Functional polymorphisms within the inflammatory pathway regulate expression of extracellular matrix components in a genetic risk dependent model for anterior cruciate ligament injuries

Accepted in Journal of Science and Medicine in Sports

Mathijs A.M. Suijkerbuijk
Marco Ponzetti
Masouda Rahim
Michale Posthumus
Charlotte K. Häger
Evalena Stattin
Kjell G. Nilsson
Anna M. Teti
Duncan E. Meuffels
Bram J.C. van der Eerden
Malcolm Collins
Alison V. September

ABSTRACT

Objectives: To investigate the functional effect of genetic polymorphisms of the inflammatory pathway on structural extracellular matrix components (ECM) and the susceptibility to an anterior cruciate ligament (ACL) injury.

Design: Laboratory study, case-control study.

Methods: Eight healthy participants were genotyped for interleukin (IL)*1B* rs16944 C>T and *IL6* rs1800795 G>C and classified into genetic risk profile groups. Differences in type I collagen (*COL1A1*), type V collagen (*COL5A1*), biglycan (*BGN*) and decorin (*DCN*) gene expression were measured in fibroblasts either unstimulated or following IL-1β, IL-6 or tumor necrosis factor (TNF)-α treatment.

Moreover, a genetic association study was conducted in: (i) a Swedish cohort comprised of 116 asymptomatic controls (CON) and 79 ACL ruptures and (ii) a South African cohort of 100 CONs and 98 ACLs. Participants were genotyped for *COL5A1* rs12722 C>T, *IL1B* rs16944 C>T, *IL6* rs1800795 G>C and *IL6R* rs2228145 G>C.

Results: *IL1B* high-risk fibroblasts had decreased *BGN* (p=0.020) and *COL5A1* (p=0.012) levels after IL-1β stimulation and expressed less *COL5A1* (p=0.042) following TNF-α treatment. Similarly, unstimulated *IL6* high-risk fibroblasts had lower *COL5A1* (p=0.012) levels than *IL6* low-risk fibroblasts.

In the genetic association study, the *COL5A1-IL1B-IL6* T-C-G (p=0.034, Haplo-score: 2.1) and the *COL5A1-IL1B-IL6R* T-C-A (p=0.044, Haplo-score: 2.0) combinations were associated with an increased susceptibility to ACL injury in the Swedish cohort when only male participants were evaluated.

Conclusions: This study shows that polymorphisms within genes of the inflammatory pathway modulate the expression of structural and fibril-associated ECM components in a genetic risk dependent manner, contributing to an increased susceptibility to ACL injuries.

Key words: Anterior cruciate ligament injury; Extracellular Matrix; Genetics; Polymorphisms; Personalized Medicine.

INTRODUCTION

Anterior cruciate ligament (ACL) rupture is a common sports-related injury of the knee[1]. The ability of the ACL to maintain its extracellular matrix (ECM) integrity is critical to its function to effectively resist mechanical loads and prevent injury[2]. Loading activates matrix-remodeling pathways to maintain ECM homeostasis, such as the inflammatory pathways (Figure 1). Therefore, it is not surprising that polymorphisms within these pathways contribute to the susceptibility of ACL injuries[3, 4].

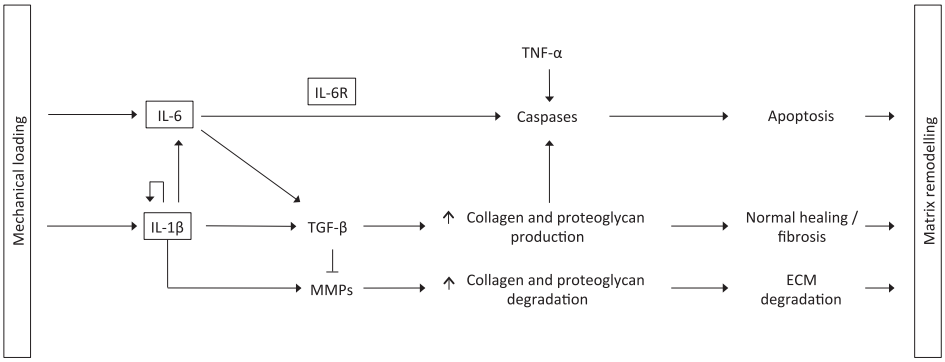


Figure 1. Schematic representation of the proposed downstream effects of cytokines IL-1β and IL-6 which are upregulated in response to mechanical loading of a ligament [3, 4, 28, 32].
Activation/upregulation is represented by a pointed arrow head (→) and inhibition/down regulation is represented by a perpendicular line at the end (---|). The boxed molecules are the ones investigated in the current study. Abbreviations: ECM, extracellular matrix; IL-1β, interleukin-1β; IL-6, interleukin-6; IL-6R, interleukin-6 receptor; MMPs, matrix metalloproteinases; ROS, reactive oxygen species; TGF-β, transforming growth factor β; TNF-α, tumor necrosis factor.

Type V collagen is a functionally important collagen for the maintenance of tissue structure and integrity. The major isoforms consists of two α1 (V) and one α2 (V) chains encoded by *COL5A1* and *COL5A2* respectively[5]. Polymorphisms within the 3'UTR of *COL5A1* were previously implicated in ACL rupture[6] and tendinopathy[7]. In addition, polymorphisms within genes encoding the α1(I) chain of type I collagen (*COL1A1*)[8], biglycan (*BGN*)[9] and decorin (*DCN*)[9] were associated with ACL injury susceptibility. Together, these molecules form the basic building blocks of the ECM and are involved in collagen fibrillogenesis.

Interleukin (IL) -1β is a pro-inflammatory cytokine encoded by *IL1B* and up-regulates the production of matrix metalloproteinases, regulating the degradation of specific ECM components, such as collagen types V and X[10]. In addition, IL-1β induces its own expression and the expression of other pro-inflammatory cytokines such IL-6 (**Figure 1**)[11]. The C-allele of the *IL1B* promoter polymorphism rs16944 C>T increases IL-1β

mRNA expression levels[12] and is hypothesised to increase susceptibility to tendinopathy and ACL injury[3, 4].

IL-6 is known to induce apoptotic cell death[15] affecting the production of extracellular matrix components and thereby the ECM integrity. Polymorphisms within the *IL6* gene that increase IL-6 expression, such as the G-allele of the *IL6* rs1800795 G>C polymorphism, can therefore potentially be associated with increased risk of ligament injuries. IL-6 needs to bind and form complexes with the interleukin-6 receptor (IL-6R) in order to exert its biological function. IL-6R exists as two isoforms: a membrane-bound receptor and a soluble receptor. *IL6R* rs2228145 A>C is located in the cleavage site and is thought to affect cleavage efficiency. The A-allele is associated with decreased levels of soluble IL-6R and an increased response to IL-6[13]. Therefore, we hypothesize that the *IL6R* rs2228145 AA genotype is associated with an increased susceptibility to ligament injury.

Although currently no genetic loci within the gene encoding tumor necrosis factor α (TNF-α) have been associated with either ACL injuries or tendinopathies, this pro-inflammatory protein is considered to be key in the inflammatory pathway[14]. The biological function of TNF-α is executed after binding to its receptor, the tumor necrosis factor receptor superfamily member 1A (TNFRSF1A). Similar to IL-6, TNF-α is involved in apoptosis and thereby possibly contributes to matrix remodeling capacity[14].

The main aim of the current study was to investigate the effects of specific genetic loci within the inflammatory pathway on the production of ECM components in an injury risk model. Additionally, the association of these genetic loci with susceptibility to ACL injury was evaluated in two independent populations of different ancestry. Based on the *a priori hypothesis* it was proposed that the *IL1B* rs16944 CC and the *IL6* rs1800795 GG downregulate the production of ECM components and should therefore be associated with an increased susceptibility to ligament injuries.

METHODS

All participants completed questionnaires regarding personal details, medical history, sporting history and a family history of tendon and ligament injury. Written informed consent was obtained from all participants according to the Declaration of Helsinki. Ethics approval was attained from the Human Research Ethics Committee (HREC) of Faculty of Health Sciences, University of Cape Town, South Africa (HREC 164/2006 and 645/2014) and the Regional Ethical Review Board in Umeå, Sweden (dnr. 2011-200-31M), where relevant.

For the in vitro work eight healthy, unrelated South African participants of self-reported Caucasian ancestry with no history of musculoskeletal soft tissue injuries were recruited. Venous blood and skin biopsies were taken from each participant.

For the Swedish cohort 195 physically active and unrelated participants (age 19-65 years) were recruited between 2011 and 2013 from either the Västerbotten or Norrbotten regions of Sweden, via the orthopedic clinics in two major hospitals in the cities of Umeå: Västerbotten and Luleå: Norrbotten. The majority of the participants were recruited from a long-term follow-up of ACL injury[15]. This cohort consisted of 79 participants with ACL rupture (SWE-ACL) and 116 asymptomatic participants without any history of ACL or tendon injury (SWE-CON). ACL ruptures were diagnosed based on physical examination, magnetic resonance imaging and arthroscopically confirmed at the University hospital in Umeå. Mechanism of injury data was categorized into direct contact, indirect contact, non-contact and skiing sports as previously defined[16]. All 79 cases reported a non-contact mechanism (SWE-NON) of injury.

For the South African cohort 198 physical active and unrelated participants were recruited from South Africa as previously described[17]. This cohort comprised of 100 asymptomatic controls (SA CON) and 98 participants with an ACL rupture (SA ACL) of which 51 reported a non-contact mechanism of injury.

A previously described protocol with slight modifications[18] was used to extract genomic DNA from venous blood. Participants participating in the *in vitro* study were genotyped for the *IL1B* rs16944 C>T and *IL6* rs1800795 G>C polymorphisms. Restriction fragment length polymorphism (RFLP) analysis was used for *IL1B* rs16944 (*AvaI*)[4] while custom designed fluorescence-based Taqman PCR assays (Applied Biosystems, Foster City, CA, USA) were used to genotype *IL6* rs1800795. Several controls were included in the genotyping protocols which included repeated controls and negative controls. Two independent researchers scored genotype results obtained from RFLP analysis. If no consensus was reached, or genotyping was not possible, samples were re-analysed. TaqMan PCR determines genotype calls automatically, however were manually checked by a researcher. In general, all samples were analysed only once. Based on their genotypes, participants were either classified in the high-risk or low-risk profile group. More specifically, the *IL1B* TT and CT genotypes were considered as low-risk, whereas the CC genotype was considered as high-risk. Additionally, the *IL6* CC and GC were classified in the low-risk group, and the GG-genotype was classified in the high-risk group.

Participants for the genetic association study were additionally screened for *COL5A1* rs12722 C>T and *IL6R* rs2228145 A>C. Genotyping of all four single nucleotide polymorphisms (SNPs) was conducted in the current study on all samples in both the Swedish (n=195) and the South African cohort (n=198). It should be noted that the DNA samples of the South African cohort were previously collected[17]. Restriction fragment length polymorphism (RFLP) analysis was used for the *COL5A1* rs12722 (*BstuI)*, and

IL6R rs2228145 (*HindIII*) SNPs. All SNPs: *COL5A1* rs12722 C>T, *IL1B* rs16944 C>T, *IL6* rs1800795 G>C and *IL6R* rs2228145 A>C were selected based on their previously reported genetic associations with risk of ACL ruptures and Achilles tendinopathy[19, 20].

To establish primary fibroblast cultures, skin biopsies were processed according to a modified Baumgarten protocol[21]. Human dermal fibroblasts were cultured to 70% confluency in Dulbecco's modified Eagle medium (DMEM, Gibco, Carlsbad, CA, USA) with 200 units/mL penicillin, 100 µg/mL streptomycin, 3.97 mM GlutaMAX (Gibco) and 10% FBS. Cells were serum-starved for 8h in DMEM and subsequently treated with 10ng/ml human recombinant (hr) IL-6[22], 20ng/ml hrIL-1β[23] or 10ng/ml hrTNF-α[22] (all from Peprotech, Rocky Hills, NJ, USA). After 24h, cells were washed twice in ice-cold Dulbecco's phosphate-buffered saline (DPBS, Sigma-Aldrich, Saint Louis, MO, USA) and frozen at -80°C until ready for RNA extraction, using the RNAeasy kit (Qiagen, Venlo, The Netherlands). Subsequently, a cDNA synthesis kit including a recombinant RNAse inhibitor (Thermo Scientific) using oligo (dT)s as primers was used.

A SYBR green-based buffer (Thermo Scientific), 10ng of cDNA, and primers specific for the transcript of interest, to a final concentration of 500nM each were mixed. PCR cycles were as follows: (2'30" at 50°C: 2'30" at 95°C) x1, (15" at 95°C: 30" at 60°C) x50 followed by melt-curve analysis (95°C-60°C-95°C). RT-PCR analyses were performed using a Quantstudio3 real-time PCR machine (Thermo Scientific). The mRNA expression levels of structural matrix components, such as *COL5A1, COL1A1, DCN* and *BGN* were assessed for each sample including components of the inflammatory pathway, namely *IL1B, IL6R1, IL6, IL6R* and *TNFRSF1A* (Invitrogen, Carlsbad, CA, USA) (Supplementary Table 1). Cofilin (Invitrogen) was previously found stable and linearly correlated with RNA quantity (data not shown), and was therefore used to normalize qPCR data. Both positive and negative controls were always included.

Statistical analyses were performed with the programming environment R (R Development Core Team). In the cytokine stimulation experiments, statistics were performed using Unpaired, two-tailed Student's t-test. Power analysis was performed using QUANTO v.1.2.4 (http://biostats.usc.edu/software) to calculate sample size for the Swedish cohort. Assuming minor allele frequencies between 0.1 and 0.5 a sample size of 79 cases would be adequate to detect an allelic odds ratio (OR) of 2.3 and greater at a power of 80%. Basic descriptive statistics were compared using the one-way analysis of variance to detect significant differences between characteristics of the SWE-CON group and the SWE-NON group. The R package *genetics*[24] and *SNPassoc*[25] were used to analyse differences in genotype and allele frequencies between the groups and to calculate Hardy-Weinberg equilibrium probabilities. Inferred allele constructs were created for *COL5A1-IL1B-IL6-IL6R* genes from both the Swedish and South African genotype data respectively using the *haplo.stats* package in R[26]. The analysed models were based on previously reported associations[3, 4].

RESULTS

Fibroblasts derived from 8 donors had a high-risk genetic profile for either *IL1B* rs16944 C>T or *IL6* rs1800795 G>C (Supplementary Table 2). No significant differences in basal expression were observed in any of the ECM genes when fibroblasts were classified based on *IL1B* genotypes (Figure 2A). A reduced (p=0.012) *COL5A1* expression was noted in *IL6* high-risk fibroblasts compared to low-risk fibroblasts. As for the cytokine-related genes, we found that *TNFRSF1A* was less (p=0.003) expressed in the untreated *IL6* high-risk fibroblasts (Supplementary Figure 1A, B).

In *IL1B* high-risk fibroblasts, *COL5A1* (p=0.012) and *BGN* (p=0.020) expression were reduced following hrIL-1β. Additionally, treatment with hrTNF-α resulted in decreased *COL5A1* (p=0.042) levels (Figure 2G). In untreated *IL6* high-risk fibroblasts, *COL5A1* was reduced (p=0.012) compared to *IL6* low-risk fibroblasts (Figure 2B). No stimulation of the fibroblasts with hrIL-6 (Supplementary Figure 1C, D), hrIL-1β(Supplementary Figure 1E, F) or hrTNF-α (Supplementary Figure 1G, H) did not significantly alter the expression of any of the cytokine-related genes analysed.

Polymorphisms within *IL1B* and *IL6* alter the expression of structural and fibril-associated ECM components and herewith possibly modulate the susceptibility of ligament injuries. Therefore, these associations were further investigated in other population groups from (i) Sweden and (ii) an indigenous mixed ancestry population from South Africa.

The South African population was previously described in detail[17]. Swedish participants were matched for height, body mass and body mass index (BMI) (Supplementary Table 3). However, participants in the SWE-CON group consisted of significantly less men (34.5%, n=40) than the SWE-NON group (54.4%, n=43, p=0.014) and were significantly older (44.7 ± 11.9, n=114) than participants in the SWE-NON group (36.5 ± 13.7, n=78, p<0.001). Differences in medical and family history are displayed in Supplementary Table 4. No significant genotype effects were noted on age, sex, height, body mass or body mass index for the investigated polymorphisms (Supplementary Table 5).

No significant differences in genotype or allele frequency distributions were observed for either *COL5A1* rs12722 C>T, *IL1B* rs16944 C>T and *IL6* rs1800795 G>C in both the South African and Swedish cohorts (Table 1). However, for the South African cohort the *IL6R* rs2228145 A>C CC genotype was significantly overrepresented (p=0.028) in the SA-CON group (13%, n=12) compared to the SA-ACL group (3%, n=3). Although not significant (p=0.054), a similar trend was observed when comparing the SA-CON group (13%, n=12) to the SA-NON subgroup (11%, n=6). Furthermore, the genotype and allele frequency distributions significantly differed between the South African and Swedish cohorts (Supplementary Table 6) for all the polymorphisms tested. Therefore, cohorts could not be combined for further analysis.

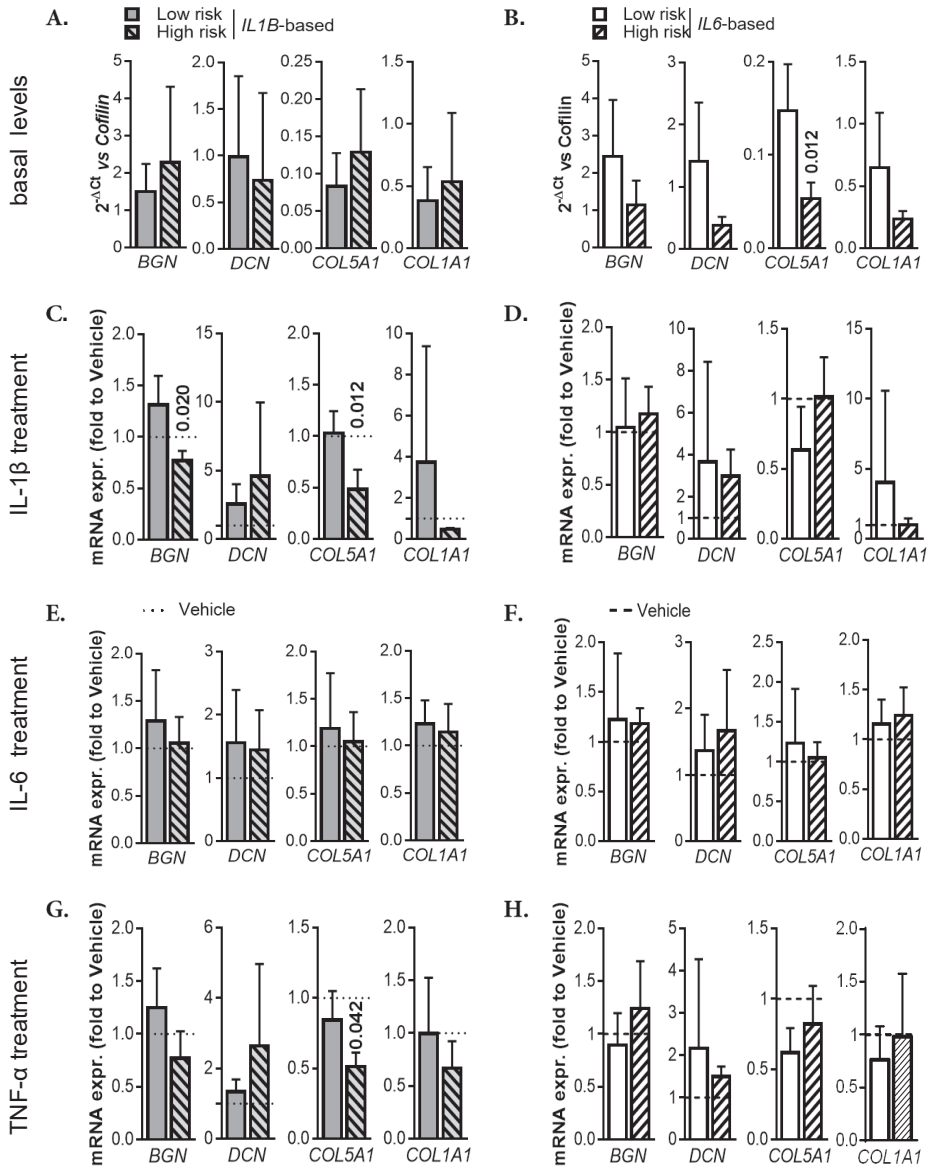


Figure 2: mRNA expression of extracellular matrix genes in stimulated and unstimulated fibroblasts.
*Primary human fibroblasts were obtained from 8 healthy volunteers and classified in high-risk or low-risk for ligament injuries based on IL1B rs12722 C>T (**A, C, E, G**) and IL6 rs1800795 G>C (**B, D, F, H**). Fibroblasts were treated with vehicle (PBS) to evaluate basal levels (**A, B**)or hr-IL-6 (**C, D**), hrIL-1β (**E, F**) or hrTNF-α (**G, H**) to evaluate fold-response to the treatment, compared to vehicle of the expression of extracellular matrix genes type I collagen α1 (COL1A1), type V collagen α1 (COL5A1), decorin (DCN), biglycan (BGN). Data is presented as (**A, B**) $2^{-\Delta Ct}$ to assess gene expression compared to CFL1 (housekeeping gene) or (**C-H**) fold to vehicle (dotted lines). Data is presented as mean with standard deviation (SD). Unpaired two-tailed Student's t-test. P-values in bold typeset indicate significance (p< 0.050).*

Table 1. Genotype and minor allele frequency distributions, and p-values for Hardy-Weinberg exact test for the four selected polymorphisms in the control (SWE-CON and SA-CON), the anterior cruciate ligament (SA-ACL) rupture group and ACL subgroup with a noncontact (SWE-NON, and SA-NON) mechanism of injury within the South African and Swedish cohorts[a].

		South Africa					Sweden		
		SA-CON	SA-ACL	p-value[a]	SA-NON	p-value[b]	SWE-CON	SWE-NON	p-value[b]
***COL5A1* rs12722 C > T**	n	96	93		48		109	77	
	CC	37 (36)	27 (25)	0.866	35 (17)	0.793	22 (24)	23 (18)	0.773
	CT	49 (47)	52 (48)		54 (26)		43 (47)	47 (36)	
	TT	14 (13)	22 (20)		10 (5)		35 (38)	30 (23)	
	T allele	38 (73)	47 (88)	0.851	38 (36)	1.000	56 (123)	53 (82)	0.617
	HWE	0.829	0.824		0.367		0.241	0.648	
***IL1B* rs16944 C > T**	n	*98*	*93*		*48*		*112*	*78*	
	CC	24 (24)	27 (25)	0.530	27 (13)	0.923	39 (44)	44 (34)	0.799
	CT	47 (46)	52 (48)		44 (21)		41 (46)	40 (31)	
	TT	29 (28)	22 (20)		29 (14)		20 (22)	17 (13)	
	T allele	52 (102)	47 (88)	0.411	51 (49)	0.971	40 (90)	37 (57)	0.542
	HWE	0.549	0.836		0.395		0.120	0.224	
***IL6* rs1800795 G > C**	n	*98*	*98*		*51*		*113*	*77*	
	GG	72 (71)	64 (63)	0.445	67 (34)	0.339	22 (25)	26 (20)	0.606
	GC	26 (25)	31 (30)		27 (14)		59 (67)	52 (40)	
	CC	2 (2)	5 (5)		6 (3)		19 (21)	22 (17)	
	C allele	15 (29)	20 (40)	0.185	20 (20)	0.369	48 (109)	48 (74)	1.000
	HWE	1.000	0.539		0.373		0.060	0.821	
***IL6R* rs2228145 A > C**	n	*95*	*95*		*49*		*112*	*76*	
	AA	54 (51)	58 (55)	**0.028**	55 (27)	0.054	53 (59)	46 (35)	0.618
	AC	34 (32)	39 (37)		43 (21)		37 (41)	46 (35)	
	CC	13 (12)	3 (3)		2 (1)		11 (12)	8 (6)	
	C allele	29 (56)	23 (43)	0.161	23 (23)	0.346	29 (65)	31 (47)	0.779
	HWE	0.082	0.385		0.257		0.253	0.599	

Genotype and allele frequencies are expressed as a percentage with the number of participants (n) in parentheses.
[a]CON vs. ACL (unadjusted p-value).
[b]CON vs. NON (unadjusted p-value).
P-values in bold typeset indicate significance (p< 0.050).

Allele combinations were inferred for *COL5A1-IL1B-IL6 and COL5A1-IL1B-IL6R*. For each of the two allele combinations, eight possible constructs were inferred at a frequency above 4%. For the South African cohort, no significant differences in the frequency distributions of these combinations were observed when all participants were evaluated or when only male or only female participants were compared (Supplementary Figure 2). The frequency distributions for the *COL5A1-IL1B-IL6* and the *COL5A1-IL1B-IL6R* allele combinations were similar between the control and cases when all participants or only

the female participants in the Swedish cohort were compared (Supplementary figures 3A and 3C). However, for the *COL5A1-IL1B-IL6* allele combination, when only males participants were evaluated, the T-C-G combination was significantly underrepresented (p=0.034 Haplo-score: 2.1) in the SWE-CON (7.7%, n=3) compared to the SWE-NON (18.0%, n=8) group (Supplementary Figure 3B). Furthermore, the frequency distributions for the *COL5A1-IL1B-IL6R* allele combinations, showed the T-C-A combination to be significantly underrepresented (p=0.044, Haplo-score: 2.0) in the SWE-CON (28.0%, n=11) compared to the SWE-NON (14.0%, n=6) group when only the male participants were compared in the Swedish cohort.

DISCUSSION

Considering the ligament as an integrative part of the knee joint, it is plausible that the ACL is subjected to cues derived from its surrounding anatomical structures, such as the synovium or synovial fluid. It is proposed, that as a response to repetitive mechanical overloading, macrophages might infiltrate tissues surrounding the ligaments[27]. Thereby, potentially exposing the ligamentocytes to an additional amount of specific inflammatory cytokines as part of the matrix remodeling mechanism. It is interesting, that some of the genetic susceptibility loci implicated in tendon and ligament injuries encode proteins involved in the homeostatic regulation of ECM components of both tendon and ligament, and components of the proinflammatory pathway[20]. This study, therefore used a hypothesis-based approach to evaluate the potential impact of the inflammatory pathway on modulating susceptibility to ligament injuries using an *in vitro* risk associated model, complimented with a genetic association approach.

For the functional *IL1B* rs16944 polymorphism, treatment with hrIL-β resulted in a 1.3-fold decrease (p=0.020) of *BGN* and a 2.1-fold (p=0.012) decrease of *COL5A1* in a genetic risk associated dependent manner. In addition, hrTNF-α treatment displayed a 2.0-fold (p=0.042) reduction in *COL5A1* mRNA levels in the fibroblasts with an *IL1B* rs16944 CC genotype. We suggest that, given an inflammatory micro-environment where these cytokines are abundant, matrix production is differently affected in *IL1B* high-risk compared to *IL1B* low-risk genetic profiles.

The *IL6* rs1800795 G-allele increases IL-6 mRNA expression levels, inducing apoptosis[28] which might decrease the production of ECM components. Our experiments indirectly support this hypothesis since fibroblasts having the *IL6* rs1800795 GG genotype displayed a 2.8-fold reduction (p=0.012) in *COL5A1* mRNA. Although not significant, a similar trend (p=0.07) was observed for other associated ECM components such as *DCN*. This is an important finding, since both *COL5A1* and *DCN* are required for normal fibrillogenesis[29].

At basal levels, the expression of proinflammatory genes was relatively low for all groups. However, with the exception of *TNFRSFA1* mRNA expression, mRNA levels of all the investigated cytokines were increased on average between 1.03 and 6109 fold in all the groups after treatment with hrIL-1β (Supplementary Figure 1C, D), hrIL-6 (Supplementary Figure 1 E, F) and TNF-α (Supplementary Figure 1G, H). More specifically, treatment with hrIL-1β significantly upregulated *IL1B* and *IL6* mRNA levels 3690 and 3948-fold respectively, although no statistically significant differences in their expression were noted between the high- and low-risk groups. This is in agreement with the hypothesis as shown in Figure 1 and with previous work[11].

These results support the proposal that polymorphisms within *IL1B* and *IL6* alter the expression of structural and fibril-associated ECM components and herewith possibly modulate the susceptibility to ligament injuries. This holds true in specific cohorts where these loci were implicated in risk models for the susceptibility to tendon and ligament injuries[3, 4]. These associations were therefore evaluated in two independent population groups from different ancestries, one from Sweden and the other from South Africa in an attempt to identify the susceptibility significance of these genetic loci in different populations. In the South African cohort, the *IL6R* rs2228145 CC genotype was significantly overrepresented (p=0.028) in the controls, compared to individuals that sustained an ACL injury. Although the CC genotype frequencies appeared to be similar in our Swedish cohort and in a previously reported South African Caucasian cohort[3], it did not reach the level of significance. As shown previously, the *COL5A1-IL1B-IL6* T-C-G and the *COL5A1-IL1B-IL6R* T-C-A allele combinations were found to be associated with an increased susceptibility to sustain an ACL rupture in the Swedish cohort when only male participants were evaluated[3, 4]. These associations were not reproduced in the South African cohort evaluated in this study, which might be explained by the different genetic background of the cohorts, as illustrated by the significant differences in genotype frequencies. Based on our power analysis, the sample size in this study is adequate to detect an allelic odds ratio (OR) of 2.4 at approximately 80% statistical power for Type 1 error detection. Although the study is underpowered to detect smaller effects, it is unlikely to reflect false positive data. In addition, it is important to note the current study used an *a priori* hypothesis and that reported associations are in line with previous ones. However, the findings should be cautiously interpreted and require confirmation in a larger cohort. We believe that all genetic data should be interpreted in the context of an individual's ancestral background. More important, all risk factors should be considered in a complex multifactorial disease, such as ACL injuries, to inform susceptibility. Risk susceptibility is most likely a combination of the interaction between a variety of extrinsic and intrinsic risk factors, including genetics.

A finely balanced inflammatory response is required for remodeling of the ECM[30] and that genetic polymorphisms potentially affect the production of inflammatory cytokines[12, 31].

The specific identity of these biological key role players however still remains unknown, including the threshold number and the time course of when they are required to direct the remodeling process within tendon and ligament. Therefore, future research should focus on the identification and quantification of inflammatory factors and on their time courses in tendon and ligament injuries. This may provide insights for biology-based therapies, such as anti-cytokine antibodies or cytokine antagonists and the most effective treatment period. Another approach might be to target cells that are responsible for the production of inflammatory cytokines, such as macrophages.

The *in vitro* experiments used dermal fibroblasts of eight individuals. Although dermal fibroblasts might have similar characteristics as tenocytes or ligamentocytes, their function and exact composition differ, possibly influencing their response to stimuli. In addition, a tissue-specific culture model applying a tensile force is required to study the effect of polymorphisms on matrix remodeling in more detail. Future research should aim to increase the number of donors. The difference in sex distribution in the genetic association study is explained by the fact that females participate less frequently in pivoting sports and therefore males and females were both tested together and separately for potential genetic associations with susceptibility for ACL injury.

In conclusion, this study describes specific polymorphisms within the inflammatory pathway to modulate the synthesis and degradation of structural and fibril-associated ECM components and thereby potentially contributing to an increased susceptibility to ACL injuries. This provisional evidence improves our understanding of the underlying mechanism for the genetic susceptibility to ACL ruptures and might lead to early identification of individuals who are of increased susceptibility to ACL injury and the potential application of personalized preventive or therapeutic interventions.

ACKNOWLEDGEMENTS

The authors would like to thank Prof. dr. G.J.V.M. van Osch for her contribution ot the study design and data interpretation. In addition, we are grateful to Dr. C. D'Alton and Dr. M. Nel for taking the skin biopsies. This research was funded in part by funds from the National Research Foundation of South Africa grant CPRR 106068 and the South University of Cape Town Research Council, Africa. M.A.M.S. and M.P. were financially supported by the European Union funded project RUBICON H2020-MSCA-RISE-2015-690850. M.R. was funded by the Harry Crossley Postdoctoral fellowship. In Sweden, financial support was obtaind from the Swedisch Scientific Research Council (Grants Nos. K2011-69X-21877-01-6, K2014-99X-21876-04-4), Västerbotten County Council (Grant no. ALF VLL548501 and Strategic funding VLL-358901; Project no. 7002795)

REFERENCES

1. Spindler KP, Wright RW. Clinical practice. Anterior cruciate ligament tear. *The New England journal of medicine.* 2008; 359(20):2135-2142.
2. Cox TR, Erler JT. Remodeling and homeostasis of the extracellular matrix: implications for fibrotic diseases and cancer. *Disease models & mechanisms.* 2011; 4(2):165-178.
3. Rahim M, Mannion S, Klug B, et al. Modulators of the extracellular matrix and risk of anterior cruciate ligament ruptures. *Journal of science and medicine in sport.* 2017; 20(2):152-158.
4. September AV, Nell EM, O'Connell K, et al. A pathway-based approach investigating the genes encoding interleukin-1beta, interleukin-6 and the interleukin-1 receptor antagonist provides new insight into the genetic susceptibility of Achilles tendinopathy. *British journal of sports medicine.* 2011; 45(13):1040-1047.
5. Malfait F, Wenstrup RJ, De Paepe A. Clinical and genetic aspects of Ehlers-Danlos syndrome, classic type. *Genetics in medicine : official journal of the American College of Medical Genetics.* 2010; 12(10):597-605.
6. Lulinska-Kuklik E, Rahim M, Domanska-Senderowska D, et al. Interactions between COL5A1 Gene and Risk of the Anterior Cruciate Ligament Rupture. *Journal of human kinetics.* 2018; 62:65-71.
7. September AV, 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. *British journal of sports medicine.* 2009; 43(5):357-365.
8. Posthumus M, September AV, Keegan M, et al. Genetic risk factors for anterior cruciate ligament ruptures: COL1A1 gene variant. *British journal of sports medicine.* 2009; 43(5):352-356.
9. Mannion S, Mtintsilana A, Posthumus M, et al. Genes encoding proteoglycans are associated with the risk of anterior cruciate ligament ruptures. *British journal of sports medicine.* 2014; 48(22):1640-1646.
10. Kalhor R, Kalhor K, Mejia L, et al. Developmental barcoding of whole mouse via homing CRISPR. *Science (New York, N.Y.).* 2018; 361(6405).
11. Tsuzaki M, Guyton G, Garrett W, et al. IL-1 beta induces COX2, MMP-1, -3 and -13, ADAMTS-4, IL-1 beta and IL-6 in human tendon cells. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society.* 2003; 21(2):256-264.
12. Landvik NE, Hart K, Skaug V, Stangeland LB, Haugen A, Zienolddiny S. A specific interleukin-1B haplotype correlates with high levels of IL1B mRNA in the lung and increased risk of non-small cell lung cancer. *Carcinogenesis.* 2009; 30(7):1186-1192.
13. Ferreira RC, Freitag DF, Cutler AJ, et al. Functional IL6R 358Ala allele impairs classical IL-6 receptor signaling and influences risk of diverse inflammatory diseases. *PLoS genetics.* 2013; 9(4):e1003444.
14. Qidwai T, Khan F. Tumour necrosis factor gene polymorphism and disease prevalence. *Scandinavian journal of immunology.* 2011; 74(6):522-547.
15. Tengman E, Brax Olofsson L, Nilsson KG, Tegner Y, Lundgren L, Hager CK. Anterior cruciate ligament injury after more than 20 years: I. Physical activity level and knee function. *Scandinavian journal of medicine & science in sports.* 2014; 24(6):e491-500.
16. Posthumus M, September AV, O'Cuinneagain D, van der Merwe W, Schwellnus MP, Collins M. The COL5A1 gene is associated with increased risk of anterior cruciate ligament ruptures in female participants. *The American journal of sports medicine.* 2009; 37(11):2234-2240.

17. Rahim M, Hobbs H, van der Merwe W, Posthumus M, Collins M, September AV. Investigation of angiogenesis genes with anterior cruciate ligament rupture risk in a South African population. *Journal of sports sciences.* 2018; 36(5):551-557.

18. Mokone GG, Schwellnus MP, Noakes TD, Collins M. The COL5A1 gene and Achilles tendon pathology. *Scandinavian journal of medicine & science in sports.* 2006; 16(1):19-26.

19. Rahim M, Collins M, September A. Genes and Musculoskeletal Soft-Tissue Injuries. *Medicine and sport science.* 2016; 61:68-91.

20. September A, Rahim M, Collins M. Towards an Understanding of the Genetics of Tendinopathy. *Advances in experimental medicine and biology.* 2016; 920:109-116.

21. Baumgarten IM. A comparison of metabolic pathway dynamics in man and other mammals [Internet]. 1993. Available from: http://digitalknowledge.cput.ac.za:8081/xmlui/handle/11189/796. Accessed 2018 Nov 2.

22. John T, Lodka D, Kohl B, et al. Effect of pro-inflammatory and immunoregulatory cytokines on human tenocytes. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society.* 2010; 28(8):1071-1077.

23. de Mos M, Joosten LA, Oppers-Walgreen B, et al. Tendon degeneration is not mediated by regulation of Toll-like receptors 2 and 4 in human tenocytes. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society.* 2009; 27(8):1043-1047.

24. Warnes G. Genetics: A package for population genetics. : R package (version 1.3.6); 2011.

25. Gonzalez JR, Armengol L, Sole X, et al. SNPassoc: an R package to perform whole genome association studies. *Bioinformatics (Oxford, England).* 2007; 23(5):644-645.

26. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *American journal of human genetics.* 2002; 70(2):425-434.

27. Wu P, Holguin N, Silva MJ, Fu M, Liao W, Sandell LJ. Early response of mouse joint tissue to noninvasive knee injury suggests treatment targets. *Arthritis & rheumatology (Hoboken, N.J.).* 2014; 66(5):1256-1265.

28. Millar NL, Wei AQ, Molloy TJ, Bonar F, Murrell GA. Cytokines and apoptosis in supraspinatus tendinopathy. *The Journal of bone and joint surgery. British volume.* 2009; 91(3):417-424.

29. Lieberthal J, Sambamurthy N, Scanzello CR. Inflammation in joint injury and post-traumatic osteoarthritis. *Osteoarthritis and cartilage.* 2015; 23(11):1825-1834.

30. Dakin SG, Dudhia J, Smith RK. Resolving an inflammatory concept: the importance of inflammation and resolution in tendinopathy. *Veterinary immunology and immunopathology.* 2014; 158(3-4):121-127.

31. Fishman D, Faulds G, Jeffery R, et al. The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. *The Journal of clinical investigation.* 1998; 102(7):1369-1376.

32. Chan KM, Fu SC, Wong YP, Hui WC, Cheuk YC, Wong MW. Expression of transforming growth factor beta isoforms and their roles in tendon healing. *Wound repair and regeneration : official publication of the Wound Healing Society [and] the European Tissue Repair Society.* 2008; 16(3):399-407.

SUPPLEMENTARY FIGURES AND TABLES

Supplementary Table 1. Primers used for RT-PCR analysis.

Gene	Forward primer (5' à 3')	Reverse primer (5' à 3')
COL5A1	GACAAGAAGTCCGAAGGGGC	TAGGAGAGCAGTTTCCCACG
COL1A1	TGAAGGGACACAGAGGTTTCAG	GTAGCACCATCATTTCCACGA
DCN	CAGACCAAGCACGCAAAACA	TCACAACCAGGGAACCTTGC
BGN	CACCGGACAGATAGACGTGC	CATGGCGGATGGACCTGGAG
IL6	GGATTCAATGAGGAGACTTGCC	GGGTCAGGGGTGGTTATTGC
IL1B	TTGCTCAAGTGTCTGAAGCAGC	CTTGCTGTAGTGGTGGTCGG
IL6R	CACGCCTTGGACAGAATCCA	TCCAGCAACCAGGAATGTGG
IL1R1	GGTAGACGCACCCTCTGAAG	GCATTTATCAGCCTCCAGAGAAGA
TNFRSFA1	ATTGGACTGGTCCCTCACCT	GTAGGTTCCTTTGTGGCACTT
CFL1	ATAAGGACTGCCGCTATGCC	CGGGGGCCCAGAAGATAAAC

Supplementary Table 2. Genetic risk profiles of the participants based on their (A) *IL1B* rs16944 C>T and (B) *IL6* rs1800795 G>C genotypes.

A.	Participant ID	*IL1B* rs16944	Sex
Low risk	5	CT	F
	7	CT	M
	8	CT	F
	11	CT	M
	22	TT	M
High risk	3	CC	M
	10	CC	M
	24	CC	F

B.	Participant ID	*IL6* rs1800795	Sex
Low risk	3	CC	M
	5	GC	F
	22	GC	M
	24	GC	F
High risk	7	GG	M
	8	GG	F
	10	GG	M
	11	GG	M

M, male; F, female.

Supplementary Table 3. Patient characteristics of the control (SWE-CON) group and the anterior cruciate ligament group with a noncontact (SWE-NON) mechanism of injury in a Swedish cohort[a].

	SWE-CON (n=116)	**SWE-NON** (n=79)	**p-value**
Age (years)[b]	44.7 ± 11.9 (114)	36.5 ± 13.7 (78)	**<0.001**
Sex (% male)	34.5 (116)	54.4 (79)	**0.014**
Height (cm)	172.3 ± 10.1 (108)	173.4 ± 8.6 (71)	0.438
Body mass (kg)[b]	72.1 ± 13.6 (107)	75.0 ± 12.8 (71)	0.149
Body mass index (kg/m^2)[b]	24.4 ± 2.9 (107)	24.7 ± 2.9 (70)	0.466

[a] *Values are presented as mean ± standard deviations except for sex, which is expressed as a percentage. The number of participants (n) with available data for each variable is in parenthesis.*

[b] *Self-reported values at the time of recruitment for the SWE-CON group, and at time of ACL rupture for the SWE-NON group.*

P-values in bold typeset indicate significance (p< 0.050).

Supplementary Table 4. Medical history and family injury for the control (SWE-CON) and non-contact (SWE-NON) anterior cruciate ligament rupture group of the Swedish cohort[a].

	Male			Female			
	SWE-CON (n=40)	SWE-NON (n=42)	p-value[b]	SWE-CON (n=76)	SWE-NON (n=37)	p-value[b]	p-value[c]
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
• Other	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

[a]Values are expressed as percentages with the number of participants (n) with available data in parentheses.
[b]SWE-CON vs. SWE -NON, p-values in bold typeset indicate significance (p< 0.050).
[c]SWE-CON (male + female) vs. SWE-NON (male + female).
P-values in bold typeset indicate significance (p< 0.050).

Supplementary Table 5. Genotype effects per patient characteristic in the Swedish cohort.

A. Genotype effects *COL5A1* rs2228145 C>T.

	C/C (n=42)	**C/T** (n=83)	**T/T** (n=61)	**p-value**
Age (years)	40.2 ± 12.6 (40)	39.9 ± 13.2 (83)	44.0 ± 13.5 (60)	0.145
Sex (% male)	50.0 (21)	43.3 (36)	34.4 (21)	0.197
Height (cm)	174.7 ± 11.0 (38)	172.1 ± 9.0 (78)	171.9 ± 8.6 (55)	0.308
Body mass (kg)	74.1 ± 11.8 (38)	77.3 ± 13.8 (77)	73.3 ± 13.2 (55)	0.782
Body mass index (kg/m^2)	24.0 ± 2.7 (37)	24.6 ± 2.8 (77)	24.7 ± 3.1 (55)	0.536

B. Genotype effects *IL1B* rs16944 C>T.

	C/C (n=78)	**C/T** (n=77)	**T/T** (n=35)	**p-value**
Age (years)	42.2 ± 13.2 (77)	39.8 ± 13.3 (75)	42.2 ± 13.5 (35)	0.483
Sex (% male)	50.0 (39)	31.1 (24)	48.6 (17)	0.091
Height (cm)	173.8 ± 9.3 (72)	171.2 ± 9.1 (72)	173.7 ± 9.5 (30)	0.203
Body mass (kg)	75.7 ± 12.2 (70)	71.6 ± 12.8 (72)	73.0 ± 16.0 (31)	0.172
Body mass index (kg/m^2)	24.9 ± 2.8 (70)	24.3 ± 2.7 (72)	24.7 ± 3.2 (30)	0.360

C. Genotype effects *IL6* rs1800795 G>C.

	C/C (n=45)	**C/G** (n=107)	**G/G** (n=38)	**p-value**
Age (years)	44.8 ± 12.5 (38)	40.8 ± 13.7 (104)	40.2 ± 12.7 (45)	0.211
Sex (% male)	47.4 (18)	37.4 (40)	46.7 (21)	0.536
Height (cm)	172.7 ± 8.3 (36)	172.3 ± 9.7 (96)	173.0 ± 9.7 (43)	0.927
Body mass (kg)	73.1 ± 11.5 (72)	73.5 ± 14.6 (96)	73.1 ± 12.1 (43)	0.983
Body mass index (kg/m^2)	24.5 ± 2.6 (35)	24.7 ± 3.0 (95)	24.3 ± 2.7 (43)	0.679

D. Genotype effects *IL6R* rs2228145 A>C.

	A/A *(n=94)*	**A/C** *(n=76)*	**C/C** (n=18)	**p-value**
Age (years)	41.4 ± 13.2 (93)	40.9 ± 14.2 (75)	42.9 ± 8.0 (17)	0.850
Sex (% male)	42.6 (40)	36.8 (28)	55.6 (10)	0.882
Height (cm)	172.2 ± 9.7 (84)	172.5 ± 8.6 (70)	174.6 ± 9.6 (18)	0.610
Body mass (kg)	72.4 ± 11.7 (85)	73.5 ± 14.4 (69)	75.4 ± 12.4 (17)	0.653
Body mass index (kg/m^2)	24.2 ± 2.6 (84)	24.8 ± 3.0 (69)	24.7 ± 2.6 (17)	0.407

Supplementary Table 6. Genotype and minor allele frequency distributions, and p-values for Hardy-Weinberg exact test of the four selected polymorphisms within the South African (SA) and Swedish (SWE) cohort for asymptomatic controls (CON), the anterior cruciate ligament (ACL) rupture group and the ACL subgroup with a noncontact (NON) mechanism[a].

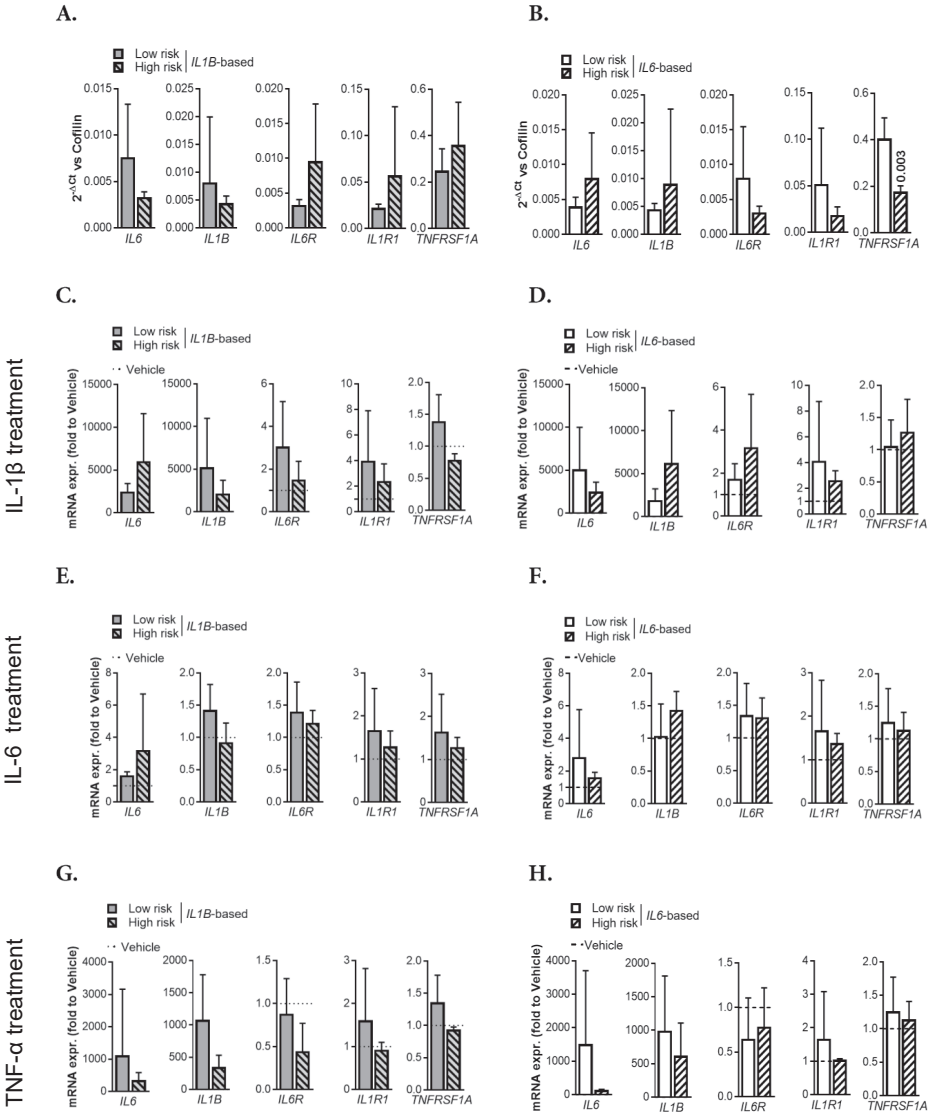
		CON			ACL				
		SA-CON	SWE-CON	P-value[b]	SA -ACL	SWE -ACL	P-value[1]	SA -NON	P-value[c]
	n	*96*	*109*		*93*	*71*		*48*	
	CC	38 (36)	22 (24)	**<0.001[d]**	41 (38)	23 (18)	**0.019[d]**	35 (17)	**0.021[d]**
COL5A1 **rs12722 C > T**	CT	49 (47)	43 (47)		45 (42)	47 (36)		54 (26)	
	TT	14 (13)	35 (38)		14 (13)	30 (23)		10 (5)	
	T allele	38 (73)	56 (123)	**<0.001[d]**	37 (68)	53 (82)	**0.003[d]**	38 (36)	**0.022[d]**
	HWE	0.829	0.241		0.824	0.648		0.367	
	n	*93*	*112*		*93*	*78*		*48*	
	CC	24 (24)	39 (44)	**0.033[d]**	27 (25)	44 (34)	**0.033[d]**	27 (13)	0.095
IL1B **rs16944 C > T**	CT	47 (46)	41 (46)		52 (48)	40 (31)		44 (21)	
	TT	29 (28)	20 (22)		22 (20)	17 (13)		29 (14)	
	T allele	52 (102)	40 (90)	**0.019[d]**	47 (88)	37 (57)	0.058	51 (49)	**0.033[d]**
	HWE	0.549	0.120		0.836	0.224		0.395	
	n	*98*	*113*		*98*	*77*		51	
	GG	72 (71)	22 (25)	**<0.001[d]**	64 (63)	26 (20)	**<0.001[d]**	67 (34)	**<0.001[d]**
IL6 **rs1800795 G > C**	GC	26 (25)	59 (67)		31 (30)	52 (40)		27 (14)	
	CC	2 (2)	19 (21)		5 (5)	22 (17)		6 (3)	
	C allele	15 (29)	48 (109)	**<0.001**	20 (40)	48 (74)	**<0.001[d]**	20 (20)	**<0.001[d]**
	HWE	1.000	0.061		0.539	0.821		0.373	
	n	*112*	*95*		*76*	*95*		49	
	AA	53 (59)	54 (51)	0.996	46 (35)	58 (55)	0.165	55 (27)	0.886
IL6R **rs2228145 A > C**	AC	37 (41)	34 (32)		46 (35)	39 (37)		43 (21)	
	CC	11 (12)	13 (12)		8 (6)	3 (3)		2 (1)	
	C allele	29 (65)	29 (56)	1.000	31 (47)	23 (43)	0.108	23 (23)	0.256
	HWE	0.082	0.253		0.385	0.599		0.257	

[a]Genotype and allele frequencies are expressed as a percentage with the number of participants (n) in parentheses.

[b]SA vs. SWE cohort (unadjusted p-value).

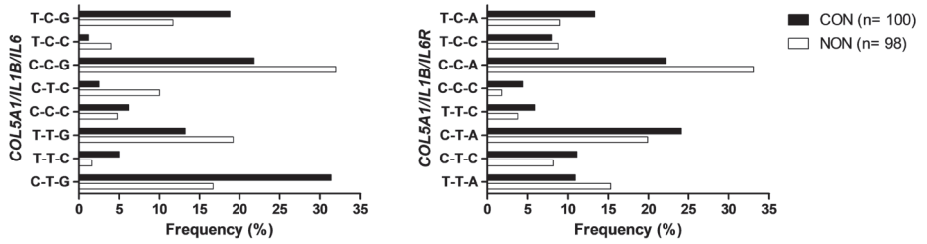
[c]SA with non-contact mechanism vs. SWE cohort (unadjusted p-value).

[d]P-values in bold typeset indicate significance (p< 0.050).

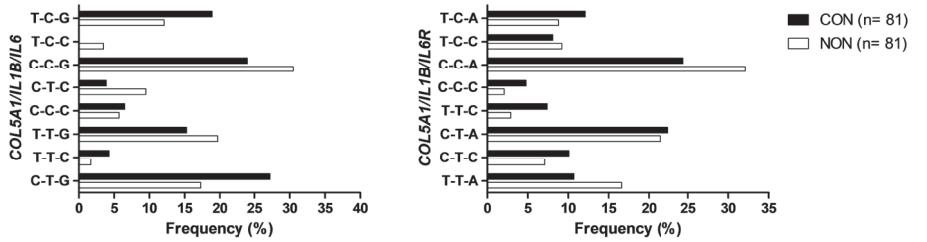


Supplementary Figure 1: mRNA expression in unstimulated fibroblasts (A, B) and fibroblasts stimulated with IL-1β (C, D), IL-6 (E, F) or TNF-α (G, H). mRNA expression levels of the cytokine-related genes, IL6, IL1B, IL6R, IL1R1 and TNFRSF1AR in fibroblasts classified in high-risk or low-risk for ACL injuries based on (A) *IL6* rs1800795 G>C or (B) *IL1B* rs12722 C>T. Data is presented as $2^{-\Delta Ct}$ to assess gene expression compared to CFL1 (housekeeping gene). Data is presented as mean with standard deviation (SD). Unpaired two-tailed Student's *t*-test, P-values in bold indicate significance (p<0.050).

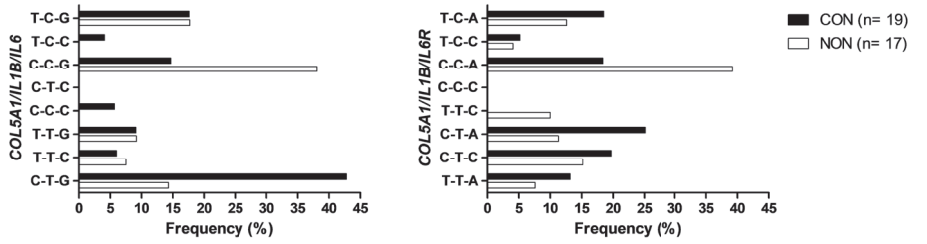
A. All participants (males and females).



B. Male participants.

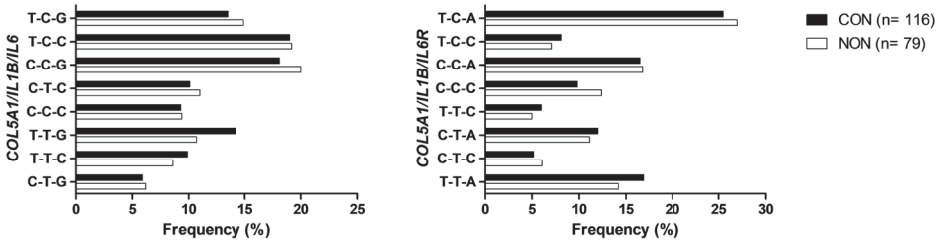


C. Female participants.

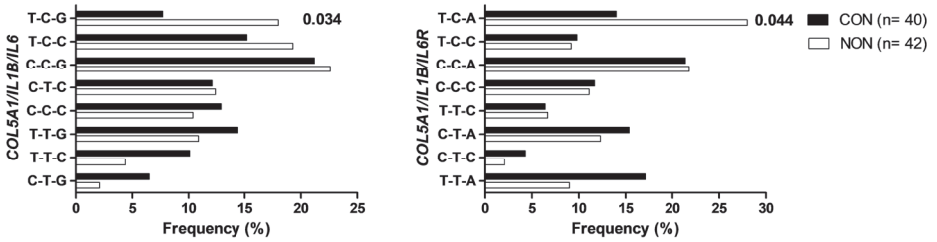


Supplementary Figure 2: Frequency distributions in the South African cohort for the *COL5A1* rs12722 C>T, *IL1B* rs16944 C>T, *IL6* rs1800795 G>C or *IL6R* rs2228145 A>C polymorphisms in the control group (CON; black bars) and the anterior cruciate ligament rupture group (ACL; white bars) for (A) all participants (males and females), (B) the male participants and (C) female participants. The number of participants (n) in each group is in parentheses.

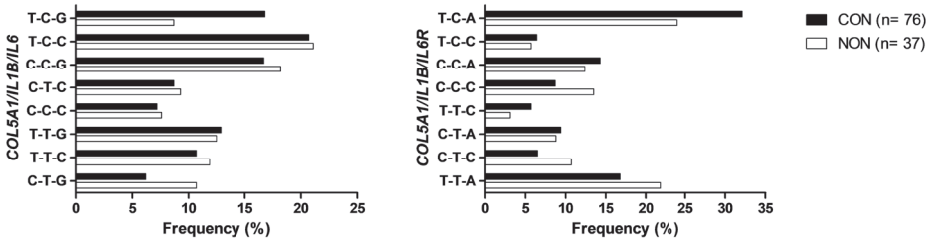
A. All participants (males and females).



B. Male participants.



C. Female participants.



Supplementary Figure 3: Frequency distributions in the Swedish (SWE) cohort for the *COL5A1* rs12722 C>T, *IL1B* rs16944 C>T, *IL6* rs1800795 G>C or *IL6R* rs2228145 A>C polymorphisms in the control group (SWE-CON; black bars) and the non-contact anterior cruciate ligament rupture group (SWE-NON; white bars) for (A) all participants (males and females), (B) male participants and (C) female participants in the Swedish cohort. The number of participants (n) in each group is in parentheses. P-values in bold typeset indicate significance (p< 0.050).