Functional COL1A1 variants are associated with the risk of acute musculoskeletal soft tissue injuries

Andrea Gibbon1| Stuart M. Raleigh2 | William J. Ribbans3 | Michael Posthumus1 | Malcolm Collins1 | Alison V. September1

[1]Division of Exercise Science and Sports Medicine, Department of Human Biology, Health Through Physical Activity, Lifestyle and Sport Research Centre, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa
[2]Centre for Sport, Exercise and Life Science, School of Life Science, Coventry University, Coventry, UK
[3]Centre for Physical Activity and Chronic Disease, Institute of Health and Wellbeing, University of Northampton, Northampton, UK

Correspondence
Alison V. September, Division of Exercise Science and Sports Medicine, Health Through Physical Activity, Lifestyle and Sport Research Centre, Sport Science Institute of South Africa, 3rd Floor, Boundary Rd, Newlands, Cape Town 7725, South Africa.
Email: alison.september@uct.ac.za

Abstract

Studies have reported the association of the COL1A1 Sp1 binding site variant (rs1800012) with the risk of acute musculoskeletal soft tissue injuries. Inter- action with the COL1A1 promoter variant (rs1107946) has also been proposed to modulate acute injury risk. Conversely, neither of these loci have been associated with chronic musculoskeletal soft tissue phenotypes. Therefore, the primary aim of this study involved characterizing these variants in a cohort of participants with chronic Achilles tendinopathy. Second, this study aimed to support the contribution of the rs1107946 and rs1800012 variants to the profile predisposing for acute musculoskeletal soft tissue injuries including Achilles tendon and anterior cruciate ligament (ACL) ruptures. A hypothesis- driven association study was conducted. In total, 295 control participants, 210 participants with clinically diagnosed Achilles tendinopathy, and 72 parti- cipants with Achilles tendon ruptures recruited independently from South Africa and the United Kingdom were genotyped for the prioritized variants. In addition, a cohort including 232 control participants and 234 participants with surgically diagnosed ACL ruptures was also analyzed. Although no associations were observed in the recruited cohorts, the rare rs1800012 TT genotype was associated with decreased ACL injury risk when the results from the current study were combined with that from previously published studies (P = .040, OR: 2.8, 95% CI: 1.0-11.0). In addition, the G-T (rs1107946-rs1800012) inferred haplotype was associated with decreased risk for Achilles tendon ruptures. These results support previous observations and reiterate the heterogeneity of musculoskeletal phenlotypes whereby certain markers may be common to the predisposing profiles while others may be unique.

INTRODUCTION

Emerging research suggests that regular exercise is associated with well-being, promoting longevity delaying disease onset and improving the overall quality of life.[1] However, paralleled with the increase in physical activity is the adverse increase in the incidence of associated injuries such as musculoskeletal injuries.[2] Approximately, 30%-50% of all sports-related musculoskeletal soft tissue injuries affect tendons and ligaments,[3] with the incidence of these injuries remaining high despite increased injury awareness and some advances in preventative strategies.[4,5] Type I collagen is the major fibrillar collagen found in tendons and ligaments.[6] This heterotrimer molecule consists of two α1- and one α2- chains encoded by the COL1A1 (chr17q21.33) and COL1A2 (chr7q21.3) genes, respectively. Several COL1A1 mutations have been associated with severe musculoskeletal disorders including osteogenesis imperfecta and Ehlers-Danlos syndrome.[7,8] Moreover, several independent studies have demonstrated the association of two func- tional variants within the COL1A1 gene (rs1107946 and rs1800012) with the risk of acute sports-related musculoskeletal soft tissue injuries including shoulder dislocations and cruciate ligament (CL) and Achilles tendon ruptures (ATR) (Table 1).[9-14] Conversely, two studies investigating the possible association of these variants with the risk of chronic musculoskeletal soft tissue injuries including Achilles tendino- pathy (TEN) and tennis elbow have failed to demonstrate similar associations (Table 1).[15,16] These results suggest that there are important differences in the genetic profiles predisposing for muscu- loskeletal soft tissue injuries, whereby, some loci may be common to both chronic and acute injuries, while other loci may be unique.

Therefore, this study aimed to test this hypothesis by repeating the association of the COL1A1 rs1107946 (G>T) and rs1800012 (G>T) variants with the risk of ACL rupture in a cohort of self- identified white individuals. In addition, the potential contribution of these variants to the genetic profile predisposing for acute Achilles tendon injuries and chronic Achilles TEN was explored. The final aim involved conducting combined statistical analyses, in which the data generated in several previously published studies were combined with the data generated in the current study.

MATERIALS AND METHODS

Participant recruitment

This study followed a case-control genetic association study approach (level of evidence III). Three cohorts, which included banked DNA samples and information of ancestry matched control and case participants, were used for this study. All recruited participants self-identified as being of white European ancestry. The two cohorts' representative of Achilles tendon injuries were sampled independently from South Africa (SA) and the United Kingdom (UK) and were subsequently pooled to increase the statistical power of the study. The SA-Achilles tendon injury cohort consisted of 165 apparently healthy asymptomatic control participants (SA-CONAT), 123 participants with chronic Achilles SA-TEN and 47 participants with acute SA-ATR.[22,23] It must be noted that the rs1800012 variant was previously genotyped in the SA-Achilles tendon injury cohort and the result presented by Posthumus et al.[15] For this reason, this data was only included for the combined analyses. The UK-Achilles tendon injury cohort consisted of 130 apparently healthy asymptomatic control participants (UK-CON), 87 participants with chronic Achilles UK-TEN and 25 participants with acute UK-ATR.[24] The single ACL rupture cohort consisted of 232 ap- parently healthy asymptomatic control participants (SA-CONACL) and
233 participants with clinically diagnosed ACL ruptures (SA-ACL). Moreover, 135 participants reported a noncontact mechanism of injury (SA-NON).[10,25] Similar to the participants in the SA-Achilles tendon injury cohort who were previously genotyped for the specificity-1 (Sp1) binding site variant, 130 SA-CONACL and 117 SA-ACL participants were genotyped for rs1800012 in the SA-ACL rupture cohort and the results presented by Posthumus et al.[10] Since publication of this initial manuscript, an additional 102 controls and 115 cases were recruited and analyzed in this study.

The control participants within the SA cohorts were recruited from various sports clubs and gyms around the greater Cape Town area. Case participants were recruited from the Sports Science Orthopaedic Clinic at the Sports Science Institute of South Africa and other clinical practices in Cape Town between 2004 and 2016. Control

participants in the UK-Achilles tendon injury cohort were recruited from the East Midlands region of the UK, whereas affected individuals were recruited from The County Clinic in Northampton, UK (2011-2014). Participants were included in the study based on the inclusion and exclusion criteria as first presented by Mokone et al[23] and Posthumus et al.[26] All participants provided written informed consent in accordance with the Declaration of Helsinki and completed questionnaires detailing their personal particulars and medical and sporting histories. This study was approved by the Human Research Ethics Committee (159/2011, 164/2006, and 370/2013) and the review board at the collaborating institution (University of Northampton).

DNA extraction

Through venepuncture of the forearm vein, each participant in the SA-Achilles tendon injury and ACL rupture cohorts provided approximately 5 mm of venous blood, which was collected in an ethylenediaminetetraacetic acid vacutainer tube. The blood samples were stored at −20°C until the total DNA extraction was performed using the protocol described by Lahiri and Nurnberger[27] and later modified by Mokone et al.[23] Approximately, 2 mm of saliva was collected from each participant in the British cohort using the OrageneTM OG-500 DNA collection tubes. Participant DNA was subsequently extracted using the prepIT DNA extraction kit (DNA Genotek Inc, Ontario, Canada) as previously described by Rickaby et al.[24] Samples yielded A260/A280 ratios between 1.8 and 2.0. With reference to the UK-cohort, all DNA extractions were conducted at the University of Northampton (UK).

SNP selection and genotyping

The two functional COL1A1 variants, rs1107946 (G>T) and rs1800012 (G>T), previously associated with the risk of musculoskeletal soft tissue injuries were prioritized for genotyping in the independent cohorts. The rs1107946 variant maps to the distal promoter region of the gene, 1997 bp upstream from the transcription start site. The rs1800012 variant locates 1245 bp downstream of the transcription start site within intron one and overlaps the binding site of the Sp1 transcription factor (Figure 1).
All participants were genotyped for the selected COL1A1 variants using inventoried Taqman genotyping assays (Applied Biosystems, Foster City, CA). The polymerase chain reaction (PCR) reactions were performed using the Applied Biosystems StepOnePlus and QuantStudio 3 Real-Time PCR systems (Applied Biosystems) following the manufacturer's recommendations. For quality control purposes, negative (DNA-free) and positive controls (samples of known genotype) were included in each 96-well PCR plate during the genotyping process. Two investigators confirmed genotype calls with all laboratory work conducted at the Division of Exercise Science and Sports Medicine at the University of Cape Town. This study was conducted in accordance with the recommendations of the STREGA Statement for reporting the results of genetic association studies.[28]

Statistical analysis

The QUANTO v1.2.4 software (http://biostats.usc.edu/software) was used to determine the statistical power of all analyses based on the respective sample sizes. Assuming "risk" allele frequencies between 0.2 and 0.8, a sample size of 100 cases and 200 controls would be sufficient to detect an allelic odds ratio of ≥2.0 at a power of 80% and a significance level of 5%. All statistical analyses were conducted in the R programming environment using specific R packages.[29] Quantitative variables were expressed as means and standard deviations, whereas qualitative characteristics were presented as frequencies. One-way analysis of variance was used to determine characteristic differences between the diagnostic groups and the possible genotype effects on age, sex, body mass, height, and body mass index (BMI). The Genetics and SNPassoc packages were used to determine genotype and allele frequencies, in addition to Hardy-Weinberg equilibrium (HWE) probabilities.[30,31] HWE P values are provided as exact P values. Logistic regression models testing the association of gene variants with the risk of musculoskeletal soft tissue injuries were adjusted for potential nongenetic confounders. Confounding parameters were equally weighted in the logistic regression models and entered by means of a forward method. In the instance when the genotype categories demonstrated frequencies less than 10%, a Fisher's exact test was conducted. The haplo.stats package was used to test the association between inferred haplotypes and the risk of musculoskeletal soft tissue

injury.[32] The strength of the association is reflected by the score statistic (Haplo.score). Negative values infer decreased risk, while positive scores infer an increased risk for injury. Statistical significance was accepted when P < .05. Adjustments for multiple testing were not performed on the genotype data as there is existing evidence for the association of both investigated COL1A1 variants with the risk of musculoskeletal soft tissue injuries.[9,10,14] Therefore, this study followed an evidence-based hypothesis approach and not a hypothesis-free approach in which there would be an argument for multiple testing.

RESULTS

Participant characteristics

The participants in the SA-Achilles tendon injury cohort were matched for sex, height, and country of birth (Table S1). However, The SA-CONAT participants were significantly younger at the time of re- cruitment (36.5 ± 10.9 years) compared to the age of participants in the SA-TEN (41.5 ± 14.0 years; P = .002) and SA-ATR (40.4 ± 11.9 years; P = .038) groups at the time of injury. In addition, the SA-CONAT participants weighed significantly less at the time of recruitment (72.1 ± 11.9 kg) compared to the SA-TEN (78.6 ± 17.5 kg; P < .001) and SA-ATR (86.0 ± 15.3 kg; P < .001). Consequently, the BMI of the SA-CONAT group (23.5 ± 2.8 kg/m2) was significantly lower than that of the SA-TEN (24.8 ± 3.3 kg/m2; P = .002) and SA-ATR (27.8 ± 4.0 kg/m2; P < .001) groups. The mean difference between age at the time of recruitment and age at the time of injury for the SA-TEN and SA-ATR groups was 6.9 ± 8.4 years and 6.8 ± 8.4 years, respectively. The participants in the UK-Achilles tendon injury cohort were matched for sex, height, body mass, BMI, and country of birth (Table S1). However, the age of participants at recruitment for the UK-CON group (41.6 ± 11.7 years) was significantly younger than the age at the time of injury for the UK-TEN group (45.2 ± 14.4 years; P = .043). The participants in the UK-TEN and UK-ATR groups were significantly older (8.3 ± 8.6 and 10.8 ± 11.3 years) and heavier (1.6 ± 6.0 and 5.0 ± 7.3 kg) at the time of recruitment than at the time of injury. When the SA- and UK-Achilles tendon injury cohorts were combined, only age was significantly different between the combined-CON group (38.7 ± 11.5 years) and combined-TEN group (43.2 ± 14.3 years; P < .001). On the other hand, the combined-CON group weighed significantly less (75.6 ± 16.1 kg; P < .001) and demonstrated lower BMI ratios (24.6 ± 3.9 kg/m2; P < .001) compared to the combined-ATR group (body mass: 84.1 ± 15.5 kg; BMI: 27.4 ± 4.0 kg/m2) (Table S1).

The participants in the SA-ACL rupture cohort were matched for the country of birth and height (Table S2). However, the SA-CONACL participants were significantly older (29.9 ± 11.8 years) at the time of recruitment compared to participants in the SA-ACL group (26.0 ± 10.8 years; P = .002) and SA-NON subgroup (26.3 ± 10.3 years; P = .004) at the time of injury. A significantly larger proportion of SA-ACL cases (73%; n = 171) including those reporting a non-contact mechanism of injury (75.6%; n = 102) were males compared to the SA-CONACL group (56%; n = 130; P < .001). Furthermore, the SA-CONACL participants weighed significantly less (73.5 ± 14.5 kg) with a corresponding lower BMI (23.8 ± 4.0 kg/m2) than the case participants in the SA-ACL group (body mass: 80.5 ± 16.9 kg; P < .001; BMI: 24.7 ± 5.9; P < .001) and SA-NON subgroup (body mass: 80.4 ± 16.1 kg; P < .001; BMI: 24.7 ± 5.4; P < .001). These differences remained significant after adjusting for age and sex (body mass SA-ACL: P = .002; SA-NON: P = 0.008; BMI SA-ACL: P = .001; SA-NON: P = .011). The SA-ACL participants were 4.6 ± 8.0 years older and 1.2 ± 4.0 kg heavier at the time of recruitment than at the time of the first ACL rupture. No significant genotype effects on participant characteristics including age, sex, body mass, height, and BMI were noted for the three investigated cohorts (Table S3).

Genotype and allele frequencies

Participants with clinically diagnosed chronic AT in the SA- and UK-Achilles tendon injury cohorts demonstrated similar genotype and allele frequency distributions for the COL1A1 rs1107946 and rs1800012 variants. Therefore, the genotype datasets from these cohorts were combined to increase the statistical power of this study. The genotype and allele frequency distributions of the rs1107946 variant did not differ significantly between the combined CON and TEN groups (P = .650) or the combined CON and ATR groups (P = 1.000) (Table 2). Similarly, the genotype frequencies for the rs1800012 variant were similarly distributed between the combined CON and TEN

groups (P = 1.000) and combined CON and ATR groups (P = .106) (Table 2). Of worth noting was that no participants within the ATR group demonstrated the rare TT genotype.

With respect to the SA-ACL rupture cohort, the genotype frequencies for the rs1107946 and rs1800012 variants were similarly distributed between the CON and ACL groups (P = .730 and P = .605, respectively) (Table 3). In addition, the rs1107946 and rs1800012 variants demonstrated similar genotype frequency distributions between the CON group and NON subgroup (P = 1.000 and P = 1.000, respectively) (Table 3).

However, an additional analysis, which involved combining results from several previously conducted studies of participants of European ancestry, with those presented for ACL ruptures in the current study, demonstrated the significant association of the rs1800012 variant with the risk of ACL ruptures.[9,11,12,20] Specifically, the rs1800012 TT genotype (TT vs GT+GG) was significantly over- represented in the group of controls (2.7%; n = 39) compared to the group of cases with ACL ruptures (0.98%; n = 4; P = .040; OR: 2.8; 95% CI: 1.0-11.0) (Figure 2). Moreover, when CL ruptures were considered; including both ACL and posterior CL ruptures, the rs1800012 TT genotype (TT vs GT+GG) was similarly associated with decreased risk of injury (CON: 2.7%; CL: 0.78%; P = .003; OR: 3.8; 95% CI: 1.4-11.7) (Figure 2).

Haplotype analysis

Inferred haplotypes were constructed using the genotype data from the two COL1A1 variants, rs1107946 (G>T) and rs1800012 (G>T). Unfortunately, combined haplotype analysis was not performed in this study as the raw data from the previously published studies were not available. With respect to the data generated in the current study, three allele combinations (G-G, T-G, and G-T) were inferred with a frequency greater than 10% in all investigated cohorts. None of the participants in any of the cohorts presented with the rare T-T inferred haplotype. There were no significant differences in the distribution of the three inferred COL1A1 haplotypes between the combined-CON and -TEN groups (P = .816) (Figure 3A).

However, the G-T inferred haplotype was significantly over- represented in the combined-CON group (16%; n = 35) compared to the combined-ATR group (8%; n = 4; P = .022; Haplo.score: −2.287) (Figure 3A). Conversely, none of the haplotypes constructed from the two investigated COL1A1 loci were significantly associated with the risk of ACL ruptures in the SA-ACL rupture cohort (ACL: P = 0.749; NON: P = .929) (Figure 3B).

DISCUSSION

This study follows several published studies that have investigated the contribution of the COL1A1 rs1107946 and rs1800012 variants to the profile predisposing for acute and chronic musculoskeletal soft tissue injuries (Table 1). These specific COL1A1 variants have been highlighted as plausible candidates based on their functionality and the fact that the protein encoded by the COL1A1 gene is the pre-dominant fibrillar collagen in noncartilaginous connective tissues.

With a specific focus on the studies profiling tendon and ligament injuries, emerging evidence suggests that the genetic signature predisposing for chronic phenotypes mapping to the COL1A1 rs1800012 locus may contrast with that underpinning acute phenotypes. For example, the rare rs1800012 TT genotype has frequently been associated with reduced risk of acute musculoskeletal soft tissue injuries including shoulder dislocations and CL ruptures, specifically in white populations.[9,10] On the contrary, other studies have failed to demonstrate an association between this variant and the risk of chronic musculoskeletal soft tissue injuries including Achilles TEN and tennis elbow.[15,16]

Recently, Rodas et al[33] used a multivariate modeling approach to predict risk for tendinopathies in elite athletes. In line with the earlier case-control genetic association studies, neither rs1800012 nor rs1107946 was identified as robust markers for injury prediction.

In accordance, the current study failed to highlight a significant association with injury risk for chronic Achilles TEN and the investigated COL1A1 variants. However, the rs1800012 variant was associated with the risk of acute musculoskeletal soft tissue injuries. Specifically, the rare rs1800012 TT genotype was associated with a significant decrease in risk for ACL ruptures when the results from the current study were combined with previously published data.[9,11,20]

Although the rs1800012 T allele is more common in the European population (19%) compared to non-European populations (African: 6%, East Asian: 0%, and South Asian: 10% [www.ensembl. org]), the rs1800012 TT genotype is still rare. Consequently, many independent studies, especially those investigating non-European populations, have been underpowered to effectively study the association of this variant with the risk of injury.[17,19] In addition, the rare TT genotype, may also explain the lack of association in the independent SA-ACL rupture cohort investigated in the current study and other studies reporting small sample sizes. Therefore, these results highlight the importance of testing the association of rare variants in larger cohorts, as performed in the combined analyses of this study.

Besides, the independent association of the rs1800012 variant with the risk of acute musculoskeletal soft tissue injuries, Ficek et al[11] previously described the association of the G-T (rs1107946-rs1800012) inferred haplotype with reduced risk for ACL rupture. This haplotype was similarly associated with reduced risk for acute ATR in the current study. However, due to the small number of participants with ATR, this result would require validation in a larger sample set. Although no association of the G-T inferred haplotype was reported to alter the risk of ACL ruptures, it was not possible to analyze previously published haplotype data with that produced in the current study.

In comparison to rs1800012 variant, the COL1A1 promoter variant, rs1107946, is less well characterized in the profiles predisposing for musculoskeletal soft tissue injuries. To the best of our knowledge, this variant has not been profiled for its contribution to the risk of chronic tendon and ligament injuries. Furthermore, this variant has demonstrated contrasting associations with ACL injuries sustained through different mechanisms. Specifically, this variant was not independently associated with the risk of ACL ruptures in a group of professional soccer players of Polish ancestry or with the risk of ACL injury in the SA-ACL rupture cohort presented in the current study,[11] yet the rs1107946 GG genotype was associated with increased risk for ACL ruptures sustained during skiing activities.[13] Skiers commonly present with ACL ruptures; however, the mechanisms resulting in these injuries are suggested to be unique to the sport. It may be proposed that external factors such as ground surface and equipment may influence the outcome of skiing-associated ACL injuries to a greater degree to that of noncontact ACL ruptures sustained during movement on land, which are believed to occur because of an individual's own movement with no external perturbation. Therefore, the association of the COL1A1 rs1800012 and rs1107946 variants with the risk of ACL ruptures in skiers requires validation in other independent cohorts.[12,13] The rs1107946 T allele has a higher frequency in non-European populations (African: 31%, East Asian: 30%, and South Asian: 24%) and is, therefore, more informative at inferring risk in these populations compared to that in the European population for which the allele frequency is only 14% (www.ensembl.org).

These COL1A1 variants have also been explored functionally and have been proposed to influence DNA binding, gene transcription, and protein production.[34–36] The rs1107946 variant demonstrates allele-dependent transcriptional differences with the ancestral G allele associated with a higher level of transcription compared to the T allele, both independently and in combination with a neighboring promoter variant, rs11327935 (indel/T).[36] Similarly, the rare

rs1800012 T allele demonstrates an increased binding affinity for the Sp1 transcription factor using gel shift assays, in addition to demonstrating a threefold increase in the level of primary COL1A1 messenger RNA (mRNA) compared to the ancestral G allele.[34,35] The rs1800012 T allele is also associated with a higher α1 (I) to α2 (I) collagen protein ratio, which reflects the increased ratio of COL1A1 mRNA relative to COL1A2 (2.3:1).[34] This has been observed to result in the production of homotrimeric type I collagen molecules. Comparatively, type I collagen production by osteoblasts derived from individuals with the rs1800012 GG genotype demonstrated ratios close to the expected (2:1, respectively).[34] Furthermore, the strength and stiffness of bone derived from rs1800012 heterozygote individuals were significantly lower compared with that of bone derived from individuals with the common rs1800012 GG genotype.

Interestingly, Jin et al[35] reported that the level of COL1A1 transcription is influenced by the interaction of variants located within 5' regulatory region of the gene, rather than that of the individual variants, for which there were no consistent effects on gene transcription observed. These variant interactions were hypothesized to involve chromatin looping. Specifically, the COL1A1 G-del-T (rs1107946-rs11327935-rs1800012) inferred haplotype drove a significantly higher level of gene transcription compared to the other investigated allele combinations.[35]

However, it must be noted that the rs1800012 and rs11327935 variants are tagged loci (D': 1.000, r2: 1.000). Therefore, genotyping only one of the tagged loci as conducted in the current study is necessary. It was postulated that based on the effects on COL1A1 transcription and collagen type-I chain production, the G-del-T inferred haplotype reduced bone mineral density and consequently predisposed for osteoarthritis. Interestingly, this haplotype appears to have the opposite effect within soft musculoskeletal tissues, as suggested by the results presented in the current study and that by Ficek et al,[11] in which the G-T inferred haplotype (rs1107946-rs1800012) was associated with decreased risk of Achilles tendon and ACL ruptures. It may be postulated that an increase in homotrimeric type-I collagen molecules may increase the tensile strength of tendons and ligaments; thereby, reducing the risk of acute soft tissue ruptures.

In conclusion, this study contributes to the growing body of evidence implicating the COL1A1 Sp1 binding site variant in the genetic profile predisposing for acute musculoskeletal soft tissue injuries, specifically ACL ruptures. This study also provided evidence for the association of the COL1A1 G-T inferred haplotype (rs1107946-rs1800012) with the risk of ATR. This is the same haplotype previously associated with the risk of ACL ruptures. Moreover, this research has reiterated the heterogeneity of musculoskeletal soft tissue injuries whereby some loci predisposing for acute musculoskeletal soft tissue injuries may not necessarily be informative risk markers for chronic phenotypes. To understand the biological mechanisms underpinning these risk associations, it is important that we begin to explore the effects of these markers on the mechanical properties of the tissue. This is more likely achievable with the pooling of resources and expertise, in addition to the sharing of omics data generated through the use of next-generation sequencing technologies.[37] Lastly, with the generation of more genomic data, researchers should aim to develop integrative multi- variate models; expeditions which are more likely to reveal robust risk conferring signatures for both acute and chronic phenotypes.

CONFLICT OF INTERESTS
The authors declare that there are no conflict of interests.

AUTHOR CONTRIBUTIONS

REFERENCES

1 Penedo FJ, Dahn JR. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. Curr Opin Psychiatry. 2005;18:189-193.

2 Morrow JR Jr, Defina LF, Leonard D, Trudelle-Jackson E, Custodio MA. Meeting physical activity guidelines and musculoskeletal injury: the WIN study. Med Sci Sports Exerc. 2012;44: 1986-1992.

3 Järvinen TA, Kannus P, Maffulli N, Khan KM. Achilles tendon dis- orders: etiology and epidemiology. Foot Ankle Clin. 2005;10:255-266.

4 Renstrom P, Ljungqvist A, Arendt E, et al. Non-contact ACL injuries in female athletes: an International Olympic Committee current concepts statement. Br J Sports Med. 2008;42:394-412.

5 de Jonge S, van den Berg C, de Vos RJ, et al. Incidence of midportion Achilles tendinopathy in the general population. Br J Sports Med. 2011; 45:1026-1028.

6 Amiel D, Frank C, Harwood F, Fronek J, Akeson W. Tendons and ligaments: a morphological and biochemical comparison. J Orthop Res. 1984;1:257-265.

7 Willing MC, Deschenes SP, Scott DA, et al. Osteogenesis imperfecta type I: molecular heterogeneity for COL1A1 null alleles of type I collagen. Am J Hum Genet. 1994;55:638-647.

8 D'alessio M, Ramirez F, Blumberg BD, et al. Characterization of a COL1A1 splicing defect in a case of Ehlers-Danlos syndrome type VII: further evidence of molecular homogeneity. Am J Hum Genet. 1991; 49:400-406.

9 Khoschnau S, Melhus H, Jacobson A, et al. Type I collagen α1 Sp1 polymorphism and the risk of cruciate ligament ruptures or shoulder dislocations. Am J Sports Med. 2008;36:2432-2436.

10 Posthumus M, September AV, Keegan M, et al. Genetic risk factors for anterior cruciate ligament ruptures: COL1A1 gene variant. Br J Sports Med. 2009;43:352-356.

11 Ficek K, Cieszczyk P, Kaczmarczyk M, et al. Gene variants within the COL1A1 gene are associated with reduced anterior cruciate ligament injury in professional soccer players. J Sci Med Sport. 2013;16:396-400.

12 Stępień-Słodkowska M, Ficek K, Eider J, et al. The +1245G/T polymorphism in the collagen type I alpha 1 (COL1A1) gene in polish skiers with anterior cruciate ligament injury. Biol Sport. 2013;30: 57-60.

13 Stępień-Słodkowska M, Ficek K, Kaczmarczyk M, et al. Influence of biological factors on injuries occurrence in the Polish population. Ann Agric Environ Med. 2016;23:315-318.

14 Collins M, Posthumus M, Schwellnus MP. The COL1A1 gene and acute soft tissue ruptures. Br J Sports Med. 2010;44:1063-1064.

15 Posthumus M, September AV, Schwellnus MP, Collins M. Investigation of the Sp1-binding site polymorphism within the COL1A1 gene in participants with Achilles tendon injuries and controls. J Sci Med Sport. 2009;12:184-189.

16 Erduran M, Altinisik J, Meric G, Ates O, Ulusal AE, Akseki D. Is Sp1 binding site polymorphism within COL1A1 gene associated with tennis elbow. Gene. 2014;537:308-311.

17 Pruna R, Artells R, Ribas J, et al. Single nucleotide polymorphisms associated with non-contact soft tissue injuries in elite professional soccer players: influence on degree of injury and recovery time. BMC Musculoskelet Disord. 2013;14:221.

18 Stepien-Slodkowska M, Ficek K, Zietek P, et al. Is the combination of COL1A1 gene polymorphisms a marker of injury risk? J Sport Rehabil. 2017;26:234-238.

19 Prabhakar S, John R, Dhillon MS, Anand A, Sharma K, Bammidi S. Are COL1A1 gene polymorphisms associated

with anterior cruciate ligament tear in the Indian population? Results of a preliminary case-control study. Muscles Ligaments Tendons J. 2018;8:15-22.

20 Sivertsen EA, Haug KBF, Kristianslund EK, et al. No association between risk of anterior cruciate ligament rupture and selected candidate collagen gene variants in female elite athletes from high-risk team sports. Am J Sports Med. 2019;47:52-58.

21 Wang C, Li H, Chen K, Wu B, Liu H. Association of polymorphisms rs1800012 in COL1A1 with sports-related tendon and ligament injuries: a meta-analysis. Oncotarget. 2017;8:27627-27634.

22 Mokone GG, Gajjar M, September AV, et al. The guanine-thymine dinucleotide repeat polymorphism within the tenascin-C gene is associated with achilles tendon injuries. Am J Sports Med. 2005;33: 1016-1021.

23 Mokone GG, Schwellnus MP, Noakes TD, Collins M. The COL5A1 gene and Achilles tendon pathology. Scand J Med Sci Sports. 2006;16: 19-26.

24 Rickaby R, El Khoury L, Ribbans WJ, Raleigh SM. Variation within three apoptosis associated genes as potential risk factors for Achilles tendinopathy in a British based case-control cohort. Gene. 2015;571: 167-171.

25 Mannion S, Mtintsilana A, Posthumus M, et al. Genes encoding proteoglycans are associated with the risk of anterior cruciate liga- ment ruptures. Br J Sports Med. 2014;48:1640-1646.

26 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. Am J Sports Med. 2009;37:2234-2240.

27 Lahiri DK, Nurnberger JI Jr. A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. Nucleic Acids Res. 1991;19:5444.

28 Little J, Higgins JPT, Ioannidis JPA, et al. STrengthening the REporting of genetic association studies (STREGA)–an extension of the STROBE statement. Eur J Clin Invest. 2009;39:247-266.

29 R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.

30 Gonzalez R, Armengol L, Guino E, et al. 2014. SNP-based whole genome association studies. Version 1.9-2.

31 Warnes G, Gorjanc G, Leisch F, Man M 2013. Population genetics. Version 1.3.8.1.

32 Sinnwell J, Schaid D 2015. Statistical analysis of haplotypes with traits and covariates when linkage is ambiguous. Version 1.7.1.

33 Rodas G, Osaba L, Arteta D, Pruna R, Fernández D, Lucia A. Genomic prediction of tendinopathy risk in elite team sports. Int J Sports Physiol Perform. 2019;14:1-7.

34 Mann V, Hobson EE, Li B, et al. A COL1A1 Sp1 binding site polymorphism predisposes to osteoporotic fracture by affecting bone density and quality. J Clin Invest. 2001;107:899-907.

35 Jin H, van't Hof RJ, Albagha OM, Ralston SH. Promoter and intron 1 polymorphisms of COL1A1 interact to regulate transcription and susceptibility to osteoporosis. Hum Mol Genet. 2009;18: 2729-2738.

36 Garciagiralt N, Enjuanes A, Bustamante M, et al. In vitro functional assay of alleles and haplotypes of two COL1A1-promoter SNPs. Bone. 2005;36:902-908.

37 Gibbon A, Saunders CJ, Collins M, Gamieldien J, September AV. Defining the molecular signatures of Achilles tendinopathy and anterior cruciate ligament ruptures: a whole-exome sequencing approach. PLoS One. 2018;13(10):e0205860.