



VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

Gene variants within the COL1A1 gene are associated with reduced anterior cruciate ligament injury in professional soccer players

This is the Accepted version of the following publication

Ficek, Krzysztof, Cieszczyk, Pawel, Kaczmarczyk, Mariusz, Maciejewska-Karłowska, Agnieszka, Sawczuk, Marek, Cholewinski, Jerzy, Leonska-Duniec, Agata, Stepien-Słodkowska, Marta, Zarebska, Aleksandra, Stepto, Nigel, Bishop, David and Eynon, Nir (2013) Gene variants within the COL1A1 gene are associated with reduced anterior cruciate ligament injury in professional soccer players. *Journal of Science and Medicine in Sport*, 16 (5). pp. 396-400. ISSN 1440-2440 (print) 1878-1861 (online)

The publisher's official version can be found at
<http://www.sciencedirect.com/science/article/pii/S1440244012002058>
Note that access to this version may require subscription.

Downloaded from VU Research Repository <https://vuir.vu.edu.au/23466/>

**Gene variants within the *COL1A1* gene are associated with reduced Anterior Cruciate
Ligament injury in professional soccer players**

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Abstract

Objective: To examine the association of *COL1A1*-1997G/T and +1245G/T polymorphisms, individually and as haplotypes, with ACL ruptures in professional soccer players. **Design:** Subjects were 91 male professional soccer players with surgically diagnosed primary ACL ruptures. The control group consisted of 143 apparently healthy male professional soccer players, who were without any self-reported history of ligament or tendon injury. Both subjects and healthy controls are from the same soccer teams, of the same ethnicity (Polish, East-Europeans for ≥ 3 generations), a similar age category, and had a comparable level of exposure to ACL injury.

Methods: Genomic DNA was extracted from the oral epithelial cells using GenElute Mammalian Genomic DNA Miniprep Kit (Sigma, Germany). All samples were genotyped using a Rotor-Gene real-time polymerase chain reaction (PCR). **Results:** Genotype distributions for both polymorphisms met the Hardy-Weinberg expectations in both subjects and controls ($p > 0.05$). Higher frequency of the *COL1A1* G-T (-1997G/T and +1245G/T polymorphisms) haplotype was significantly associated with reduced risk for ACL rupture (Hap.score -1.98, $p = 0.048$). The TT genotype was under-represented in the ACL rupture group. However, this result was not statistically significant ($p = 0.084$ Fisher's exact test, recessive mode: TT vs GT+GG).

Conclusions: Higher frequency of the *COL1A1* G-T haplotype is associated with reduced risk of ACL injury in a group of professional soccer players. Consequently, carrying two copies the *COL1A1* G-T haplotype may be protective against ACL injury.

Key words: ACL, Sport injuries, Soccer, Athletes, Genetic Polymorphism, *COL1A1*

1 **1. Introduction**

2 Soccer is an intermittent team sport with high physiological demands. Professional players cover
3 8,000–12,000 m during a match, with up to 20% of the overall distance corresponding to maximal
4 or near maximal running efforts.^{1, 2} Rapid movements such as jumping and tackling are also
5 frequently performed by players during a match.^{3,4} Given the high physiological demands, it is not
6 surprising that soccer is also associated with a relatively large number of injuries. The incidence
7 of soccer-related injuries is estimated to be 10-35 per 1000 hours of exposure in adult male soccer
8 players,^{5, 6} with approximately 60-80% of injuries occurring in the lower extremities, most
9 commonly at the knee or ankle.⁷

10 The anterior cruciate ligament (ACL) rupture is one of the most severe musculoskeletal
11 soft tissue injuries in professional sport.^{8, 9} ACL ruptures are complex, multifactorial disorders
12 determined by the interaction of extrinsic and intrinsic risk factors.¹⁰ The familial aggregation
13 observed in ACL injuries^{11, 12} has prompted researchers to investigate a possible genetic linkage.

14 Genetic polymorphisms within the major alpha chains of the collagen type I gene
15 (*COL1A1*), which is located on chromosome 17q21, have been shown to influence the
16 predisposition for ACL rupture in non-athletic cohorts.¹³ In particular, the functional *COL1A1* Sp1
17 binding site polymorphism (*COL1A1* Sp1 +1245G/T, rs1800012), initially described in 1996,¹⁴
18 has been associated with the risk for ACL injury. The uncommon *COL1A1* Sp1 TT genotype is
19 significantly underrepresented in participants with ACL ruptures, and therefore a higher frequency
20 of *COL1A1* TT genotype was proposed to have a protective effect against ACL rupture.^{9, 15, 16} A
21 second polymorphism (G/T) has been identified in the proximal promoter of *COL1A1*, at position
22 -1997 (rs1107946) relative to the transcription start site, which is in linkage disequilibrium with
23 the Sp1 polymorphism.¹⁷ This *COL1A1*-1997 G/T polymorphism has been associated with bone
24 mineral density (BMD), and has also been reported to interact with the *COL1A1* Sp1 +1245G/T
25 polymorphism to regulate BMD.¹⁷⁻¹⁹ Furthermore, a functional significance of the *COL1A1*-1997
26 G/T polymorphism was demonstrated in the in-vitro regulation of the *COL1A1* gene in
27 osteoblasts; the G allele showed a higher transcriptional activity than the T allele.²⁰ Despite its
28 functionality, thus far, the potential influence of the *COL1A1*-1997 G/T polymorphism on the

1 incidence of ACL rupture injury has never been studied.

2 Recently, a transcription analysis that included *COL1A1* -1997G/T, and *COL1A1* Sp1
3 +1245G/T polymorphisms in the 5' flank of *COL1A1* revealed that the levels of transcription are
4 influenced by haplotype (i.e., a combination of alleles at adjacent locations on the chromosome
5 that are transmitted together), rather than by the genotype, at individual polymorphic sites,
6 indicating that *COL1A1* polymorphisms interact with each other to form a haplotype that regulates
7 transcription.²¹ That leads us to hypothesis that interaction between two or more polymorphisms
8 within the *COL1A1* gene may influence the predisposition for ACL injury.

9 Soccer players are obviously more exposed to ACL rupture injury than the general
10 population. Therefore, the aim of this study was to examine the association of the *COL1A1* -
11 1997G/T and *COL1A1* Sp1 +1245G/T polymorphisms in the *COL1A1* gene, individually and as
12 haplotypes, with ACL ruptures in professional male soccer players. We hypothesized that the: 1)
13 *COL1A1* Sp1+1245G/T and the *COL1A1* -1997G/T polymorphisms would be individually
14 associated with the incidence of ACL rupture and; 2) The interaction between the *COL1A1*
15 Sp1+1245G/T and the *COL1A1*-1997G/T polymorphisms will form a haplotype that predisposes
16 athletes to a greater risk of ACL rupture.

17 18 **2. Methods**

19 The study was approved by the Pomeranian Medical University Ethics Committee, Poland and
20 written informed consent was obtained from each participant according to the declaration of
21 Helsinki. A total of 91 professional male soccer players (age=23±3 years) with surgically-
22 diagnosed primary ACL ruptures who qualified for ligament reconstruction were recruited for this
23 study. All players had non-contact ACL injuries. For the obvious reason that soccer teams are
24 homogenous in term of gender, we have recruited only male subjects. Soccer players were all
25 participants in the Polish 1st division professional soccer league, with an overall training time of
26 14-18 hours per week (7-9 training sessions a week, 2 hours each training session). Subjects were
27 treated in the Galen Orthopaedics, Bieruń, Poland. The control group consist of 143 apparently
28 healthy male professional soccer players (age=25.2±2.6 years), who were without any self-

1 reported history of ligament or tendon injury. Both the ACL rupture group and the healthy
2 controls were from the same soccer teams, of the same ethnicity (as self-reported, all Polish, East-
3 Europeans for ≥ 3 generations), of similar age (ACLR group= 23 ± 3 , control group= 24 ± 5), and had
4 a comparable level of exposure to risk of ACL injury (same volume and intensity of training and
5 match play).

6 We followed recent recommendations for genotype-phenotype association studies.^{22, 23} Genomic
7 DNA was extracted from the oral epithelial cells using GenElute Mammalian Genomic DNA
8 Miniprep Kit (Sigma, Germany) according to manufacturer's protocol. Allelic discrimination of
9 *COL1A1* Sp1 +1245G/T (rs1800012) and -1997G/T (rs1107946) polymorphic sites was
10 performed using a TaqMan Pre-Designed SNP Genotyping Assays (Applied Biosystems, USA),
11 including primers and fluorescently labelled (FAM and VIC) probes for the detection of alleles.
12 All samples were genotyped on a Rotor-Gene real-time polymerase chain reaction (PCR)
13 instrument (Corbett, Australia). Thermal cycler conditions were as follows: an initial step at 95°C
14 for 5 min, followed by 45 cycles of denaturation at 94°C for 15 s and anneal/extend at 60°C for 1
15 min. We used positive and negative controls for the detection of both polymorphisms. The results
16 were scored by two experienced and independent investigators who were blind to the participants'
17 data.

18 Genotype and allele frequency were analysed using χ^2 or Fisher exact tests. Allelic based odds
19 ratios (OR) with 95% confidence intervals (95%CI) were calculated using logistic regression
20 analysis. A post hoc power calculation was detected for each gene variant. The genotypes between
21 cases and controls were compared in three ways: first, in a general test of association in the 2-by-3
22 table of phenotype-by-genotype, then two different modes of inheritance of minor allele were
23 assumed: dominant, in which homozygotes and heterozygotes for the minor allele were pooled
24 and compared to homozygotes for the major allele and recessive, in which homozygotes and
25 heterozygotes for the major alleles were pooled and compared to homozygotes for the minor
26 allele. In addition, the programming language and environment R <http://www.r-project.org> was
27 used for Hardy-Weinberg and linkage disequilibrium (LD) testing (package genetics) and for the

1 haplotype analysis (package haplo.stats). The haplo.em, haplo.group and haplo.score functions of
2 the haplo.stats package were used to infer haplotype frequencies and to test the association
3 between reconstructed haplotypes and the risk of ACL rupture assuming three possible haplotype
4 effects: additive, dominant and recessive. Hap.score is the statistical score for haplotypes
5 reflecting the strength of association; the positive value of hap.score indicates increased risk of
6 ACL injury for a particular haplotype, while the negative value indicates reduced risk. The pair-
7 wise linkage disequilibrium between -1997G/T and +1245G/T was estimated by D' and r^2 . For all
8 tests, significance was set at $p < 0.05$.

9 10 **3. Results**

11 Both the ACL rupture group and the healthy controls are at a similar age category (ACL rupture
12 group= 23 ± 3 , control group= 24 ± 5 , $p = 0.093$), with a similar height (ACL rupture
13 group= 179 ± 4.6 cm, control group= 178 ± 5.2 cm, $p = 0.13$), and a similar body mass (ACL rupture
14 group= 75 ± 3.9 kg, control group= 74.3 ± 4.7 kg, $p = 0.22$). The genotype and allele frequencies for
15 the *COL1A1*-1997G/T and *COL1A1* Sp1 +1245G/T polymorphisms of *COL1A1* are shown in
16 Table 1. The genotype distributions for both polymorphisms met the Hardy-Weinberg
17 expectations in both groups (ACL rupture and controls) ($p > 0.2$).

18 There were no significant differences in genotype distribution and allele frequencies of the
19 *COL1A1* -1997G/T (rs1107946) and +1245G/T (rs1800012) polymorphisms between the ACL
20 rupture group and the control group using the 2-by-3 general test of association (Table 1).
21 However, none of the 91 participants with ACL rupture harboured the TT genotype of the
22 *COL1A1* +1245G/T (rs1800012) polymorphisms, whereas six (4.2%) TT homozygotes were
23 present in the control group. There was a trend toward under-representation of the TT genotype in
24 the ACL rupture group ($p = 0.084$ Fisher's exact test, recessive mode: TT vs GT+GG). A post hoc
25 power calculation revealed that assuming confidence level is at 95%, the statistical power to
26 detect the differences in TT genotype between ACL rupture group and the control group is 47%.
27 If the difference was two times higher (8.4% vs. 0%) the statistical power would increase to 82%.
28 Likewise, there were also no significant differences in the dominant or the recessive tests for the

1 *COL1A1* -1997G/T polymorphism.

2 *COL1A1* -1997G/T and +1245G/T were found to be in linkage disequilibrium (Table 2). A total of
3 3 reconstituted haplotypes with estimated frequency >0.05 were found, and only those were
4 evaluated for an association with ACL rupture. The G-G (-1997G, +1245G) haplotype was the
5 most common (frequency 67.9%). Two other haplotypes, G-T (frequency 16.9%) and T-G
6 (frequency 15.2%), had similar frequencies. The rare T-T haplotype (frequency 0%) was not
7 present in any of the subjects. We then tested the association between these haplotypes and ACL
8 rupture assuming three haplotype effects: additive (considering the count of a particular haplotype
9 as 0, 1 and 2), dominant (heterozygous or homozygous carrier of a particular haplotype versus
10 otherwise) and recessive (homozygous for a particular haplotype versus otherwise). Under the
11 recessive mode of inheritance, haplotype analysis yielded a mild significant association with ACL
12 rupture (p=0.048, Table 3), as two copies of the G-T haplotype conferred decreased risk of this
13 injury.

14 15 **4. Discussion**

16 We examined the association between *COL1A1* -1997G/T (rs1107946) and *COL1A1*
17 Sp1+1245G/T (rs1800012) and incidents of ACL ruptures in professional soccer players. In
18 contrast to our initial hypotheses, the *COL1A1* Sp1+1245G/T and the *COL1A1* -1997G/T
19 polymorphisms were not individually associated with the incidence of ACL rupture. However, a
20 novel finding in the present study was that the G-T haplotype (*COL1A1*-1997G, +1245T) is
21 significantly underrepresented in the ACL rupture group compared with healthy controls
22 (p=0.048), suggesting that harbouring this particular haplotype may have a protective effect
23 against ACL rupture injury.

24 In the present study, no association was found between the *COL1A1* -1997G/T
25 polymorphism and incidence of ACL ruptures. The -1997G/T polymorphism has been associated
26 with bone mineral density (BMD) in several studies^{19, 24, 25} and was found to be in high degree of
27 linkage disequilibrium (LD) with the +1245G/T loci.¹⁷ In general, promoter polymorphisms are

1 suspected to affect gene transcription activity, and thereby gene functions. That, together with the
2 assumption that the -1997G/T polymorphism has never been tested with respect to ACL ruptures,
3 led us to consider this polymorphism separately, as a candidate to influence the risk of sustaining
4 an ACL rupture. Our results suggest that the -1997G/T polymorphism is not associated with
5 incidents of ACL rupture; however it contributes to the combined influence of the *COL1A1*-
6 1997G, +1245T haplotype on the incidence of ACL ruptures.

7 Higher frequency distribution of the *COL1A1* +1245TT genotype has been reported to be
8 associated with a substantially lower risk of cruciate ligament ruptures in both Swedish¹⁶ and
9 South African participants.⁹ Combined analysis of these two studies²⁶ suggests that the TT
10 genotype has a possible preventive role not only in ACL rupture, but also in other soft tissue
11 injuries. The TT genotype frequency, when compared to combined control groups (4.1%) was
12 about ten times less frequent in subjects with all other soft tissue injuries, namely, cruciate
13 ligament ruptures, shoulder dislocation, and Achilles tendon rupture (2 cases out of 517, 0.4%). In
14 keeping with these observations, the *COL1A1* Sp1 (+1245G/T) polymorphism has been proposed
15 as a functional variant which modulates Sp1 binding and *COL1A1* gene regulation, increasing the
16 production of collagen $\alpha 1(I)$ chain relative to $\alpha 2(I)$ and reducing bone strength.²⁷ Similar to
17 previous studies, we found lower TT genotype frequency among our ACL rupture athletes (0%)
18 compared to controls (6/143 participants, 4.2%). This difference was however only a trend
19 ($p=0.084$), probably owing to the relatively low sample size, and, more importantly, the low minor
20 allele frequency (MAF) in both subjects and controls (combined MAF=16.8%), which makes
21 meaningful (MAF>40%) association harder to detect.²⁸

22 We have used a haplotype-based approach and reconstructed two-locus haplotypes that
23 were also tested for an association with the risk of ACL rupture. It was assumed that
24 predisposition for ACL rupture might be a polygenic trait, and therefore haplotype in two
25 candidate polymorphisms would provide more information on the complex relationship between
26 DNA sequence variation and traits than any single polymorphism.²⁹ In fact, studies have shown
27 that haplotype analyses are more powerful than marker-by-marker analyses when the genotypes
28 are in LD with the causative locus.³⁰ The G-T haplotype (estimated frequency among subjects

1 14.3%) was significantly lower in our ACL rupture group compared to controls, suggesting that
2 participants with two copies of this haplotype have a decreased risk for ACL injury. Additionally,
3 haplotype frequencies and score (D' and r^2) in our ACL rupture group (data not shown) were
4 similar to those in a large cohort of The Rotterdam Study.¹⁹

5 A possible functional explanation for the aforementioned results was recently provided by
6 Jin et al.,²¹ who found a significantly higher transcriptional activity of *COL1A1* gene for the
7 haplotype G-inde-T (-1997G/T, -1663indelT, +1245G/T) compared to all other haplotypes, by
8 using a functional luciferase gene reporter analysis. This study suggests that higher transcriptional
9 activity of *COL1A1* results in an unusual ratio of alpha-1(I) chains relative to alpha-2(I). However,
10 it remains unclear if the higher transcription activity of *COL1A1* has either a negative or positive
11 effect on the incidence of soft tissue injuries, including ACL ruptures.^{21, 27} In the present study, we
12 were unable to report the association between the -1663indelT polymorphism and ACL rupture
13 due to study limitations. It has been shown, however, that the -1663indelT polymorphism, that
14 corresponds to a deletion of a T within a tract of eight T residues, has a minor, if any, effect on the
15 transcription activity of *COL1A1* gene. In fact, haplotype analysis of the *COL1A1* -1997G/T and -
16 1663indelT polymorphisms revealed that haplotypes containing the G allele (-1997 G/T
17 polymorphism) yielded high transcriptional activity regardless the -1663 allele status.²⁰

18 We believe that the results of our study are overall valid, as we strictly followed the latest
19 genotype:phenotype study recommendations,^{22, 23} and all of the following criteria have been met:
20 both cases and controls (athletes) clearly presented the main study phenotype (i.e., being a
21 professional soccer players) and where equally exposed to risk of ACL rupture, participants within
22 each cohort were both age and ethnically-matched, genetic assessment was accurate and unbiased.
23 Further, genotype distributions were in Hardy-Weinberg equilibrium (HWE) in both cases and
24 controls and were similar to control genotype frequencies in other studies.^{17, 19}

25 This study has some limitations. In general, genetic association studies must be interpreted
26 with caution, since there is a non-trivial possibility of false positive results attributable to chance,
27 particularly in studies involving multiple gene-trait analyses. Furthermore, the present study is
28 hampered by small sample size, which reflects the low number of professional soccer players,

1 from the same origin, who had non-contact ACL injuries.

2 3 **5. Conclusion**

4 The *COL1A1* G-T haplotype was associated with a reduced risk of ACL injury in a group of
5 professional soccer players. Consequently, carriers of two copies of this haplotype may have a
6 reduced risk of ACL injury. Although functional analyses of *COL1A1* 5' flank suggest that the G-
7 T haplotype is responsible for enhanced transcriptional activity of the *COL1A1* gene, further
8 investigation, including evaluation of Type I/ Type 3 collagen ratio, and the possible effect of
9 microRNAs, is required to explain the relationship between *COL1A1* expression and susceptibility
10 to ACL injury.

11 12 **6. Practical implications**

- 13 • The results of the present study will assist to understand which genetic profiles contribute
14 to higher ACL injury risk.
- 15 • Discovering the complex relationship between gene variants and ACL rupture among
16 athletes may assist clinicians and coaches to optimize training , and to reduce the risk for
17 ACL rupture.
- 18 • Since it is suggested that ACL rupture is a polygenic trait, our results suggest that
19 identifying the genetic profile associated with ACL ruptures via haplotype analysis have
20 become a worthy alternative to single-locus analysis.

21 22 23 **Acknowledgments**

24 There has been no financial assistance with the project
25
26
27
28
29

References

1. Bishop D, Girard O, Mendez-Villanueva A. Repeated-sprint ability - part II: recommendations for training. *Sports Med* 2011; 41(9):741-756.
2. Stolen T, Chamari K, Castagna C, et al. Physiology of soccer: an update. *Sports Med* 2005; 35(6):501-536.
3. Bangsbo J. Energy demands in competitive soccer. *J Sports Sci* 1994;12:S5-S12.
4. Zois J, Bishop DJ, Ball K, et al. High-intensity warm-ups elicit superior performance to a current soccer warm-up routine. *J Sci Med Sport* 2011; 14(6):522-528.
5. Agel J, Evans TA, Dick R, et al. Descriptive epidemiology of collegiate men's soccer injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2002-2003. *J Athl Train* 2007; 42(2):270-277.
6. Dvorak J, Junge A. Football injuries and physical symptoms. A review of the literature. *Am J Sports Med* 2000; 28(5 Suppl):S3-S9.
7. Brophy R, Silvers HJ, Gonzales T, et al. Gender influences: the role of leg dominance in ACL injury among soccer players. *Br J Sports Med* 2010; 44(10):694-697.
8. Brooks JH, Fuller CW, Kemp SP, et al. Epidemiology of injuries in English professional rugby union: part 2 training Injuries. *Br J Sports Med* 2005; 39(10):767-775.
9. 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(5):352-356.
10. Hewett TE, Lynch TR, Myer GD, et al. Multiple risk factors related to familial predisposition to anterior cruciate ligament injury: fraternal twin sisters with anterior cruciate ligament ruptures. *Br J Sports Med* 2010; 44(12):848-855.
11. Flynn RK, Pedersen CL, Birmingham TB, et al. The familial predisposition toward tearing the anterior cruciate ligament: a case control study. *Am J Sports Med* 2005; 33(1):23-28.
12. Harner CD, Paulos LE, Greenwald AE, et al. Detailed analysis of patients with bilateral anterior cruciate ligament injuries. *Am J Sports Med* 1994; 22(1):37-43.
13. Collins M, Raleigh SM. Genetic risk factors for musculoskeletal soft tissue injuries. *Med Sport Sci* 2009; 54:136-149.

- 1 14. Grant SF, Reid DM, Blake G, et al. Reduced bone density and osteoporosis
2 associated with a polymorphic Sp1 binding site in the collagen type I alpha 1
3 gene. *Nat Genet* 1996; 14(2):203-205.
4
- 5 15. Collins M, Posthumus M, Schwellnus MP. The COL1A1 gene and acute soft
6 tissue ruptures. *Br J Sports Med* 2010; 44(14):1063-1064.
7
- 8 16. Khoschnau S, Melhus H, Jacobson A, et al. Type I collagen alpha1 Sp1
9 polymorphism and the risk of cruciate ligament ruptures or shoulder dislocations.
10 *Am J Sports Med* 2008; 36(12):2432-2436.
11
- 12 17. Garcia-Giralt N, Nogues X, Enjuanes A, et al. Two new single-nucleotide
13 polymorphisms in the COL1A1 upstream regulatory region and their relationship
14 to bone mineral density. *J Bone Miner Res* 2002; 17(3):384-393.
15
- 16 18. Stewart TL, Jin H, McGuigan FE, et al. Haplotypes defined by promoter and
17 intron 1 polymorphisms of the COL1A1 gene regulate bone mineral density in
18 women. *J Clin Endocrinol Metab* 2006; 91(9):3575-3583.
19
- 20 19. Yazdanpanah N, Rivadeneira F, van Meurs JB, et al. The -1997 G/T and Sp1
21 polymorphisms in the collagen type I alpha1 (COL1A1) gene in relation to
22 changes in femoral neck bone mineral density and the risk of fracture in the
23 elderly: the Rotterdam study. *Calcif Tissue Int* 2007; 81(1):18-25.
24
- 25 20. Garcia-Giralt N, Enjuanes A, Bustamante M, et al. In vitro functional assay of
26 alleles and haplotypes of two COL1A1-promoter SNPs. *Bone* 2005; 36(5):902-
27 908.
28
- 29 21. Jin H, van't Hof RJ, Albagha OM, et al. Promoter and intron 1 polymorphisms of
30 COL1A1 interact to regulate transcription and susceptibility to osteoporosis. *Hum*
31 *Mol Genet* 2009; 18(15):2729-2738.
32
- 33 22. Chanock SJ, Manolio T, Boehnke M, et al. Replicating genotype-phenotype
34 associations. *Nature* 2007; 447(7145):655-660.
35
- 36 23. Attia J, Ioannidis JP, Thakkinstian A, et al. How to use an article about genetic
37 association: B: Are the results of the study valid? *Jama* 2009; 301(2):191-197.
38
- 39 24. Liu PY, Lu Y, Long JR, et al. Common variants at the PCOL2 and Sp1 binding
40 sites of the COL1A1 gene and their interactive effect influence bone mineral
41 density in Caucasians. *J Med Genet* 2004; 41(10):752-757.
42
- 43 25. Zhang YY, Lei SF, Mo XY, et al. The -1997 G/T polymorphism in the COL1A1
44 upstream regulatory region is associated with hip bone mineral density (BMD) in
45 Chinese nuclear families. *Calcif Tissue Int* 2005; 76(2):107-112.
46
- 47 26. Collins M. Genetic risk factors for soft-tissue injuries 101: a practical summary to
48 help clinicians understand the role of genetics and 'personalised medicine'. *Br J*
49 *Sports Med* 2010; 44(13):915-917.
50

- 1 27. Mann V, Hobson EE, Li B, et al. A COL1A1 Sp1 binding site polymorphism
2 predisposes to osteoporotic fracture by affecting bone density and quality. *J Clin*
3 *Invest* 2001; 107(7):899-907.
4
- 5 28. Tabangin ME, Woo JG, Martin LJ. The effect of minor allele frequency on the
6 likelihood of obtaining false positives. *BMC Proc* 2009; 3(Suppl 7):S41.
7
- 8 29. Stephens JC, Schneider JA, Tanguay DA, et al. Haplotype variation and linkage
9 disequilibrium in 313 human genes. *Science* 2001; 293(5529):489-493.
10
- 11 30. Akey J, Jin L, Xiong M. Haplotypes vs single marker linkage disequilibrium tests:
12 what do we gain? *Eur J Hum Genet* 2001; 9(4):291-300.
13
14
15
16
17
18
19
20
21

1
2
3
4
5
6
7
8
9
10
11
12
13

Tables

Table 1. Association between the presence of either *COL1A1* -1997G/T (rs1107946) or *COL1A1*+1245G/T (rs1800012) polymorphisms and incidence of ACL rupture

SNP	Group	HWE	Genotype distribution	p value	Allele frequency	p value	OR (95% CI)
-1997G/T	ACLR n=91	0.288	GG: 60 (65.9%) GT: 30 (33.0%) TT: 1(1.1%)	0.228 p_D 0.182 p_R 1.000	G: 82.4 T: 17.6	0.246	1.35 (0.79-2.32)
	Control n=143	0.726	GG: 107 (74.8%) GT: 33 (23.1%) TT: 3 (2.1%)		G: 86.4 T: 13.6		
+1245G/T	ACLR n=91	0.201	GG: 65 (71.4%) GT: 26 (28.6%) TT: 0 (0%)	0.138 p_D 0.563 p_R 0.084	G: 85.7 T: 14.3	0.232†	0.73 (0.43-1.26)
	Control n=143	0.578	GG: 96 (67.1%) GT: 41 (28.7%) TT: 6 (4.2%)		G: 81.5 T: 18.5		

a) P values correspond to genotype distribution and allele frequency.

b) OR correspond to the odds ratio for the incidence of ACL rupture.

ACLR=anterior cruciate ligament rupture; HWE =Hardy-Weinberg Equilibrium; OR=odds ratio; CI=confident interval

p_D and p_R are two-sided Fisher's exact test probabilities for dominant (TT+GT vs GG) and recessive (TT vs GT+GG) modes of inheritance of the minor allele (-1997T, +1245T), respectively.

1 **Table 2. Pair-wise linkage disequilibrium (LD) and inferred haplotype frequencies for *COL1A1* -1997G/T and +1245G/T**
 2
 3

Gene	SNP1	SNP2	D'	r ²	Haplotypes	
					Haplotype	Frequency (%)
<i>COL1A1</i>	-1997G/T (rs1107946)	+1245G/T (rs180012)	0.997†	0.036	G-G	67.9
					G-T	16.9
					T-G	15.2
					T-T	0.0
					Total	100.0

4
 5 a) D'=the deviation of the observed frequency of a haplotype from the expected. D'=0.977 suggest on very high LD between *COL1A1* -
 6 1997G/T and *COL1A1*+1245G/T

7 b) Haplotypes are in decreasing frequency

8 † p<0.001
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18

Table 3. Haplotypes frequencies in ACLR and controls (under a recessive mode of haplotype inheritance)

Haplotypes		Frequency (%)		Hap.score	p value
1997G/T	+1245G/T	combined ACLR and Controls	ACLR Controls		
G	G	67.9	68.1 67.8	-0.45	0.652
G	T	16.9	14.3 18.5	-1.98	0.048

ACLR=anterior cruciate ligament rupture group; Hap.score= statistics score for haplotypes