (MRC) of South Africa, The University of Cape Town and Discovery Health. We thank Dr G.G. Mokone for
recruiting the SA participants and Professor M.P. Schwellnus for the clinical diagnoses of the SA participants.
References


abnormalities that are associated with disordered collagen fibrillogenesis, an increased vascular density, and a bleeding diathesis. *Journal of Cell Biology, 140*, 419–430.


Posthumus, M., September, A.V., Keegan, M., O’Cuinneagain, D., van der Merwe, W., Schwellnus, M.P. et al. (2009). Genetic risk factors for anterior cruciate...


http://repository.uwc.ac.za


