HUMAN BIOLOGY AND PUBLIC HEALTH JOURNAL OF THE AUXOLOGICAL SOCIETY

ISSN 2748-9957 | Vol. 1 2025 | pp. 1–20

No correlation between *HSPG2* genetic variants and anthropometric characteristics within an ACL rupture risk modelling study

Lethabo Ramoshaba[1] · Mary-Jessica N Laguette[1, 2] · Malcolm Collins[1, 2] · Alison V September[1, 2]

[1]Division of Physiological Sciences, Health through Physical Activity, Lifestyle and Sport Research Centre (HPALS), Department of Human Biology, University of Cape Town, Newlands, South Africa
[2]International Federation of Sports Medicine (FIMS) Collaborative Centre of Sports Medicine, University of Cape Town, Newlands, South Africa

Citation:

Ramoshaba, L./Laguette, M.-J.N/Collins, M. et al. (2025). No correlation between *HSPG2* genetic variants and anthropometric characteristics within an ACL rupture risk modelling study, Human Biology and Public Health 1.
https://doi.org/10.52905/hbph2025.1.93.

Received: 2024-10-11
Accepted: 2025-01-27
Published: 2025-07-11

Conflict of Interest:

There are no conflicts of interest.

Correspondence to:

Keywords

Anterior cruciate ligament, body mass index, *HSPG2* , Heparan sulphate proteoglycan-2 gene, SNHA, St Nicholas House analysis

Abstract

Background Anterior cruciate ligament (ACL) ruptures are common musculoskeletal injuries, influenced by extrinsic and intrinsic factors such as genetic variations and anthropometric traits. While these factors contribute to ACL rupture susceptibility, their interactions are underexplored.

Objectives To investigate the relationship between *HSPG2* variants and anthropometric traits in participants from an ACL study from Poland and Sweden.

Hypothesis Genetic variability within *HSPG2* loci along with height variability may collectively contribute to ACL rupture susceptibility.

Sample and methods A genetic case-control association study was conducted with two cohorts from Poland and Sweden and a combined cohort. Participants were self-reported Caucasian and physically active. The combined cohort consisted of 265 asymptomatic controls (POL-CON=150; SWD-CON=116); 237 ACL rupture cases (POL-ACLR=141; SWD-ACLR=95) and a subgroup of 135 non-contact ACL ruptures (POL-NON=54; SWD-NON=79). Participants were genotyped for rs2291826 A>G and rs2291827 G>A and data were analysed using R, with $p<0.05$.

Results Strong correlations were found between mass and BMI across all cohorts (r=0.78–0.81), suggesting these traits may influence injury risk. Sex-mass and sex-height correlations were consistent, with a strong negative correlation between sex and height in the Swedish cohort (r=-0.75). No positive correlations were found between the *HSPG2* variants and anthropometric traits, except a moderate negative correlation between rs2291826A>G and height in the Swedish cohort (r=-0.019, $p<0.009$), suggesting possible genotype effect on height.

Conclusion Mass and BMI were highlighted as potential risk factors for ACL rupture. Height-mass relationships varied by sex and population, suggesting both genetics and environment impact injury patterns. Further testing of the variants may clarify their role in ACL injury variability.

Take-home message for students It is important to test the potential relationship between genetic variants and anthropometric measurements towards identifying potential confounders in a genetic association study. This study has shown that height is not a confounder for these variants.

Introduction

Rupture of the anterior cruciate ligament (ACL) is one of the most severe and common musculoskeletal soft tissue knee injuries (Alsayed et al. 2023; Kiapour and Murray 2014). It is associated with immediate pain, rapid swelling, weakness of the knee, and instability, which significantly impair mobility (Logerstedt et al. 2018). The incidence rate of ACL ruptures is estimated at 75–80 per 100,000 per year, especially in physically active adolescents and athletes (15–40 years) (Magnusson et al. 2020). ACL ruptures occur most frequently in athletes who participate in sporting activities that require sudden changes in direction, jumps, and twists (Joseph et al. 2013). The lifetime risk of developing knee osteoarthritis as a result of an ACL injury is 34%, 10–15 years post-injury and surgical repairs (Davis et al. 2021; Suter et al. 2017). It is particularly concerning that these injuries are increasingly affecting young individuals; consequently, there is a growing number of osteoarthritis cases appearing at a younger age, which raises further concern (Murray et al. 2000; Lohmander et al. 2007).

Ligaments have a low healing capacity, poor blood supply, and low cellular content, all of which has been ascribed to contribute to the ACL's susceptibility to injury (Kiapour and Murray 2014). As a result, ACL ruptures not only cause significant direct healthcare costs, but also lead to indirect costs, such as social and productivity losses and increased absenteeism, collectively creating a long term economic burden (Griffin et al. 2006). The burden of ACL injuries is well recognized, however, the aetiology of ACL rupture remains unclear (Kobayashi et al. 2010). It is known that ACL injuries are multifactorial, meaning that an interaction between various intrinsic and extrinsic risk factors contribute to the susceptibility to these injuries (Hewett et al. 2016; Bittencourt et al. 2016; Meeuwisse 1994). Although research is increasingly focusing on the contributions of biomechanical (modifiable) risk factors and environmental risk factors to ACL injuries (Murphy et al. 2003; Griffin et al. 2006), there is growing evidence that an individual's genetic predisposition and anthropometric traits contribute to susceptibility to musculoskeletal soft tissue injuries such as ACL ruptures (Joseph et al. 2013; Magnusson et al. 2020; Rahim et al. 2017; Snaebjörnsson et al. 2019).

More than 80 loci encoding genes related to components of the extracellular matrix (ECM) have been identified as being associated with predisposition to ACL rupture (Ribbans et al. 2022). These genes encode proteins which form structural, non-structural and regulatory components of the ECM (Posthumus et al. 2010; Ribbans et al. 2022; Willard et al. 2018; Feldmann 2022; Feldmann et al. 2022; Dlamini et al. 2023). Recently, Dlamini et al. (2023) highlighted genetic variants within two candidate genes, heparan sulphate proteoglycan-2 (*HSPG2)* and integrin beta 2 *(ITGB2)*, to be associated with ACL rupture susceptibility. They showed that the G-A-C allele combination between *HSPG2* (rs2291826 A>G-rs2291827 G>A) and *ITGB2* (rs2230528 C>T) variants was associated with reduced risk of ACL rupture as a proxy for gene–gene interaction between these two genes.

The *ITGB2* gene encodes integrin β2 subunit, a heterodimeric transmembrane receptor primarily involved in leukocyte adhesion and immune responses (Bednarczyk et al. 2020; Huang et al. 2019). The potential contribution of this gene to ACL rupture susceptibility may arise through the inflammatory pathways and gene–gene interactions (Dlamini et al. 2023). The *HSPG2* gene is a multidomain gene that encodes perlecan, a large basement mem-

brane protein (Gubbiotti et al. 2017). Perlecan belongs to the proteoglycan family, which form a major component of extracellular matrix (ECM) ground substance (Gubbiotti et al. 2017; Arikawa-Hirasawa 2022). Perlecan plays a multifaceted role in maintaining the structural and functional integrity of the ECM (Arikawa-Hirasawa 2022). It interacts with various molecules, including collagens and growth factors such as vascular endothelial growth factor (VEGF) and fibroblast growth factor (FGF) to maintain ECM stability and support cellular process (Farach-Carson and Carson 2007; Arikawa-Hirasawa 2022). The expression of *HSPG2* contributes to various developmental and cellular processes including bone and cartilage development, angiogenesis, inflammatory responses and wound healing (Farach-Carson and Carson 2007; Martinez et al. 2018). Given the role of *HSPG2* in skeletal and cartilage development, it could also contribute to the regulation of height, a known risk factor for ACL ruptures. While the *ITGB2* was identified in previous research, the study focused on the *HSPG2* due to its stronger biological relevance to ligament integrity and growth regulation.

The exact genetic variants and their interaction with extrinsic or intrinsic factors have not been interrogated and require exploration. Some of the intrinsic risk factors include the anthropometric traits such as mass, height and body mass index (BMI), which have been identified as potential factors contributing to susceptibility to musculoskeletal soft tissue injuries such as ACL rupture (Snaebjörnsson et al. 2019; Alsayed et al. 2023). A study by Alsayed et al. (2023) reported that individuals with a BMI greater than 25 kg/m^2 were at a greater risk of sustaining a sport-related ACL injury compared to individuals with a lower BMI. A high BMI may result in increased mechanical stress on the knee, resulting in knee injury (Alsayed et al. 2023; Snaebjörnsson et al. 2019). It has been noted that females have about 9 times greater risk of ACL injury than men, potentially due to biomechanical, anatomical and hormonal differences (Seneviratne et al. 2004; Bruder et al. 2023). The variability in these factors also has a genetic component, making it important to identify whether these intrinsic risk factors are also confounders when trying to characterise the genetic risk susceptibility for a complex phenotype such as ACL ruptures.

The aim of this study was therefore to investigate the relationship between genetic variants of *HSPG2* and anthropometric factors. It was hypothesised that genetic variability at the *HSPG2* loci, together with variability in height, may collectively explain the contribution to susceptibility to ACL rupture. In attempting to characterise the risk factors for ACL rupture susceptibility, it is important to identify the confounding factors and determine which factors are associated with each other.

Sample and Methods

Samples

Participant recruitment

A case-control genetic association study was conducted comprising two previously recruited cohorts from Poland and Sweden, respectively, as well as a combined cohort of members from both populations. Each participant was of self-reported European ancestry and provided written informed consent to participate in accordance with principles in the Declaration of Helsinki. All participants in each cohort completed detailed questionnaires covering demographics, lifestyle habits, occupation details, and sporting background including

type of sporting activity, duration of participation, and frequency. The questionaries requested medical details related to the ACL injuries, such as the mechanism of injury, and history of other ligament or tendon injuries. Ethical approval for this study was obtained from the Human Research Ethics Committee of the Faculty of Health Science, University of Cape Town (HREC: 026/2023). Approval was also obtained from each of the respective local ethics committees for the individual cohorts recruited from the Regional Ethical Review Board in Umeå, Sweden (dnr. 2011-200-31 M), Bioethics Committee for Clinical Research, Regional Medical Chamber, Gdansk, Poland (KB-8/16).

The ACL rupture group included individuals older than 18 years at the time of recruitment. All the participants had a clinical diagnosis of ACL rupture by physical examination confirmed by magnetic resonance imaging (MRI) or arthroscopy. This group of participants included injuries sustained through contact and non-contact mechanisms. Non-contact ACL rupture was analysed as a subgroup of the cases (NON). The control (CON) group included individuals with no history of an ACL injury or other ligament or tendon injuries. They regularly participated in similar sports activities as the affected individuals and were in the same age group. All participants engaged in regular physical activity, primarily on a recreational level.

Individuals taking chronic medications or diagnosed with a connective tissue disease or other systemic disease known to affect connective tissues were excluded. In addition, individuals with current or previous use of fluoroquinolone antibiotics (within the last 12 months from the time of recruitment) or previous use of local corticosteroids injections into the ACL or surrounding areas were excluded, as these factors can have a negative effect on collagen synthesis and result in matrix degradation thus increasing susceptibility to ACL injuries.

Participant characteristics

The combined cohort, including participants from the Polish and Swedish cohorts, consisted of 502 participants. The cohort consisted of 265 asymptomatic control participants (COMB-CON), 237 participants with ACL rupture (COMB-ACLR), and 135 participants (COMB-NON) who reported sustaining an ACL rupture by a non-contact mechanism (Table 1). The Polish cohort consisted of 291 physically active, unrelated participants recruited between 2008 and 2018 from the Galen Orthopaedics Clinic in Poland. This included 150 asymptomatic control participants (POL-CON), 141 ACL participants (POL-ACLR), and 54 participants who reported a non-contact mechanism (POL-NON). The male controls (n=112) and ACL rupture cases (n=102) were recruited from the Polish soccer league. The female cases (n=39) were recruited from soccer teams and skiing sports. The female controls (n=37) were recruited from sports clubs and wellness centres (Cięszczyk et al. 2017; Feldmann et al. 2022; Lulińska-Kuklik et al. 2019).

The Swedish cohort consisted of 211 physically active and unrelated participants recruited between 2011 and 2013 from the orthopedic clinics of two major hospitals in the city of Umeå (in the Vaesterbotten region) and Luleå (in the Norrbotten region) respectively. This cohort comprised 116 asymptomatic control participants (SWD-CON) with no history of ACL injury, 95 participants (SWD-ACLR) with ACL injury, and 79 participants who reported an ACL rupture by a non-contact mechanism (SWD-NON). Most of the recruited participants had a previously described long-term follow-up of the ACL injury (Suijkerbuijk et al. 2019).

Participant sports data were previously published by Suijkerbuijk et al. (2019) and Cięszczyk et al. (2017).The type of sport participation was categorised into non-contact jumping sports, non-contact non-jumping sports, and skiing sports. The participants in the Polish cohort were matched in terms of sport type, level, and frequency of participation. Details on the years of sports participation in the Polish were not available. Most of the participants in the Swedish cohort reported participating in non-contact jumping sports at a recreational level.

Method

DNA extraction

For the participants in the Polish cohort, genomic DNA was extracted from oral epithelial cells using a GEN Elute Mammalian Genomic DNA Miniprep Kit (Sigma; Darmstadt, Germany) according to the manufacturer's recommendations. For the Swedish cohort, genomic DNA was extracted from venous blood using rapid non-ethanol precipitation as previously described by Lahiri and Nurnberger (1991), with slight modifications from Mokone et al. (2006).

Candidate gene and genetic variants selection

The candidate gene *HSPG2* was selected because of its multifunctional role in different biological pathways involved in remodelling the extracellular matrix (ECM) components and thereby maintaining ECM homeostasis. Participants were genotyped for the following genetic variants: *HSPG2* (rs2291826 A>G, rs2281827 G>A). The genetic variants were selected from a previous Whole Exome Sequencing (WES) study, which showed that the variants had a difference in allele frequency distribution of atleast 30% between 10 tendinopathy cases and 10 controls (Gibbon et al. 2018). These genetic variants were also selected because they reported a minor allele frequency of > 5% in the Caucasian population, as reported in the ALFA Allele Frequency dataset and 1000 Genomes Project hosted by the National Centre for Biotechnology Information (NCBI database (http://www.ncbi.nlm.nih.gov/). The functional significance of the genetic variants remains unknown.

Genotyping

For the standard genotyping protocols, a predesigned TaqMan polymerase chain reaction (PCR) assay was used to genotype the *HSPG2* rs2291826 A>G (Assay ID: C_15966515_10) and *HSPG2* rs2291827 G>A (Assay ID: C__15966517_10) (ThermoFisher Scientific; Waltham, Massachusetts, USA). Real-time polymerase chain reactions (PCR) were performed for all genetic variants using the Quant Studio™ 3 Real-Time PCR System (Applied Biosystems; Waltham, Massachusetts, USA). Three negative controls (no DNA) and five replicate samples (known genotypes) were included on each 96-walled PCR plate (FG-TCII reaction plate) as a quality control for reliable genotyping and detection of contamination. Genotypes were confirmed by two independent investigators. Genotypes were called using the Thermofisher Cloud™ Suite (Thermo Fisher Scientific; Waltham, Massachusetts, USA) with an average call rate of 82.7%. All laboratory work was conducted at the University of Cape Town, at the Health through Physical Activity, Lifestyle and Sport (HPALS) Research Centre, which is located at the Sports Science Institute of South Africa (Newlands).

Statistical analysis

The sample size for the study was calculated using QUANTO version 1.2.4 (http:/

/keck.usc.edu/biostatistics/software/). For the Polish and Swedish cohorts, a sample size of 118 or more cases was required to detect an OR of 1.8 or higher. For the Swedish cohort, a sample size of 85 cases was required to detect an odds ratio (OR) of 2.0 or higher. All calculations assumed minor allele frequencies between 0.2 and 0.5 with a power of 80% and a significance of 5%. The data were analysed with the R Project for Statistical Computing Version 4.3.1 . The distribution of continuous data was analysed using the Shapiro-Wilk normality test, and data was presented as means±standard deviations (Shapiro and Wilk 1965). Non-parametric data was presented as median and interquartile ranges. Categorical data were compared using the chi-square test or Fisher's exact test (sample size <10) and presented as percentages. One-way analysis of variance (ANOVA) was used for continuous parametric data, while the Mann-Whitney U or Kruskal-Wallis test was used for non-parametric data.

The R packages genetics, version 1.3.8.1.3 (González et al. 2007) and SNPassoc version 2.1.0 (Schaid et al. 2012) were used to determine the differences in genotypes and allele frequencies between the groups and to calculate the probabilities of Hardy-Weinberg equilibrium (HWE) for the variants and linkage disequilibrium between the *HSPG2* genetic variants. The *HSPG2*: rs2291826 A>G, rs2291827G>A variants were in a moderate positive LD (D'=0.703), with a significant non-random association (X2=912.35; p<0.001) of the alleles at the two loci. GraphPad Prism version 10.1.2 (GraphPad Software., Boston, Massachusetts, USA) was used to visualize genotype and allele frequencies between the different populations. The St Nicholas House Analysis (SNHA) was used to visually represent the relationship between different variables such as genetic variables, demographic and anthropometric factors and phenotypic traits (Hermanussen et al. 2021).

Results

Participant characteristics

In the combined cohort, the COMB groups (CON vs. ACLR), were covaried for sex and country of recruitment Table 1. The COMB-ACLR group (33.8 ± 11.7 years, n=235) was significantly older than the COMB-CON group (p= 0.036, 31.3 ± 14.2 years, n = 263). The COMB-ACLR group also had a significantly higher mass (p = 0.016, 75.4 ± 14.2 kg, n = 219) compared to the COMB-CON group (72.4 ± 12.7 kg, n=256); however, after the height variable was covaried for sex, the significance was no longer noted. No significant differences were noted for the COMB-NON subgroup.

For the individual cohorts, Polish cohort had younger participants compared to the Swedish cohort participants. The POL-CON group (21.0 ± 1.8 years, n = 149) was significantly younger than the POL-ACLR group (p < 0.001, 31.6 ± 10.0 years, n = 142) and the POL-NON subgroup (p < 0.001, 30.7 ± 9.9 years, n = 56). However, for the Swedish cohort, the SWD-CON group (44.7 ± 11.9 years, n = 114) was significantly older than the SWD-ACLR group (p < 0.001, 37.1 ± 13.3 years, n = 93) and SWD-NON subgroup (p < 0.001, 36.5 ± 13.7 years, n = 78) (Table 1).

The Polish and Swedish groups were similar in height. However, with regards to mass, the Polish cohort had a significantly higher mass and BMI. The POL-ACLR group had a significantly higher mass (p < 0.001, 79.0 ± 14.8 n = 140) and a higher BMI (p < 0.001, 25.0 ± 4.0 kg.m-2, n = 137) than the POL-CON group. Similarly, the high mass and BMI was noted in the POL-NON subgroup (mass: p < 0.001, 81.4 ± 16.7 kg, n = 56; BMI: p < 0.001, 25.1 ± 5.1 kg.m-2, n =56). After the mass and

BMI were covaried for age, the significance remained only for the POL-NON subgroup (Table 1).

No significant genotype effect on any of the descriptive variables was noted for the combined cohort (Table 2). For the individual cohorts, genotype effect for height was only noted for the Swedish cohort for the *HSPG2* rs2291826 A>G variant distribution where the A/A and A/G genotype was frequently noted in the distribution (p = 0.009) (Table S4).

Genotype and allele frequency

There were no significant differences in the distribution of genotype and allele frequencies for the *HSPG2* rs2291826 A>G variant (COMB-CON vs. COMB-ACLR: p = 0.067, and COMB-CON vs. COMB-NON: p = 0.334) between the groups (Figure 1A). Similar findings were noted for the *HSPG2* rs2291827 G/A variant: (COMB-CON vs. COMB-ACLR: p=0.482, and COMB-CON vs. COMB-NON: p=0.272) (Figure 1B). No significant differences in either the genotype or allele frequency distributions for the *HSPG2* rs2291826 A>G (Figure 2A) and rs2291827 G>A (Figure 2B) variant were noted in the combined cohort when only males or females were examined.

Table 1 Participant's descriptive characteristics in the combined cohort (COMB) (male and female) as well as individual cohorts (Poland, Sweden), for the control (CON), anterior cruciate ligament ruptures groups (ACLR), and non-contact mechanism anterior cruciate ligament rupture subgroup (ACL-NON).

	COMB-CON	COMB-ACLR	p values	COMB-NON	p values
N	265	237		135	
Age(years)	31.3±14.2 (263)	33.8±11.7 (235)	0.036 (0.002)B	34.0±12.6 (134)	0.054
Sex (%male)	57.4 (152)	62.0 (147)	0.331	66.7 (90)	0.096
Height (cm)	175.6±10.2 (257)	175.6±9.5 (224)	0.985	176.3±9.3 (127)	0.503
Mass (kg)	72.4±12.7 (256)	75.4±14.2 (219)	0.016 (0.222)C	75.1±15.0 (122)	0.072
BMI (kg.m-2)	22.7±4.9 (249)	23.41±6.1 (219)	0.186	22.9±6.5 (126)	0.738
	POL-CON	POL-ACLR	p values	POL-NON	p values
N	149	141		54	
Age (years)	21.0±1.8 (149)	31.6±10.0 (142)	<0.001	30.7±9.9 (56)	<0.001
Sex (%male)	75.2 (112)	71.8 (102)	0.519	85.7 (48)	0.113
Height (cm)	178.0±9.7 (149)	177.3±9.6 (139)	0.545	180.0±8.8 (56)	0.176
Mass (kg)	72.6±12.0 (149)	79.0±14.8 (140)	<0.001 (0.086)A	81.4±16.7 (56)	0.000 (0.019)A
BMI (kg.m-2)	22.8±2.5 (149)	25.0±4.0 (137)	<0.001 (0.078)A	25.1±5.1 (56)	<0.001 (0.096)A
	SWD-CON	SWD-ACLR	p values	SWD -NON	p values
N	116	95		79	
Age (years)	44.7±11.9 (114)	37.1±13.3 (93)	<0.001	36.5±13.7 (78)	<0.001
Sex (%male)	34.5 (40)	47.4 (45)	0.059	53.2 (42)	0.010 (0.002)A
Height (cm)	172.3±10.1 (108)	172.8±8.6 (85)	0.730	173.4±8.6 (71)	0.435
Mass (kg)	72.1±13.8 (107)	69.1±10.8 (79)	0.113	69.8±10.9 (66)	0.252
BMI (kg.m-2)	23.7±2.6 (44)	22.9±2.6 (73)	0.125	22.9±2.6 (62)	0.134

Values are expressed as mean±standard deviation, and sex is represented as a percentage.
p-values in bold typeset indicate significance (p<0.05).
p-values in parentheses are adjusted for the following variables: A Age, B country of recruitment, and C Sex.

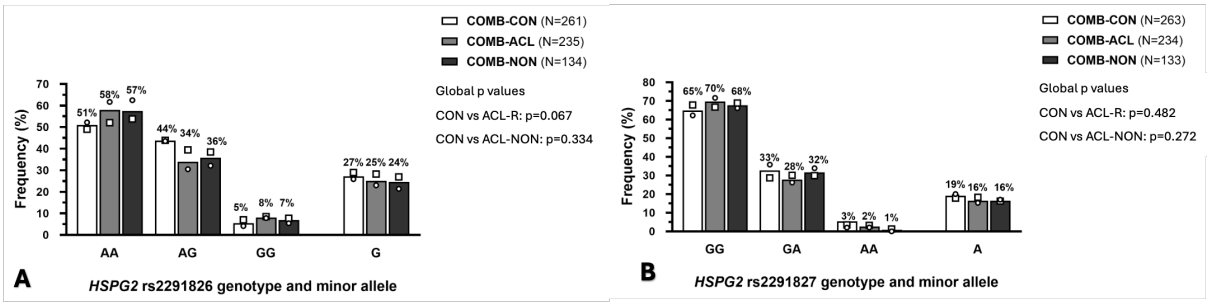

Figure 1 Genotype and allele frequency distributions of (A) *HSPG2* rs2291826 A/G and (B) *HSPG2* rs2291827 G/A variant for the combined cohort control group (COMB-CON: white bars), combined anterior cruciate ligament rupture group (COMB-ACLR: light grey bars), and ACL rupture by non-contact mechanism subgroup (COMB-NON: dark grey bars). The differences in genotype and allele distributions within each individual cohort are annotated as follows: Poland (circle) and Sweden (square blocks). Details of the HWE and AIC values can be found in Table S5.

Relationship between Genetic variants and Anthropometric Factors

A positive correlation in mass and BMI was found in all individual cohorts; COMB (r = 0.78), (POL: r = 0.81), and SWD: r = 0.78) respectively Figure 3. BMI was consistently influenced by mass in all the cohorts and height was similarly distributed between cases and controls between the cohorts. There appears to be a consistent correlation between sex with mass and height in all cohorts, with a negative correlation in the Swedish cohort (r = -0.43) (Figure 3C). The combined (r = 0.73) and Polish cohorts (r = 0.68) had a strong positive correlation between sex and height cohort while a strong negative correlation was noted in the Swedish cohort (r = -0.75), which may be ascribed to the sex-specific differences in height noted between the cohorts. A positive correlation was found between COR and height (r = 0.25) in the COMB cohort, which could possibly be influenced by the height of the Polish participants. No positive correlation between the *HSPG2*: rs2291826 A>G and rs2291827 G>A was noted with the combined and Polish co-

Table 2 Genotype effect on descriptive characteristics for the *HSPG2*: rs2291826 A>G, and rs2291827 G>A genetic variants in the combined cohort (COMB).

		Age (yrs.)	p value	Height (cm)	p value	Mass (kg)	p value	BMI (kg/m2)	p value	Sex (%M)	p value
HSPG2 rs2291826 A/G	A/A	31.8 ±12.4 (267)		175.7±9.97 (258)		74.2 ±14.0 (254)		23.1±5.8 (250)		61.0 (164)	
	A/G	32.6 ±13.7 (192)	0.199	175.1±9.9 (186)	0.266	72.6±12.7 (183)	0.083	22.8±5.3 (182)	0.345	57.7 (112)	0.705
	G/G	36.1±15.1 (33)		178.2±9.2 (31)		78.2±14.7 (32)		24.3±3.2 (30)		63.6 (21)	
HSPG2 rs2291827 G/A	G/G	32.6±12.5 (330)		175.8±10.5 (320)		74.2 ±13.9 (313)		22.9±5.9 (310)		60.1 (200)	
	G/A	31.6 ±13.9 (151)	0.349	175.2±8.8 (144)	0.666	72.7±12.2 (144)	0.299	23.3±4.3 (142)	0.458	60.3 (91)	0.291
	A/A	36.9 ±15.5 (12)		173.9±5.6 (12)		77.9±19.7 (13)		24.8±4.5 (11)		38.5 (5)	

All variables except sex are expressed as mean±standard deviation with the number of participants presented in parentheses.
Sex is expressed as percentages with the number of participants written in parentheses.
HSPG2: Heparan sulphate proteoglycan 2.

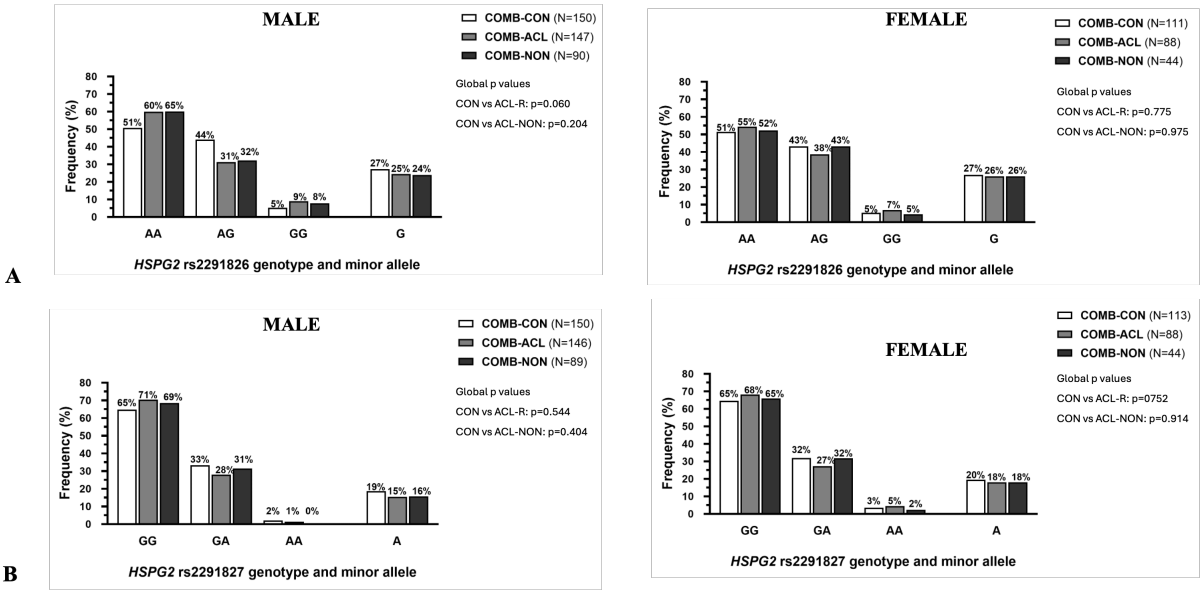

Figure 2 Genotype and allele frequency distributions of (A) *HSPG2* rs2291826 A/G and (B) *HSPG2* rs2291827 G/A variant for the combined cohort control group (COMB-CON: white bars), combined anterior cruciate ligament rupture group (COMB-ACLR: light grey bars), and ACL rupture by non-contact mechanism subgroup (COMB-NON: dark grey bars). Details of the HWE and AIC values can be found in Table S4 (males) and Table S5 (females).

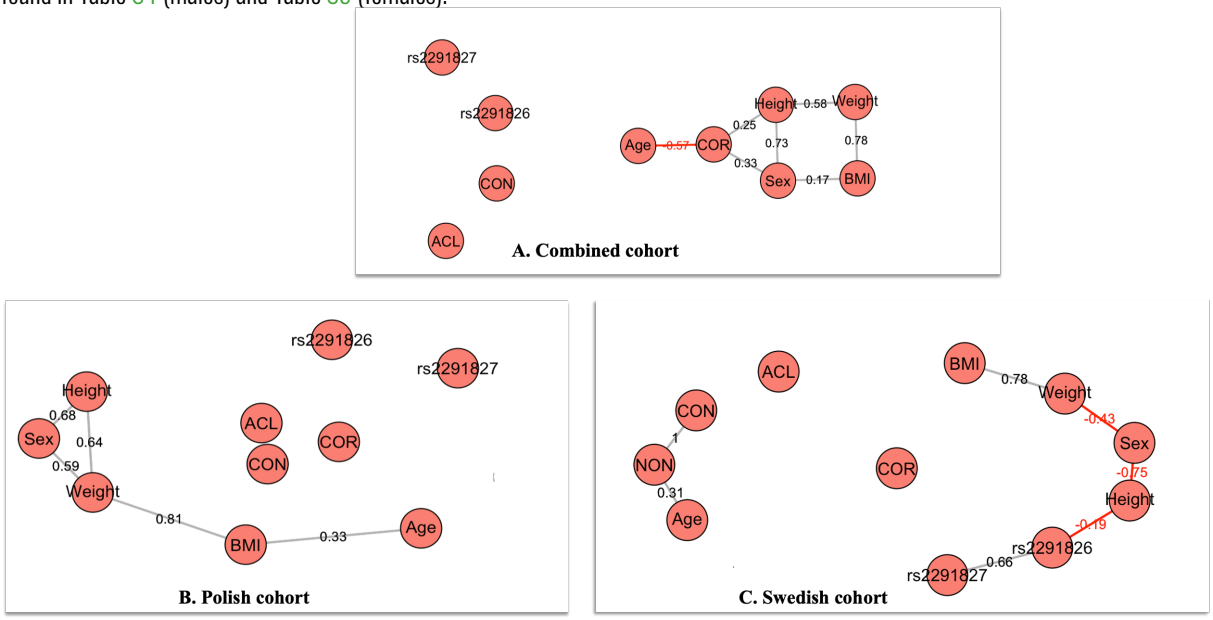

Figure 3 St Nicholas House analysis (SNHA) plot for the combined cohort and individual populations. The grey lines indicate a positive correlation and the red line a negative correlation (COR: country of recruitment, BMI: body mass index).

horts. However, a negative correlation of the *HSPG2* rs2261826 A/G with height (r = -0.19) was found in the Swedish cohort. A positive correlation was found between the two *HSPG2* rs2291826 A>G and rs2291827 G>A (r = 0.66) variants, which is expected, considering that the genetic variants are in a moderate positive linkage disequilibrium (D' = 0.703).

Discussion

The study aimed to investigate the relationship between genetic variants of *HSPG2*: rs2291826 A>G and rs2291827 G>A and anthropometric traits with the risk of ACL rupture in a combined group of participants recruited from Poland and Sweden. It was hypothesised that genetic variability at the HSPG2 loci together with variabil-

ity in height may collectively explain the contribution to susceptibility to ACL rupture. There was no significant association between anthropometric traits and the risk of ACL rupture. A genotype effect for height was found in the Swedish cohort for the *HSPG2* rs2291826 A>G variant, where genotypes A/A and A/G were frequently noted in the distribution. The SNHA plot also showed a negative correlation between genetic variant *HSPG2* rs2291826 A>G and height in the Swedish cohort, suggesting that individuals with genotypes A/A and A/G tend to be shorter.

In the combined cohort, significant associations in age and mass were found between the ACL rupture group and the control groups. The COMB-ACLR group was significantly older (p=0.036) and had a higher mass (p = 0.016) compared to the CON groups. However, when mass was adjusted for sex, significance was removed, suggesting that sex and mass may be confounding factors. The results are consistent with previous studies that emphasize the influence of age and mass on the risk of ACL rupture (Hurd et al. 2008; Adouni et al. 2024). Excessive mass can increase mechanical stress on the knee joint and change the kinematics of the joint, which can lead to knee injuries such as ACL ruptures (Adouni et al. 2024). In the individual cohorts, the Polish cohort was significantly younger compared to the Swedish cohort, with the POL-ACLR group having a higher mass and BMI compared to the CON group (p<0.001) (including non-contact ACL rupture). After adjusting for age, the significance of higher BMI was only maintained for the POL-NON subgroup. This suggests that age may play a role in the mass and BMI differences observed in the Polish cohort. Similar trends have been observed in previous studies where younger people tended to participate in high risk sporting activities, possibly predisposing them to injury (Motififard et al. 2024; Snaebjörnsson et al. 2019).

The SNHA revealed a strong positive correlation between mass and BMI in all cohorts: COMB cohort (r = 0.78), the SWD cohort (r = 0.78), and an even stronger correlation in the Polish cohort (r = 0.81) (Figure 3). This is to be expected as BMI is a mass-related index and the heights between the cases and controls were similar. However, the broader associations involving anthropometric traits such as mass and BMI, suggests these factors may influence ACL injury risk (Kızılgöz et al. 2019; Snaebjörnsson et al. 2019; Alsayed et al. 2023). A higher BMI can lead to excessive stress on the knee joint and can place additional strain on ligaments, cartilage, and tendons as well as decreased neuromuscular control around the knee joints, potentially increasing the risk of ACL ruptures (Widmyer et al. 2013; Snaebjörnsson et al. 2019; Motififard et al. 2024). When exploring the relationship between sex, height and mass, a negative correlation was found in the Swedish cohort, especially between sex and height (r = -0.75) and sex and mass (r = -0.43). However, a strong positive correlation between sex and height was also found in the combined cohort and the Polish cohort. The differences in the patterns could reflect the sex-specific differences in anthropometric traits influenced by the differences in the average height of individuals recruited in the two countries. In particular, males in the Swedish cohort were shorter and weighed less than males in the Polish cohort, which may have influenced the direction and strength of the observed correlation.

Although a strong association was found between anthropometric traits and the risk of ACL rupture, no correlation was found between the genetic variants *HSPG2*: rs2291826 A>G, rs2291827 G>A and anthropometric traits in the combined as well as in the Polish cohort. However, a mod-

erate negative correlation between genetic variant *HSPG2* rs2291826 A/G and height was found in the Swedish cohort (r=-0.019) (Figure 3C). A genotype effect on height related to *HSPG2* rs2291826 A>G was also noted in the Swedish cohort. Although the effect size is small, the results suggest that individuals with an A/A and A/G genotype of *HSPG2* rs2291826 A>G may be shorter. The findings are important because short individuals who weigh more (or have a higher BMI) may be at increased risk of ACL rupture due to the biomechanical load on the knee joint relative to stature. Although the HSPG2 rs2291826 variant may not directly influence ACL rupture risk, but it could have an indirect modifying effect on height. The *HSPG2* gene encodes perlecan, a multifunctional protein found in the basement membrane (Iozzo 2005; Farach-Carson and Carson 2007). Perlecan is mainly expressed in the cartilage matrix and plays an important role in regulating the availability and activities of growth factors such as fibroblast growth factors (FGF) and bone morphogenetic proteins (BMPs) (Iozzo 2005). These growth factors play an important role in the proliferation and differentiation of chondrocytes, an important process for skeletal development and endochondral ossification (Hayes et al. 2022). The exact biological mechanisms are still unclear and should be investigated in a larger cohort.

The study had several limitations, including the sample size of each cohort, in which controls and cases were not matched by sex and mass. In addition, the cohorts were not matched by age. The Polish cohort had younger participants, while the Swedish cohort had older participants. In addition, positive correlations of COR with height and sex could indicate possible recruitment bias or population-specific characteristics in the cohort, such as regional differences in height and sex distribution, which could have an influence on the analysis if not considered. Sport participation data were self-reported, and participants were not matched for participation in contact and non-contact sports.

Conclusion

The combined analysis of participant characteristics and the St. Nicholas house plot highlighted the role of anthropometric traits, particularly mass and BMI. The positive correlations between mass and BMI across all cohorts emphasize the importance of mass as a key factor in the risk of ACL rupture, while the sex-and population-specific differences in the relationships between height and mass suggest that ACL injury patterns are influenced by both genetic and environmental factors. There was no positive association between the genetic variants tested and anthropometric characteristics of participants. For this reason, it is hypothesised that for these two genetic variants tested, the variability in ACL risk is most likely not largely influenced by a genetic component of these measurements as influenced by *HSPG2*. The variability observed in the risk to ACL rupture can therefore be tested using the two variants *HSPG2* rs2291826 A>G and rs2291827 G

Acknowledgements

The statistical analysis and manuscript were the result of participating in the 7th International Summer School (University of Potsdam) funded by KoUP University of Potsdam, in Gülpe, Brandenburg, Germany from the 3rd of July to 12th July 2024. I extend great thanks to PD Dr. Christiane Scheffler, Dr. Detlef Groth, Prof. Michael Hermanussen and Prof. Barry Bogin for

accommodation and patience, their expertise and adventures they created to make the summer schools enjoyable. I extend utmost appreciation to my supervisors, Professor Alison September, Professor Malcolm Collins and Dr Mary-Jane Laguette, for the opportunity to participate in the Summer School and for their invaluable support and guidance during my MSc journey.

Funding statement

The National Research Foundation (NRF) supported L.Ramoshaba. The research study was supported by the South African Medical Research Council (SAMRC) through its Division of Research Capacity Development under the Mid-Career Scientist Programme, which is funded by the South African National Department of Health (SANDH).

References

Adouni, M./Aydelik, H./Faisal, T. R./Hajji, R. (2024). The effect of body weight on the knee joint biomechanics based on subject-specific finite element-musculoskeletal approach. Scientific Reports 14 (1), 13777. https://doi.org/10.1038/s41598-024-63745-x.

Alsayed, H. N./Alkhateeb, M. A./Aldossary, A. A./Houbani, K. M./Aljamaan, Y. M./Alrashidi, Y. A. (2023). Risk of anterior cruciate ligament injury in population with elevated body mass index. Medicinski Glasnik 20 (1). https://doi.org/10.17392/1517-22.

Arikawa-Hirasawa, E. (2022). Impact of the heparan sulfate proteoglycan perlecan on human disease and health. American Journal of Physiology. Cell Physiology 322 (6), C1117-C1122. https://doi.org/10.1152/ajpcell.00113.2022.

Bednarczyk, M./Stege, H./Grabbe, S./Bros, M. (2020). β2 Integrins-Multi-Functional Leukocyte Receptors in Health and Disease. International Journal of Molecular Sciences 21 (4). https://doi.org/10.3390/ijms21041402.

Bittencourt, N. F. N./Meeuwisse, W. H./Mendonça, L. D./Nettel-Aguirre, A./Ocarino, J. M./Fonseca, S. T. (2016). Complex systems approach for sports injuries: moving from risk factor identification to injury pattern recognition-narrative review and new concept. British Journal of Sports Medicine 50 (21), 1309–1314. https://doi.org/10.1136/bjsports-2015-095850.

Bruder, A. M./Culvenor, A. G./King, M. G./Haberfield, M./Roughead, E. A./Mastwyk, J./Kemp, J. L./Ferraz Pazzinatto, M./West, T. J./Coburn, S. L./Cowan, S. M./Ezzat, A. M./To, L./Chilman, K./Couch, J. L./Whittaker, J. L./Crossley, K. M. (2023). Let's talk about sex (and gender) after ACL injury: a systematic review and meta-analysis of self-reported activity and knee-related outcomes. British Journal of Sports Medicine 57 (10), 602–610. https://doi.org/10.1136/bjsports-2022-106099.

Cięszczyk, P./Willard, K./Gronek, P./Zmijewski, P./Trybek, G./Gronek, J./Weber-Rajek, M./Stastny, P./Petr, M./Lulińska-Kuklik, E./Ficek, K./Kemeryte-Riaubiene, E./Maculewicz, E./September, A. V. (2017). Are genes encoding proteoglycans really associated with the risk of anterior cruciate ligament rupture? Biology of Sport 34 (2), 97–103. https://doi.org/10.5114/biolsport.2017.64582.

Davis, A. M./Wong, R./Steinhart, K./Cruz, L./Cudmore, D./Dwyer, T./Li, L./Marks, P./McGlasson, R./Urquhart, N./Wilson, J. A./Nimmon, L./Ogilvie-Harris, D./Chahal, J. (2021). Development of an intervention to manage knee osteoarthritis risk and symptoms following anterior cruciate ligament injury. Osteoarthritis and Cartilage 29 (12), 1654–1665. https://doi.org/10.1016/j.joca.2021.08.011.

Dlamini, S. B./Saunders, C. J./Laguette, M.-J. N./Gibbon, A./Gamieldien, J./Collins, M./September, A. V. (2023). Application of an in silico approach identifies a genetic locus within ITGB2, and its interactions with HSPG2 and FGF9, to be associated with anterior cruciate ligament rupture risk. European Journal of Sport Science 23 (10), 2098–2108. https://doi.org/10.1080/17461391.2023.2171906.

Farach-Carson, M. C./Carson, D. D. (2007). Perlecan–a multifunctional extracellular proteoglycan scaffold. Glycobiology 17 (9), 897–905. https://doi.org/10.1093/glycob/cwm043.

Feldmann, D. (2022). Whole genome sequencing approach to identifying genetic risk factors underlying anterior cruciate ligament injuries in a twin family study. Faculty of Health Sciences, Department of Human Biology, 2022. Available online at http://hdl.handle.net/11427/36615.

Feldmann, D. C./Rahim, M./Suijkerbuijk, M. A. M./Laguette, M.-J. N./Cieszczyk, P./Ficek, K./Huminska-Lisowska, K./Häger, C. K./Stattin, E./Nilsson, K. G./Alvarez-Rumero, J./Eynon, N./Feller, J./Tirosh, O./Posthumus, M./Chimusa, E. R./Collins, M./September, A. V.

(2022). Investigation of multiple populations highlight VEGFA polymorphisms to modulate anterior cruciate ligament injury. Journal of Orthopaedic Research: Official Publication of the Orthopaedic Research Society 40 (7), 1604–1612. https://doi.org/10.1002/jor.25192.

Gibbon, A./Saunders, C. J./Collins, M./Gamieldien, J./September, A. V. (2018). Defining the molecular signatures of Achilles tendinopathy and anterior cruciate ligament ruptures: A whole-exome sequencing approach. PloS One 13 (10), e0205860. https://doi.org/10.1371/journal.pone.0205860.

González, J. R./Armengol, L./Solé, X./Guinó, E./Mercader, J. M./Estivill, X./Moreno, V. (2007). SNPassoc: an R package to perform whole genome association studies. Bioinformatics 23 (5), 644–645. https://doi.org/10.1093/bioinformatics/btm025.

Griffin, L. Y./Albohm, M. J./Arendt, E. A./Bahr, R./Beynnon, B. D./Demaio, M./Dick, R. W./Engebretsen, L./Garrett, W. E./Hannafin, J. A./Hewett, T. E./Huston, L. J./Ireland, M. L./Johnson, R. J./Lephart, S./Mandelbaum, B. R./Mann, B. J./Marks, P. H./Marshall, S. W./Myklebust, G./Noyes, F. R./Powers, C./Shields, C./Shultz, S. J./Silvers, H./Slauterbeck, J./Taylor, D. C./Teitz, C. C./Wojtys, E. M./Yu, B. (2006). Understanding and preventing noncontact anterior cruciate ligament injuries: a review of the Hunt Valley II meeting, January 2005. The American Journal of Sports Medicine 34 (9), 1512–1532. https://doi.org/10.1177/0363546506286866.

Gubbiotti, M. A./Neill, T./Iozzo, R. V. (2017). A current view of perlecan in physiology and pathology: A mosaic of functions. Matrix Biology: Journal of the International Society for Matrix Biology 57–58, 285–298. https://doi.org/10.1016/j.matbio.2016.09.003.

Hayes, A. J./Farrugia, B. L./Biose, I. J./Bix, G. J./Melrose, J. (2022). Perlecan, A Multi-Functional, Cell-Instructive, Matrix-Stabilizing Proteoglycan With Roles in Tissue Development Has Relevance to Connective Tissue Repair and Regeneration. Frontiers in Cell and Developmental Biology 10, 856261. https://doi.org/10.3389/fcell.2022.856261.

Hermanussen, M./Aßmann, C./Groth, D. (2021). Chain Reversion for Detecting Associations in Interacting Variables-St. Nicolas House Analysis. International Journal of Environmental Research and Public Health 18 (4). https://doi.org/10.3390/ijerph18041741.

Hewett, T. E./Myer, G. D./Ford, K. R./Paterno, M. V./Quatman, C. E. (2016). Mechanisms, prediction, and prevention of ACL injuries: Cut risk with three sharpened and validated tools. Journal of Orthopaedic Research: Official Publication of the Orthopaedic Research Society 34 (11), 1843–1855. https://doi.org/10.1002/jor.23414.

Huang, J./Li, X./Shi, X./Zhu, M./Wang, J./Huang, S./Huang, X./Wang, H./Li, L./Deng, H./Zhou, Y./Mao, J./Long, Z./Ma, Z./Ye, W./Pan, J./Xi, X./Jin, J. (2019). Platelet integrin αIIbβ3: signal transduction, regulation, and its therapeutic targeting. Journal of Hematology & Oncology 12 (1), 26. https://doi.org/10.1186/s13045-019-0709-6.

Hurd, W. J./Axe, M. J./Snyder-Mackler, L. (2008). Influence of age, gender, and injury mechanism on the development of dynamic knee stability after acute ACL rupture. The Journal of Orthopaedic and Sports physical Therapy 38 (2), 36–41. https://doi.org/10.2519/jospt.2008.2609.

Iozzo, R. V. (2005). Basement membrane proteoglycans: from cellar to ceiling. Nature Reviews. Molecular Cell Biology 6 (8), 646–656. https://doi.org/10.1038/nrm1702.

Joseph, A. M./Collins, C. L./Henke, N. M./Yard, E. E./Fields, S. K./Comstock, R. D. (2013). A multisport epidemiologic comparison of anterior cruciate ligament injuries in high school athletics. Journal of Athletic Training 48 (6), 810–817. https://doi.org/10.4085/1062-6050-48.6.03.

Kiapour, A. M./Murray, M. M. (2014). Basic science of anterior cruciate ligament injury and repair. Bone & Joint Research 3 (2), 20–31. https://doi.org/10.1302/2046-3758.32.2000241.

Kızılgöz, V./Sivrioğlu, A. K./Aydın, H./Ulusoy, G. R./Çetin, T./Tuncer, K. (2019). The Combined Effect of Body Mass Index and Tibial Slope Angles on Anterior Cruciate Ligament Injury Risk in Male Knees: A Case-Control Study. Clinical Medicine Insights. Arthritis and Musculoskeletal Disorders 12, 1179544119867922. https://doi.org/10.1177/1179544119867922.

Kobayashi, H./Kanamura, T./Koshida, S./Miyashita, K./Okado, T./Shimizu, T./Yokoe, K. (2010). Mechanisms of the anterior cruciate ligament injury in sports activities: a twenty-year clinical research of 1,700 athletes. Journal of Sports science & Medicine 9 (4), 669–675.

Lahiri, D. K./Nurnberger, J. I. (1991). A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. Nucleic Acids Research 19 (19), 5444. https://doi.org/10.1093/nar/19.19.5444.

Logerstedt, D. S./Scalzitti, D. A./Bennell, K. L./Hinman, R. S./Silvers-Granelli, H./Ebert, J./Hambly, K./Carey, J. L./Snyder-Mackler, L./Axe, M. J./McDonough, C. M. (2018). Knee Pain and Mobility Impairments: Meniscal and Articular Cartilage Lesions Revision 2018. The Journal of Orthopaedic and Sports physical Therapy 48 (2), A1-A50. https://doi.org/10.2519/jospt.2018.0301.

Lohmander, L. S./Englund, P. M./Dahl, L. L./Roos, E. M. (2007). The long-term consequence of anterior cruciate ligament and meniscus injuries: osteoarthritis. The American Journal of Sports Medicine 35 (10), 1756–1769. https://doi.org/10.1177/0363546507307396.

Lulińska-Kuklik, E./Leźnicka, K./Humińska-Lisowska, K./Moska, W./Michałowska-Sawczyn, M./Ossowski, Z./Maculewicz, E./Cięszczyk, P./Kaczmarczyk, M./Ratkowski, W./Ficek, K./Zmijewski, P./Leońska-Duniec, A. (2019). The VEGFA gene and anterior cruciate ligament rupture risk in the Caucasian population. Biology of Sport 36 (1), 3–8. https://doi.org/10.5114/biolsport.2018.78902.

Magnusson, K./Turkiewicz, A./Hughes, V./Frobell, R./Englund, M. (2020). High genetic contribution to anterior cruciate ligament rupture: Heritability ~69. British Journal of Sports Medicine. https://doi.org/10.1136/bjsports-2020-102392.

Martinez, J. R./Dhawan, A./Farach-Carson, M. C. (2018). Modular Proteoglycan Perlecan/HSPG2: Mutations, Phenotypes, and Functions. Genes 9 (11). https://doi.org/10.3390/genes9110556.

Meeuwisse, W. (1994). Assessing Causation in Sport Injury: A Multifactorial Model. Clinical Journal of Sport Medicine 4 (3), 166–170.

Mokone, G. G./Schwellnus, M. P./Noakes, T. D./Collins, M. (2006). The COL5A1 gene and Achilles tendon pathology. Scandinavian Journal of Medicine & Science in Sports 16 (1), 19–26. https://doi.org/10.1111/j.1600-0838.2005.00439.x.

Motififard, M./Akbari Aghdam, H./Ravanbod, H./Jafarpishe, M. S./Shahsavan, M./Daemi, A./Mehrvar, A./Rezvani, A./Jamalirad, H./Jajroudi, M./Shahsavan, M. (2024). Demographic and Injury Characteristics as Potential Risk Factors for Anterior Cruciate Ligament Injuries: A Multicentric Cross-Sectional Study. Journal of Clinical Medicine 13 (17). https://doi.org/10.3390/jcm13175063.

Murphy, D. F./Connolly, D. A. J./Beynnon, B. D. (2003). Risk factors for lower extremity injury: a review of the literature. British Journal of Sports Medicine 37 (1), 13–29. https://doi.org/10.1136/bjsm.37.1.13.

Murray, M. M./Martin, S. D./Martin, T. L./Spector, M. (2000). Histological changes in the human anterior cruciate ligament after rupture. The Journal of Bone and Joint Surgery. American Volume 82 (10), 1387–1397. https://doi.org/10.2106/00004623-200010000-00004.

Posthumus, M./September, A. V./O'Cuinneagain, D./van der Merwe, W./Schwellnus, M. P./Collins, M. (2010). The association between the COL12A1 gene and anterior cruciate ligament ruptures. British Journal of Sports Medicine 44 (16), 1160–1165. https://doi.org/10.1136/bjsm.2009.060756.

Rahim, M./Mannion, S./Klug, B./Hobbs, H./van der Merwe, W./Posthumus, M./Collins, M./September, A. V. (2017). Modulators of the extracellular matrix and risk of anterior cruciate ligament ruptures. Journal of Science and Medicine in Sport 20 (2), 152–158. https://doi.org/10.1016/j.jsams.2016.07.003.

Ribbans, W. J./September, A. V./Collins, M. (2022). Tendon and Ligament Genetics: How Do They Contribute to Disease and Injury? A Narrative Review. Life 12 (5). https://doi.org/10.3390/life12050663.

Schaid, D. J./Sinnwell, J. P./Jenkins, G. D./McDonnell, S. K./Ingle, J. N./Kubo, M./Goss, P. E./Costantino, J. P./Wickerham, D. L./Weinshilboum, R. M. (2012). Using the gene ontology to scan multilevel gene sets for associations in genome wide association studies. Genetic Epidemiology 36 (1), 3–16. https://doi.org/10.1002/gepi.20632.

Seneviratne, A./Attia, E./Williams, R. J./Rodeo, S. A./Hannafin, J. A. (2004). The effect of estrogen on ovine anterior cruciate ligament fibroblasts: cell proliferation and collagen synthesis. The American Journal of Sports Medicine 32 (7), 1613–1618. https://doi.org/10.1177/0363546503262179.

Shapiro, S. S./Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). Biometrika 52 (3/4), 591. https://doi.org/10.2307/2333709.

Snaebjörnsson, T./Svantesson, E./Sundemo, D./Westin, O./Sansone, M./Engebretsen, L./Hamrin-Senorski, E. (2019). Young age and high BMI are predictors of early revision surgery after primary anterior cruciate ligament reconstruction: a cohort study from the Swedish and Norwegian knee ligament registries based on 30,747 patients. Knee Surgery, Sports Traumatology, Arthroscopy: Official Journal of the ESSKA 27 (11), 3583–3591. https://doi.org/10.1007/s00167-019-05487-2.

Suijkerbuijk, M. A. M./Ponzetti, M./Rahim, M./Posthumus, M./Häger, C. K./Stattin, E./Nilsson, K. G./Teti, A./Meuffels, D. E./van der Eerden, B. J. C./Collins, M./September, A. V. (2019). Functional polymorphisms within the inflammatory pathway regulate expression of extracellular matrix components in a genetic risk dependent model for anterior cruciate ligament injuries. Journal of Science and Medicine in Sport 22 (11), 1219–1225. https://doi.org/10.1016/j.jsams.2019.07.012.

Suter, L. G./Smith, S. R./Katz, J. N./Englund, M./Hunter, D. J./Frobell, R./Losina, E. (2017). Projecting Lifetime Risk of Symptomatic Knee Osteoarthritis and Total Knee Replacement in Individuals Sustaining a Complete Anterior Cruciate Ligament Tear in Early Adulthood. Arthritis care & Research 69 (2), 201–208. https://doi.org/10.1002/acr.22940.

Widmyer, M. R./Utturkar, G. M./Leddy, H. A./Coleman, J. L./Spritzer, C. E./Moorman, C. T./DeFrate, L. E./Guilak, F. (2013). High body mass index is associated with increased diurnal strains in the articular cartilage of the knee. Arthritis and Rheumatism 65 (10), 2615–2622. https://doi.org/10.1002/art.38062.

Willard, K./Mannion, S./Saunders, C. J./Collins, M./September, A. V. (2018). The interaction of polymorphisms in extracellular matrix genes and underlying miRNA motifs that modulate susceptibility to anterior cruciate ligament rupture. Journal of Science and Medicine in Sport 21 (1), 22–28. https://doi.org/10.1016/j.jsams.2017.08.017.

Appendix

Table S1 Genotype effects on descriptive characteristics for the HSPG2 genetic variants: rs2291826 A>G and rs2291827 G>A, in the Polish cohort.

		Age (yrs.)	*p* value	Height (cm)	*p* value	Weight (kg)	*p* value	BMI (kg/m^2)	*p* value	Sex (% M)	*p* value
***HSPG2* rs2291826 A/G**	**A/A**	26.6±8.5 (164)		176.8±10.0 (161)		75.6±15.1 (162)		24.0±3.8 (159)		70.7 (116)	
	A/G	25.2±8.9 (107)	0.196	178.7±9.1 (107)	0.218	75.9±12.1 (107)	0.907	23.7±2.8 (107)	0.652	78.5 (84)	0.352
	G/G	28.8±1.9 (17)		179.4±10.1 (17)		77.2±10.4 (17)		24.0±2.8 (7)		76.5 (13)	
***HSPG2* rs2291827 G/A**	**G/G**	26.6±8.3(193)		177.6±10.3 (191)		76.3±14.9 (191)		24.0±3.7 (189)		72.0 (139)	
	G/A	25.2±8.9 (107)	0.483	178.7±9.1 (107)	0.953	75.9±12.1(107)	0.663	23.7±2.8 (107)	0.594	77.8 (70)	0.547
	A/A	27.3±9.2 (6)		176.5±6.1 (6)		73.5±10.7(6)		23.5±2.8 (6)		66.7 (4)	

All variables except sex are expressed as mean±standard deviation with the number of participants presented in parentheses.
Sex is expressed as percentages with the number of participants written in parentheses.
HSPG2: Heparan sulphate proteoglycan 2, and POL: Poland.

Table S2 Genotype effects on descriptive characteristics for the HSPG2 genetic variants: rs2291826 A>G, and rs2291827 G>A in the Swedish cohort.

		Age (yrs.)	*p* value	Height (cm)	*p* value	Weight (kg)	*p* value	BMI (kg/m^2)	*p* value	Sex (%M)	*p* value
***HSPG2* rs2291826 A/G**	**A/A**	40.1±13.0 (103)		173.8±9.8 (97)		71.6±11.3 (92)		21.4±7.9 (91)		45.7 (48)	
	A/G	41.9±13.0 (85)	0.448	170.2±8.9 (79)	**0.009**	67.9±12.1(76)	0.094	21.4±7.4 (75)	0.303	32 .2(28)	0.117
	G/G	43.9±14.4(16)		176.8±8.1 (14)		79.3±18.7(15)		24.7±3.8 (13)		50.0 (8)	
***HSPG2* rs2291827 G/A**	**G/G**	41.0±12.7 (137)		173.2±10.3 (129)		71.0±11.4 (122)		21.1±8.0 (121)		43.6 (61)	
	G/A	42.9±12.9 (95)	0.587	172.6±9.3 (90)	0.324	70.5±12.9 (90)	0.755	22.3±6.34(83)	0.826	34.4 (21)	0.177
	A/A	46.5±15.0 (6)		171.3±3.9 (6)		81.6±25.5 (7)		26.3±5.9 (5)		14 .3(1)	

Table S3 Genotype and minor allele frequency distributions, p-values, Hardy-Weinberg exact test (HWE), and Akaike information criterion (AIC) for HSPG2: rs2261826 A>G and rs2291827 G>A genetic variant in the combined cohort (males and females).).

		COMBINED						
MALE + FEMALE		CON %(n)	ACL-R %(n)	*p* value	AIC	NON %(n)	*p* value	AIC
HSPG2 rs2291826 (A/G)	n	261	235			134		
	AA	51.0 (133)	57.9 (136)			57.5 (77)		
	AG	43.7 (114)	34.0 (80)	0.067	686.8	35.8 (48)	0.334	510.7
	GG	5.4 (14)	8.1(19)			6.7 (9)		
	G	27.2 (143)	25.1 (118)	0.498		24.6 (66)	0.488	
	HWE	0.118	0.164			0.647		
HSPG2 rs2291827 (G/A)	n	263	234			133		
	GG	64.6 (170)	69.7 (163)			67.7 (90)		
	GA	32.7 (86)	27.8 (65)	0.482	691.8	31.6 (42)	0.272	509.7
	AA	2.7 (7)	2.6 (6)			0.8 (1)		
	A	19.1 (100)	16.5 (77)	0.332		16.5 (44)	0.451	
	HWE	0.423	1.000			0.125		

Genotype and allele frequencies are expressed as a percentage with the number of participants (n) in parentheses.
COMB-CON vs COMB-ACLR and COMB-CON vs COMB-NON p values are unadjusted.
P-values in bold typeset indicate significance p< 0.05.

Table S4 Genotype and minor allele frequency distributions, p-values, Hardy-Weinberg exact test (HWE), and Akaike information criterion (AIC) for HSPG2: rs2261826 A>G and rs2291827 G>A genetic variant in the combined cohort (only males).).

	COMBINED							
MALE		CON %(n)	ACL-R %(n)	p value	AIC	NON %(n)	p value	AIC
HSPG2 rs2291826 (A/G)	n	150	147			90		
	AA	50.7 (76)	59.9 (88)			60.0 (54)		
	AG	44.0 (66)	31.3 (46)	0.060	412.1	32.2 (29)	0.204	321.3
	GG	5.3 (8)	8.8 (13)			7.8 (7)		
	G	27.3 (82)	24.5 (72)	0.486		23.9 (43)	0.468	
	HWE	0.400	0.128			0.218		
HSPG2 rs2291827 (G/A)	n	150	146			89		
	GG	64.7 (97)	70.5 (103)			68.5 (61)	0.404	
	GA	33.3 (50)	28.1 (41)	0.544	415.1	31.5 (28)	0.404	318.6
	AA	2.0 (3)	1.4 (2)			0.0 (0)	0.403	
	A	18.7 (56)	15.4 (45)	0.345		15.7 (28)		
	HWE	0.411	0.534			0.579		

Genotype and allele frequencies are expressed as a percentage with the number of participants (n) in parentheses.
COMB-CON vs COMB-ACLR and COMB-CON vs COMB-NON p values are unadjusted.
P-values in bold typeset indicate significance *p*< 0.05.

Table S5 Genotype and minor allele frequency distributions, p-values, Hardy-Weinberg exact test (HWE), and Akaike information criterion (AIC) for HSPG2: rs2261826 A>G and rs2291827 G>A genetic variant in the combined cohort (only females).

		COMBINED						
FEMALE		CON %(n)	ACL-R %(n)	p value	AIC	NON %(n)	p value	AIC
HSPG2 rs2291826 (A/G)	n	111	88			44		
	AA	51.4 (57)	54.5 (48)			52.3 (23)		
	AG	43.2 (48)	38.6 (34)	0.775	278.7	43.2 (19)	0.975	190.9
	GG	5.4 (6)	6.8 (6)			4.5 (2)		
	G	27.0 (60)	26.1 (46)	0.932		26.1 (23)	0.986	
	HWE	0.664	0.750			0.382		
HSPG2 rs2291827 (G/A)	n	113	88			44		
	GG	64.6 (73)	68.2 (60)			65.9 (29)		
	GA	31.9 (36)	27.3 (24)	0.752	281.0	31.8 (14)	0.914	192.1
	AA	3.5 (4)	4.5 (4)			2.3 (1)		
	A	19.5 (44)	18.2 (32)	0.843		18.2 (16)	0.920	
	HWE	1.000	0.671			0.288		

Genotype and allele frequencies are expressed as a percentage with the number of participants (n) in parentheses.
COMB-CON vs COMB-ACLR and COMB-CON vs COMB-NON p values are unadjusted.
P-values in bold typeset indicate significance $p < 0.05$.