This work has been submitted to NECTAR, the Northampton Electronic Collection of Theses and Research.

**Article**

**Title:** ELN and FBN2 gene variants as risk factors for two sports-related musculoskeletal injuries

**Creators:** El Khoury, L., Posthumus, M., Collins, M., van der Merwe, W., Handley, C. J., Cook, J. and Raleigh, S. M.

**DOI:** 10.1055/s-0034-1390492


It is advisable to refer to the publisher's version if you intend to cite from this work.

**Version:** Accepted version


**Note:** © 2016 Georg Thieme Verlag KG

http://nectar.northampton.ac.uk/8000/
**Abstract**

The ELN and FBN2 proteins are important in extracellular matrix function. The *ELN* rs2071307 and *FBN2* rs331079 gene variants have been associated with soft tissue pathologies. We aimed to determine whether these variants were predisposing factors for both Achilles tendinopathy (AT) and anterior cruciate ligament (ACL) ruptures.

For the AT study, 135 cases (TEN group) and 239 asymptomatic controls were recruited. For the ACL rupture study our cohort consisted of 141 cases (ACL group) and 219 controls. Samples were genotyped for both the *ELN* rs2071307 and *FBN2* rs331079 variants using TaqMan assays. Analysis of variance and chi-squared tests were used to determine whether either variant was associated with AT or ACL rupture with significance set at p<0.05.

The GG genotype of the *FBN2* variant was significantly over-represented within the TEN group (p=0.035; OR=1.83; 95% CI 1.04–3.25) compared to the CON group. We also found that the frequency of the G allele was significantly different between the TEN (p=0.017; OR=1.90; 95% CI 1.11–3.27) and ACL groups (p=0.047; OR=1.76; 95% CI 1.00–3.10) compared to controls. The *ELN* rs207137 variant was not associated with either AT or ACL rupture. In conclusion, DNA sequence variation within the *FBN2* gene is associated with both AT and ACL rupture.

**Keywords** Gene; Achilles Tendinopathy; Tendon rupture; Ligament rupture; Injury prevention.
Introduction

Injury to the Achilles tendon and the anterior cruciate ligament are severe traumas typically sustained during sports activities. Achilles tendon injuries, including chronic Achilles tendinopathy (AT) and acute Achilles tendon rupture, are prevalent within athletic populations [26]. Indeed, the lifetime incidence of AT is approximately 10% in the general population and as high as 50% within competitive runners [12]. Chronic AT may be due, in part, to excessive exposure of the Achilles tendon to acute or repetitive mechanical loading forces experienced during exercise [15]. Anterior cruciate ligament (ACL) ruptures have low lifetime prevalence in the general population but have been reported to be nearly 80% among netball players [7]. The most common mechanism of ACL rupture involves a sudden change in an athlete’s direction or rapid deceleration [19]. Both AT and ACL ruptures are complex multifactorial phenotypes with several intrinsic and extrinsic risk factors. However, the exact aetiology is not yet fully understood [3].

Among the intrinsic risk factors, several genetic sequence variants have been shown to increase the risk (predispose individuals) to AT and ACL ruptures. Variants within the TNC [3], MMP3 [29], GDF5 [23] and TIMP2 genes [5] are associated with the risk of AT. Furthermore, variants within the COL1A1 [22] and COL12A1 genes [24] have also been associated with ACL ruptures. Interestingly, a variant within the COL5A1 gene was found to be associated with both AT [16] and ACL ruptures [25]. These findings show that both chronic AT and ACL ruptures have a partial polygenic basis where complex interactions between genes and the environment are likely to exacerbate the risk of both types of injuries [3]. All the genes described above encode proteins with either a structural or regulatory role in maintaining the homeostasis of the soft tissue extracellular matrix (ECM). Therefore, it is fair to assume that other genes, which code for additional regulatory components of the ECM might also be candidates for AT and ACL rupture.
Elastin (ELN) is an insoluble polymer composed of several tropoelastin molecules covalently bound to each other by cross-links [31]. ELN proteins contribute to tendon and ligament elasticity by allowing them to stretch and return to their original state. These proteins have an important load-bearing role in musculoskeletal tissues and are expressed in places where mechanical energy is stored [8]. The ELN rs2071307 gene variant has been shown to be associated with other multifactorial conditions of the extracellular matrix, such as aortic stenosis [6] and aortic aneurysm [33]. Interestingly, the ELN rs2071307 variant is located within exon 20 of the gene and is a non-synonymous SNP. It is predicted to be deleterious (Queen’s University. [http://compbio.cs.queensu.ca/F-SNP/]) since it substitutes a hydrophobic amino acid glycine with a hydrophilic serine residue (National Center for Biotechnology Information. [http://www.ncbi.nlm.nih.gov/projects/SNP/]). This substitution may disrupt the integrity of the microfibrils rendering them more prone to damage [18] and therefore this variant may predispose to soft tissue damage during sports performance.

Fibrillins are large glycoproteins present in the extracellular matrix of tendons and ligaments [2]. Both fibrillin-1 (FBN-1) and fibrillin-2 (FBN-2) share high amino acid homology and are involved in providing strength and flexibility to various soft tissues. FBN-2 is abundant in elastic tissues, such as tendons and ligaments [35] where it plays an important role in the assembly of elastic fibres [2]. Mutations within the FBN2 gene are known to associate with musculoskeletal pathologies such as congenital contractual arachnodactyly [9]. Furthermore, the rs331079 variant located within intron 7 of the gene (University of Florida. [www.snppep.chip.org]) has previously been associated with intracranial aneurysms [32].

As both the FBN2 rs331079 and the ELN rs2071307 variants associate with other conditions related to the extracellular matrix we considered them as possible risk determinants for both AT and ACL rupture. Accordingly, the aim of this study was to test that hypothesis.
Material and Methods

One hundred and thirty five (60 Australian (AUS) and 75 South African (SA)) Caucasian participants with Achilles tendinopathy (TEN group) were recruited to this study from the Musculoskeletal Research Centre at La Trobe University in Melbourne, and from the Medical Practice at the Sports Science Institute of South Africa. Furthermore, 239 (143 AUS and 96 SA) asymptomatic Caucasian controls (CON groups) were recruited to this study from recreational sports clubs within the Melbourne area in Australia, and within the Cape Town area in South Africa. Chronic AT was clinically diagnosed as described by Mokone et al.[17] in the first manuscript describing the South African AT cohort. The Australian cohort used the same clinical diagnosis described by Mokone et al. In addition, diagnosis was confirmed with soft tissue ultrasound examination in all the AUS and 40 of 75 SA participants. In addition, 141 South African Caucasian participants with surgically diagnosed ACL ruptures (ACL group) and 219 apparently healthy (CON group), unrelated, physically active, gender matched South African Caucasian participants without any self-reported history of ligament or tendon injury were recruited for this study as previously described [22]. Seventy four participants sustained the injury through a non-contact mechanism and were analysed as a separate subgroup (NON subgroup).

Previous injury data was used as inclusion criteria in the various cohorts analysed. In the AUS Achilles cohort, the CON group had no history of any tendon injury, whereas in the SA Achilles cohort, the CON group merely had no previous history of Achilles tendon injuries. In the case of ACL rupture, the first ACL rupture was documented as the specific inclusion injury. Therefore, by definition, no participant in the ACL group had a previous ACL rupture

None of the participants included in this study had symptoms or signs of Ehlers-Danlos syndrome (EDS), hypermobility or benign hypermobility joint syndrome or other monogenic connective tissue disorders when their medical examinations were reviewed by the medical practitioner [16,17,34].
Physical activity data was recorded for the South African Achilles tendinopathy cohort (SA CON and SA TEN), but not for the Australian Achilles tendinopathy cohort (AUS CON and AUS TEN). In addition physical activity data was also recorded for the South African ACL cohort (SA ACL and SA CON). The data recorded for the SA CON and SA TEN groups included total years participated in running and high impact sports, as well as hours per week of participation in the last 2 years. The data reported for the ACL cohort included years of participation in contact sports, non-contact jumping sports, non-contact non-jumping sports and skiing sports. Data were collected as previously described [16, 25].

Based on our earlier work, this study had a large enough sample size to detect associations with an OR of 2.0 at p<0.05 with 80% power [28]. All participants gave informed written consent, in accordance with the journal’s recommendations [10,11], and all completed a medical and injury history questionnaire. Ethical approval was obtained from the Research Ethics Committees at the University of Cape Town, South Africa, La Trobe University, Australia, Monash University, Australia and the University of Northampton, United Kingdom prior to initiating this work.

For the Australian cohort, DNA was extracted from whole blood using Qiagen DNA extraction kits (Flexigene DNA kit, Qiagen P/L, Valencia, California, USA) as per the manufacturer’s recommendations. DNA from the South African individuals was extracted from blood using the method described by Lahiri and Nurnberg [14] and modified by Mokone et. al. [16,17]. Upon extraction, DNA was frozen at -20 °C for long-term storage, and smaller aliquots were stored at 4 °C for short term usage.

DNA from all participants was genotyped for the FBN2 rs331079 and ELN rs2071307 gene variants using fluorescence-based TaqMan assays (Applied Biosystems, Foster City, CA, USA). PCR reactions contained allele-specific probes and primers in a PCR mastermix containing AmpliTaq DNA Polymerase Gold (Applied Biosystems, Foster City, CA, USA) in a total reaction volume of 12 μL. PCR was performed on an Applied Biosystems
StepOnePlus™ real-time PCR system (Applied Biosystems, Foster City, CA, USA). Genotypes were called according to output clustering profiles using Applied Biosystems StepOnePlus™ real-time PCR software Version 2.1 (Applied Biosystems, Foster City, CA, USA). Rox was used as a passive reference to normalise fluorescence signal intensity relative to the amount of sample used.

The statistical power of the study was determined using Quanto v1.2 (http://hydra.usc.edu/gxe). The initial calculations were done using a recessive model and a disease population prevalence of 10%. Assuming a risk allele frequency of 60%, a matched case-control population of 136 individuals per group was adequate to detect an allelic OR of 2.0 at a power of 80% and a significance level of 5%.

Data were analysed using SPSS Version 20 (SPSS Science Inc, Chicago, Ill, USA) statistical program. A one-way analysis of variance was used to establish if any significant difference existed between the characteristics of the TEN and CON groups within the Australian and South African cohorts as well as between the ACL rupture and CON groups. A chi-squared ($\chi^2$) analysis or Fisher's exact test was used to determine if significance differences existed between genotype and/or allele frequencies, as well as other categorical data between the groups. In all analysis significance was accepted when $p<0.05$. Adjustments for multiple testing were not conducted as it has been previously described [21] that no appropriate method exists. Furthermore, the Bonferroni adjustment was considered too conservative [21] and inappropriate for a situation like this where there is prior evidence that the gene of interest is associated with a trait [20]. Hardy-Weinberg equilibrium was determined using the program Genepop web version 3.4 (Curtin University. http://genepop.curtin.edu.au/).
Results

Running was the predominant sporting activity resulting in Achilles tendon injuries (63.1%, N=65) in the SA cohort. The SA groups were matched for the mean number of years participating in running (CON, 8.7 ± 8.2 yrs, n=95; TEN, 10.0 ± 11.0 yrs, n=62; p=0.402). However, there was a significant difference in hours of training between the two groups (CON, 3.6 ± 3.0 hrs/week, n=91; TEN, 2.4 ± 2.7 hrs/week, n=55; p=0.011), where the SA CON group trained for more hours per week. The SA TEN participants participated in more years of high impact sports compared to the SA CON group in the past (CON, 9.4 ± 8.4 yrs, n=95; TEN, 13.1 ± 11.1 yrs, n=62; p=0.018), however, the SA CON group performed a greater amount of high impact sports during the last 2 years (CON, 3.6 ± 3.1 yrs, n=95; TEN, 2.5 ± 12.9 yrs, N=62; p=0.029). Although all AUS participants were physically active individuals, the type of sporting activity involved in, the hours of training and the frequency of activity were not recorded.

The SA ACL and SA CON groups were matched for years of participation in contact sports (SA CON, 11.7 ± 7.1 yrs, n=219; SA ACL, 11.5 ± 8.0 yrs, n=141; p=0.892), non-contact jumping sports (SA CON, 27.8 ± 19.9 yrs, n=190; SA ACL, 25.7 ± 22.6 yrs, n=141; p=0.398), non-contact non-jumping sports (SA CON, 11.5 ± 7.1 yrs, n=219; SA ACL, 10.5 ± 8.5 yrs, n=141; p=0.575), and skiing sports (SA CON, 19.1 ± 16.9 yrs, n=219; SA ACL, 8.6 ± 8.5 yrs, n=141; p=0.094).

Since the ELN rs2071307 and FBN2 rs331079 allele and genotype frequencies in both of the South African (SA) and Australian (AUS) TEN and CON groups were similar (Supplementary table 1), the data was collectively analysed. The CON and TEN groups were similarly matched for age and gender (Table 1). When co-varied for sex, the two groups were similarly matched for height. Furthermore, when co-varied for sex and age at recruitment, the TEN group was found to be significantly heavier (p<0.001) with larger BMIs (p<0.001) (Table 1). The TEN group was recruited on average 5.1 years after the initial injury.
Participants in the AUS TEN group carrying the *ELN* rs2071307 AA (53.1 ± 11.6, n=10) genotype were significantly (p=0.005) older when they reported their initial Achilles tendon injury when compared to those with a GG (37.2 ± 12.6, n=16) or GA (37.8 ± 13.6, n=32) genotype. There were, however, no significant differences in the mean ages of the three genotype groups in the CON AUS group (GG: 40.7 ± 11.8, n=48; GA: 37.4 ± 12.2, n=68; AA: 40.1 ± 12.1, n=24; p=0.323). There were no other significant genotype effects of either variants with respect to height, weight, BMI, or sex in the AT group (data not shown). Furthermore, the investigated variants did not show any interaction with age, height, weight, BMI and sex in the ACL population (data not shown).

The genotype frequency distributions of the *FBN2* rs331079 and the *ELN* rs2071307 variants within the AT and the ACL rupture groups are shown in table 2. In the combined TEN cohort, the *FBN2* rs331079 genotype frequency was significantly different (p=0.035) between the CON (GG, 76.9%; GC + CC, 23.1%) and TEN (GG, 85.9%; GC + CC, 14.1%) groups (Table 2). The GG genotype was significantly over-represented within the TEN group (p=0.035; OR=1.83; 95% CI 1.04 – 3.25). We also found a significant (p=0.017; OR=1.90; 95% CI 1.11 – 3.27) allele frequency distribution difference for the *FBN2* rs331079 variant between the CON (G, 87.4%; C, 12.6%) and TEN (G, 93.0%; C, 7.0%) groups (Table 2). Similarly, we also found a significant (p=0.047; OR=1.76; 95% CI 1.00 – 3.10) allele frequency distribution difference of the rs331079 locus between the CON (G, 89.3%; C, 10.7) and ACL (G, 93.6%; C, 6.4%) groups. Also, in the AT population, there were no significant *ELN* rs2071307 genotype (p=0.795) or allelic (p=0.741) frequency differences between the CON and TEN groups (Table 2).

Although not significant, we found a tendency towards an allelic (p=0.064) association for the *ELN* rs2071307 variant and a tendency towards a genotypic (p=0.075; p=0.112) association between the CON and ACL groups for the *FBN2* rs331079 and *ELN* rs2071307 variants respectively. There were no genotypic or allelic associations between the CON and NON
Furthermore, these gene variants did not show any significant distribution difference when participants were grouped into genders (data not shown).

Discussion

We have shown that the FBN2 rs331079 variant is significantly associated with the risk of both AT and ACL rupture. Specifically, the GG genotype was over-represented in participants with chronic AT and the G allele was over-represented in both pathologies. Therefore, it appears that individuals carrying the G allele or the GG genotype are approximately twice as likely to develop either of the two injuries. Interestingly this same variant has recently been shown to associate with intracranial aneurysms in a Dutch population [32]. However, in the Dutch study it was the C allele that was found to be the risk factor as opposed to the G allele. It is noteworthy that FBN2 mRNA levels have been shown to be elevated in rat Achilles tendon undergoing repair with expression of FBN2 reported to be increased for ten days post injury [13]. Similarly, an increase in the expression of FBN2 has been found in other pathologies such as mitral valve prolapse [27].

ELN and FBN-2 are known to form a network of microfibrils that maintains the tendon architecture [31]. An increase in FBN-2 levels might be expected to increase the density of the tendon and lead to an increase in tendon stiffness and rigidity possibly affecting the compliance of the tendon to muscle movement [4]. On the other hand, a decrease in FBN-2 levels could result in weaker tendons caused by structural deficiencies in the microfibril network [30]. Impairment of the function of FBN-2 is believed to be a major determinant of microfibrillopathy [30] which is speculated to precede a tendinopathy. Furthermore, the increase in FBN2 expression levels observed during tendon repair [13] is consistent with an important role for FBN-2 in maintaining the tendon's architectural integrity.

Mutations such as the G3532T and G3590A substitutions have been found within the FBN2 gene that lead to the development of connective tissue disorders such as congenital contractural arachnodactyly [9]. The rs331079 variant that we investigated in this study
resides within an intronic region of the FBN2 gene (University of Florida. www.snpper.chip.org). Although intronic variants do not determine the primary sequence of a protein molecule [1], they may have other, hitherto, undiscovered roles that are necessary for appropriate expression of protein molecules. However, at present the functionality of this variant has not been described and therefore we do not know why it predisposes individuals to AT and ACL rupture. The rs331079 variant is known to be part of a linkage block in Caucasians and is in high linkage disequilibrium (D'=1) with the FBN2 rs331081, rs331082, and rs331085 variants (Wellcome Trust Sanger Institute. www.ensembl.com). All three of these additional variants are also located within intron 7 of the FBN2 gene (University of Florida. www.snpper.chip.org). The linkage disequilibrium between the rs331079 variant that we investigated and rs331081, rs331082, and rs331085 means that it is conceivable that one of these linked variants may also have a role in predisposing to AT or ACL.

Our data do not support an association between the ELN rs2071307 variant and either AT or ACL ruptures. It is interesting to note however, that although we found no relationship between this variant and either pathology; the rs2071307 SNP is a non-synonymous and possibly deleterious polymorphism (Queen’s University. http://compbio.cs.queensu.ca/F-SNP/) which results in a change of amino acid from hydrophobic glycine to hydrophilic serine (University of Florida. www.snpper.chip.org). It is possible of course, that other variants within this gene may be associated with either AT or ACL ruptures.

Although our study found a significant association between the FBN2 rs331079 G allele and the risk of AT and ACL rupture, the work has some limitations. Firstly, although our SA cohorts (both TEN and ACL rupture groups) were matched for some aspects of physical activity there were some differences in training behaviour and previous exposure to high impact sports for the TEN cohort. Secondly, we did not have detailed information on sports history for the Australian cohort. Levels of physical activity should be accurately documented in future studies. Furthermore, although the study was sufficiently powered to detect associations with relatively large effects it should be repeated in bigger cohorts. Likewise,
additional association studies should be carried out in populations of different ethnicities showing different minor allele frequencies for the rs331079 (African, 3%; European, 10%; ad-mixed American, 28%; East Asian, 7%) and the rs2071307 (African, 26%; European, 39%; ad-mixed American, 30%; East Asian, 14%) variants (1000 Genomes Project, www.1000genomes.org).

Finally, the findings from this study advance our understanding of the polygenic basis of musculoskeletal injuries. We suggest that the FBN2 rs331079 variant should be considered as an additional genetic locus to include in an injury risk assessment model that might be used to identify athletes who are predisposed to AT and ACL ruptures.
References


List of legends

Table 1: General characteristics of the Achilles tendinopathy group (TEN), the anterior cruciate ligament rupture group (ACL), and the ACL subgroup with the non-contact (NON) mechanism of injury as well as their respective control groups.

Table 2: The genotype and allele frequency distribution of the two selected candidate variants within the Achilles tendinopathy (TEN), ACL ruptures (ACL) and their respective asymptomatic control (CON) groups.