COL5A1 rs12722 polymorphism is not associated with passive muscle stiffness and sports-related muscle injury in Japanese athletes

**CURRENT STATUS:** ACCEPTED

**BMC Medical Genetics**

Eri Miyamoto-Mikami
Juntendo University

Email: miyamoto.mikami@gmail.com
**Corresponding Author**
**ORCID:** https://orcid.org/0000-0002-4864-8581

Naokazu Miyamoto
Juntendo University

Hiroshi Kumagai
Juntendo University

Kosuke Hirata
Shibaura Institute of Technology

Naoki Kikuchi
Nippon Sport Science University

Hirofumi Zempo
Tokyo Seiei College

Noriko Kimura
Osaka University of Health and Sport Sciences

Nobuhiko Kamiya
Tenri University

Hiroaki Kanehisa
Ritsumeikan University

Hisashi Naito
Juntendo University

Noriyuki Fuku
Juntendo University
| DOI:               |
|-------------------|
| 10.21203/rs.2.11654/v2 |

**SUBJECT AREAS**

*Medical Genetics*

**KEYWORDS**

*muscle injury, muscle stiffness, joint flexibility, single nucleotide polymorphism, type V collagen*
Abstract
Background: Poor joint flexibility has been repeatedly proposed as a risk factor for muscle injury. The C-to-T polymorphism (rs12722) in the 3′-untranslated region of the collagen type V α1 chain gene (COL5A1) is reportedly associated with joint flexibility. Flexibility of a normal joint is largely determined by passive muscle stiffness, which is influenced by intramuscular collagenous connective tissues including type V collagen. The present study aimed to test the hypothesis that the COL5A1 rs12722 polymorphism influences joint flexibility via passive muscle stiffness, and is accordingly associated with the incidence of muscle injury. Methods: In Study 1, we examined whether the rs12722 polymorphism is associated with joint flexibility and passive muscle stiffness in 363 healthy young adults. Joint flexibility was evaluated by passive straight-leg-raise and sit-and-reach tests, and passive muscle stiffness was measured using ultrasound shear wave elastography. In Study 2, the association of the rs12722 polymorphism with sports-related muscle injury was assessed in 1559 Japanese athletes. Muscle injury history and severity were assessed by a questionnaire. In both Study 1 and Study 2, the rs12722 C-to-T polymorphism in the COL5A1 was determined using the TaqMan SNP Genotyping Assay. Results: Study 1 revealed that the rs12722 polymorphism had no significant effect on range of motion in passive straight-leg-raise and sit-and-reach tests. Furthermore, there was no significant difference in passive muscle stiffness of the hamstring among the rs12722 genotypes. In Study 2, rs12722 genotype frequencies did not differ between the muscle injury and no muscle injury groups. Moreover, no association was observed between rs12722 polymorphism and severity of muscle injury. Conclusions: The present study does not support the view that COL5A1 rs12722 polymorphism has a role as a risk factor for sports-related muscle injury, or that it is a determinant for passive muscle stiffness in a Japanese population.

Background
Despite preventive efforts, muscle injury is increasingly common in sports involving sprinting and jumping (1). Muscle injury results in considerable losses such as missed training time and unavailability for competition and thereby negatively influences the athletes’ sport success. Thus, it is essential to elucidate the etiology of sports-related muscle injury in order to develop a more effective
Poor joint flexibility has been well documented as a modifiable risk factor for muscle injury. A prospective study (2) indicated a direct link between pre-season joint flexibility and muscle injury in soccer players; with injured players having a significantly smaller joint range of motion (ROM) measured pre-season than uninjured players. Joint flexibility is influenced by environmental factors such as stretching (3, 4) and resistance training (5). In contrast, a meta-analysis showed that genetic factors explain 50% of joint flexibility variance (6). Thus, the influences of genetic and environmental factors on joint flexibility are comparable.

Other than the commonly accepted extrinsic factor, genetic factors play an important role in the risk of muscle injury (7, 8). The collagen type V α1 chain gene (COL5A1) has been proposed as a candidate gene whose variants affect not only muscle injury per se but also passive muscle stiffness and joint flexibility (9). Several studies examining the association of the C-to-T polymorphism (rs12722) in the 3’-untranslated region (UTR) of the COL5A1 with ROM in sit-and-reach (SR) or passive straight-leg-raise (PSLR) tests (10-12) suggest that this polymorphism could be a factor affecting ROM. Although the ROMs in SR and PSLR tests are multifactorial, passive muscle stiffness has been proposed to be a major contributor to the ROM of normal joints (13, 14). Passive muscle stiffness is influenced by intramuscular collagenous connective tissues, such as perimysium and endomysium (15), which contain type V collagen (16). COL5A1 is reportedly expressed in skeletal muscle tissue (17). Given this, it is assumed that the COL5A1 rs12722 polymorphism influences ROM via passive muscle stiffness, and is accordingly associated with muscle injury. To date, two studies suggest that the COL5A1 rs12722 polymorphism is associated with severity of muscle injury in Caucasian professional soccer players (18, 19). However, the association of the COL5A1 rs12722 polymorphism with passive muscle stiffness and incidence of muscle injury remains unclear.

The C-to-T rs12722 polymorphism is reported to be associated with altered stability of COL5A1 mRNA (20). The COL5A1 3’-UTR with the rs12722 polymorphism T allele exhibited enhanced mRNA stability compared to the 3’-UTR of the C allele, suggesting that more type V collagen α1 chain is synthesized from the T allele. Although type V collagen is a quantitatively minor fraction fibrillar collagen, it plays
a critical role in the regulation of collagen fibril assembly and is an important structural component of tendons, ligaments, and other connective tissues (9). As suggested by Collins and Posthumus (9), it is reasonable to assume that the T allele of the COL5A1 rs12722 polymorphism leads to higher type V collagen production and accordingly alters collagen fibril architecture, resulting in changes in the mechanical properties of the connective tissues. Based on these considerations, we hypothesized that the T allele of the COL5A1 rs12722 polymorphism is associated with high muscle stiffness, and accordingly poor joint flexibility and a high incidence of muscle injury. To test this hypothesis, we examined the association of COL5A1 rs12722 polymorphism with passive muscle stiffness, ROM, and muscle injury.

Methods

Study design

We performed two studies to fulfill the goals of this paper. The first study (Study 1) was designed to examine the association between COL5A1 rs12722 polymorphism and joint ROM and passive muscle stiffness. In the second study (Study 2), we investigated whether COL5A1 rs12722 polymorphism was associated with sports-related muscle injury. Study 2 was a part of the Japanese Human Athlome Project (J-HAP) in “Athlome Project Consortium” (21).

Study 1

A total of 363 healthy young adults (men, n = 231; women, n = 132) participated in Study 1. None of the participants had apparent neurological, orthopedic, or neuromuscular problems. None of the participants reported muscle soreness or fatigue in the lower limbs at the time of testing. Written consent was obtained from each participant. The procedure was approved by the ethics committee of the Juntendo University and National Institute of Fitness and Sports in Kanoya, and performed in accordance with the Declaration of Helsinki.

The PSLR (right leg) test, SR test, and passive muscle stiffness measurements were performed on each participant in a randomized order. Room temperature for all measurements was kept at 24 ± 2 °C to minimize potential effects of temperature-induced changes in tissue mechanical properties and participants’ sensation to muscle stretch. Prior to the measurements the participants were not
allowed to warm-up or stretch. The procedures have been described in detail previously (14). Briefly, in the PSLR test, participants lay supine with their legs straight on an examining bed. The pelvis and non-testing (left) leg were secured to the bed. The right hip joint was passively flexed, with the knee straight, by an examiner until each participant felt pain in the hamstring. The PSLR ROM (i.e., hip flexion angle from the resting position) was measured using a digital inclinometer (MLT-100, Sakai Medical, Japan) attached to the right shank. In the SR test, participants sat on the floor with their heada back, and hip against a wall, knee fully extended, and soles of the foot positioned flat against an SR box (T.K.K.5111, Takei Scientific Instruments, Japan). They were then asked to bend forward slowly and reach forward as far as possible while keeping the knees extended and slid their hands along a digital measuring scale which was placed on the box to measure SR ROM.

In accordance with a previous study, passive shear modulus (a measure of stiffness, expressed in kPa) of the biceps femoris long head (BF), semitendinosus (ST), and semimembranosus (SM) of the right leg were measured using an ultrasound shear wave elastography scanner (Aixplorer, Supersonic Imagine, France). During the measurement of passive muscle stiffness, participants sat on a bench with their hip flexed 70° and the right knee fully extended. This hip joint angle was chosen based on a recent study (14), which aimed to define an angle where the hamstring could be stretched to a tensioned state without pain for all participants and the shear modulus could be quantified at a given joint angle in all participants. An ultrasound linear probe was positioned at 50% level of the thigh length (the distance between the greater trochanter and the lateral epicondyle of the femur). For each muscle, the probe orientation was adjusted to visualize the fascicles within the B-mode image. Care was taken not to press and deform the muscles while scanning. The subjects were instructed to fully relax the leg throughout the measurements. The images were acquired after ensuring a stable color distribution of shear modulus mapping for a few seconds. The measurements were performed three times for each muscle. To evaluate the stiffness of the overall hamstring, the shear modulus of three muscles were averaged. All measurements and analyses of the elastographic data were performed by experienced examiners (>3 years of practice). For each variable (i.e., SR test, PSLR test, and passive muscle stiffness), the average values of three measurements were used for
subsequent analyses. All participants then completed a questionnaire that included information on regular stretching of the hamstring.

**Study 2**

Participants of study 2 were Japanese athletes of various sports (not limited to a specific discipline), recruited from March 2015 to November 2017. A total of 2181 participants were recruited. Written consent was obtained from each participant. The procedure was approved by the Ethics Committees of Juntendo University, Nippon Sport Science University, Tenri University, and the National Institute of Fitness and Sports in Kanoya, and performed in accordance with the Declaration of Helsinki.

In J-HAP, the history of sports-related injuries was assessed by questionnaire (8). In the present paper, we focused on non-contact muscle injury diagnosed by medical practitioners. The number of days that elapsed from the date of injury to the date of the athlete’s return to usual training was asked to evaluate the severity of injury via the questionnaire. Injury severity was categorized into one of four levels based on the number of days until return: minimal (1-3 days), mild (4-7 days), moderate (8-28 days), and severe (>28 days) (22). If a participant had a history of multiple muscle injuries, data on the most severe injury was used. In addition, precise information on the primary sport, playing years, and competition level was obtained using the questionnaire. Participants who had (1) no Japanese ancestry, (2) missing or invalid questionnaire data, or (3) less than 3 years playing experience in their primary sport were excluded. The final sample size in the analyses of history of muscle injury and severity of muscle injury were 1559 and 186, respectively.

**Genotyping analysis**

Total DNA was isolated from the saliva of all participants in both Study 1 and 2 with an Oragene® DNA Collection Kit (DNA Genotek, ON, Canada) and quantified using a NanoDrop 8000 UV-Vis Spectrophotometer (Thermo Fisher Scientific, DE, USA) or Eppendorf Bio Photometer Plus (Eppendorf, Tokyo, Japan). DNA samples were stored at 4 °C until use. The samples were analyzed for the rs12722 polymorphism in the 3’-UTR of COL5A1 using a TaqMan SNP Genotyping Assay (Assay ID: C__370252_20) and LightCycler® 480 System (Roche Molecular Systems, Mannheim, Germany) or StepOne™ Real-Time PCR System (Thermo Fisher Scientific). Five microliters of the genotyping
mixture contained 2.5 μL TaqMan® GTXpressTM Master Mix (2×) or TaqMan® Universal Master Mix II (2×), 0.0625 μL TaqMan® SNP Genotyping Assay mix (40×), 1.4375 μL sterilized water, and 1 μL genomic DNA (10 ng/μL). Two to four negative controls were included on each plate. Genotypes were called based on TaqMan® assays results using LightCycler® 480 SW (version 1.5, Roche Molecular Systems) or StepOne™ software (version 2.3, Thermo Fisher Scientific). 380 randomly selected samples were genotyped in duplicate for the rs12722 polymorphism, and we confirmed that the genotyping results perfectly agreed between duplicates.

**Statistical analysis**

Data are expressed as mean ± standard deviation (SD). Statistical significance was set at P < 0.05. Statistical analyses were performed using JMP Pro version 12 (SAS Institute, USA). The Hardy-Weinberg equilibrium of the rs12722 polymorphism was assessed using χ² test. For the data of study 1, comparisons between two groups (sex, regular stretching habit, genotype [T-dominant and T-recessive models]) were conducted using χ² test or unpaired Student’s t-test as appropriate. For comparisons under an additive model, the Cochran-Amitage trend test or Spearman correlation test was used as appropriate. Additionally, in order to examine whether the genotypes are associated with the phenotype variables independently of sex and regular stretching, analysis of covariance (ANCOVA) and multiple regression analysis were employed for T-dominant and T-recessive models and for an additive model, respectively. For the data of study 2, comparisons between the muscle injury and no muscle injury groups were conducted by χ² test or unpaired Student’s t-test. Logistic regression analysis was applied to investigate the associations between rs12722 polymorphism and history of muscle injury with adjustment for playing years and main sport (track & field or the others). Odds ratios (OR) and 95% confidence intervals (CI) were calculated under the T-dominant genetic model. Association of the rs12722 polymorphism with severity of muscle injury was assessed by ordinal regression analysis. When the sample size of the TT genotype of the rs12722 polymorphism was less than five, statistical analyses were not conducted for T-recessive and additive models. Using the data of our previous study (8), we calculated the necessary samples size to detect the expected
difference in passive muscle stiffness between the T allele carriers and CC genotype carriers ($\alpha = 0.05$, power = 0.8, difference in group means = 3.5, within group SD = 7.4, and the ratio of CC genotype carriers to T allele carriers = 2.3) and to detect the association between the T allele of the rs12722 polymorphism and a history of muscle injury with an OR of 2.0 ($\alpha = 0.05$, power = 0.8, probability of T allele carriers = 0.317, and the ratio of control to case subjects = 8.9). The critical sample sizes were estimated to be 168 and 733, respectively.

Results

The COL5A1 rs12722 genotype frequencies in studies 1 and 2 did not deviate from Hardy-Weinberg equilibrium ($P = 0.975$ and $P = 0.183$).

**Study 1**

Table 1 shows descriptive data on the characteristics of the participants in study 1. There were significant sex differences in height, body mass, PSLR ROM, SR ROM, and muscle shear modulus of the ST and SM (Table 1). When the participants were divided by stretching habit, participants conducting regular stretching showed significantly higher PSLR (84.4 ± 17.5 degree vs. 75.6 ± 14.9 degree, $P < 0.001$) and SR (11.5 ± 9.0 cm vs. 6.5 ± 10.1 cm, $P < 0.001$) ROMs and lower shear modulus of SM (38.6 ± 15.3 kPa vs. 46.4 ± 19.7 kPa, $P < 0.001$) than those without a stretching habit.

The $\chi^2$ test revealed no significant difference in sex ratio and regular stretching among those with the various rs12722 genotypes (Table 2). PSLR and SR ROMs did not differ among the genotypes (Table 2). Similarly, no significant effect of genotype on PSLR and SR ROMs was observed after adjustment for sex ratio and regular stretching (Table 2).

Figure 1 shows the muscle shear modulus for each rs12722 genotype. There was no significant difference in the shear modulus of all three muscles under three genetic models, regardless of the adjustment for sex ratio and regular stretching (unadjusted: $P \geq 0.268$, adjusted: $P \geq 0.276$, Figure 1).

When the shear moduli of the overall hamstring were compared among the rs12722 genotypes, no significant difference was observed either before (T-dominant: $P = 0.650$, T-recessive: $P = 0.747$, Additive: $P = 0.618$) or after adjustment for sex ratio and regular stretching (T-dominant: $P = 0.958$, T-recessive: $P = 0.835$, Additive: $P = 0.987$). In addition, when the participants were divided by sex,
genotype had no significant effect on PSLR/SR ROMs or the shear modulus for any of the three muscles in either sex group (Supplemental Table 1).

**Study 2**

Table 3 shows the participant characteristics in study 2. There were no significant differences in sex ratio, age, height, body mass, and competitive level between the muscle injury and no muscle injury groups. The muscle injury group had fewer playing years and a higher proportion of track and field athletes than the no muscle injury group. Logistic regression analyses of the rs12722 polymorphism (as a whole or for males and females separately) revealed that the genotype frequencies did not differ between the muscle injury and no muscle injury groups (Table 4).

Figure 2 shows the distribution of muscle injury severity in each rs12722 genotype. There was no significant difference in the distribution of injury severity between participants with TT+TC genotype and those with the CC genotype at the rs12722 polymorphism (P = 0.717).

**Discussion**

The main finding obtained here is that the COL5A1 rs12722 polymorphism is not associated with joint ROM, passive muscle stiffness, or a history of muscle injury. These results rule out our hypothesis and suggest that at least in a Japanese population, passive muscle stiffness and muscle injury risk is not affected by rs12722 polymorphism.

In recent years, a growing number of studies have paid attention to genetic susceptibilities to musculoskeletal soft tissue injuries such as tendon and ligament injuries (23-29). However, limited evidence is available on the genetic component of muscle injury. In the present paper, we focused on the COL5A1 rs12722 polymorphism and hypothesized that the COL5A1 rs12722 T allele would be associated with greater susceptibility to muscle injury. To test this hypothesis, we conducted a case-control association analysis in 1559 Japanese athletes, which was much larger than the sample size required to detect the potential association with enough statistical power (n = 733). However, contrary to our hypothesis, no association was found between the COL5A1 rs12722 polymorphism and muscle injury. Additionally, previous studies reported that the COL5A1 rs12722 polymorphism was associated with the severity of muscle injury in professional soccer players (18, 19) whereas no
association was found between the COL5A1 rs12722 polymorphism and the injury severity in the present study. Participants in the present study were athletes from various sports (not limited to a specific discipline) and with various competitive levels. The polymorphism was not associated with muscle injury history/severity even when examining each sports discipline (i.e., soccer, track & field, etc.) and competitive level (i.e., international, national, and regional) individually (data not shown). The reasons for the discrepancy between the hypothesis and the results are unclear at this time, but we will discuss some possible explanations.

Assuming that intramuscular connective tissues such as the perimysium and endomysium contain type V collagen (16) and contribute largely to passive muscle stiffness (15), we expected that passive muscle stiffness would be influenced by the rs12722 polymorphism. However, the present study found no association between the rs12722 polymorphism and passive muscle stiffness. Passive muscle stiffness is reportedly influenced by not only the connective tissue but also the presence of titin, a giant sarcomeric protein with a wide range of functions (30). In cardiac muscle, at shorter sarcomere lengths, passive tension development of muscle fibers mainly depends on titin, while at longer sarcomere lengths it is determined by extracellular collagen fibers (31). Muscle shear modulus measured in a sitting position in the present study were approximately 54% (BF), 43% (ST), and 31% (SM) of those measured at maximal ROM (32). Taking these aspects into account, it seems that the contribution of intramuscular connective tissues to passive muscle stiffness might be small in the experimental design used here to determine passive muscle stiffness. In other words, we cannot eliminate possibility that passive muscle stiffness would be influenced by the rs12722 polymorphism when measured in more stretched and therefore tensioned hamstring positions. Further studies will be required to clarify these points.

To the best of our knowledge, four studies have attempted to identify the association between the COL5A1 rs12722 polymorphism and joint flexibility (10-12, 33). Of these, three have indicated a potential association (10-12), which seems to be inconsistent with the present results. However, careful attention should be paid when interpreting the previous findings. For example, Collins et al. (11) and Brown et al. (10) included participants with a wide age range (although its effect was
statistically adjusted). Especially in the latter study, a significant association between the $COL5A1$ rs12722 polymorphism and SR ROM was observed in the old age ($\geq$ 35 years) group but not in the young age group ($< 35$ years) or combined group (i.e., all participants) (10). The present study examined only young populations ($\leq 32$ years) and found no association between the rs12722 polymorphism and joint flexibility. There remains a possibility that the influence of the rs12722 polymorphism on joint flexibility is age-dependent. Another possible explanation for not only this inconsistency regarding joint flexibility but also the aforementioned discrepancy in the association of the $COL5A1$ rs12722 polymorphism with injury severity is the difference in participant ethnicity: Caucasian (10, 11, 18, 19) vs. East Asian. A recent meta-analysis indicated that the $COL5A1$ rs12722 polymorphism was associated with tendon and ligament injuries in Caucasians but not in Asians (25).

The T allele frequency of the rs12722 polymorphism is lower in East Asian [0.220 in East Asian population of the 1000 Genomes (34), 0.162 in Japanese population of the Integrative Japanese Genome Variation Database (35)] than in Caucasian [0.585 in European population of the 1000 Genomes (34)] populations. If the optimum model in the association of the rs12722 polymorphism with joint flexibility and muscle injury is T-recessive or additive, the potential association may not be found due to the low frequency of the TT genotype in East Asian populations ((33) and the present study). In the present study, however, joint ROMs and passive muscle stiffness in TT genotype carriers were comparable to those in TC+CC genotype carriers. All in all, these facts suggest that the $COL5A1$ rs12722 polymorphism plays a minimal or negligible role in sports-related muscle injury risk.

A limitation of the present study is the cross-sectional design, which does not permit a mechanistic explanation. Thus, large-scale follow-up studies would be desirable to conclude whether the $COL5A1$ rs12722 polymorphism is associated with muscle injury. Another limitation is that we have tested only the $COL5A1$ rs12722 polymorphism. It is possible that other polymorphisms within the $COL5A1$ as well as the genes for other types of collagen are associated with the joint flexibility, passive muscle stiffness, and muscle injury. Further investigations are warranted to identify these.

Conclusions
The present study does not support a role for the $COL5A1$ rs12722 polymorphism in sports-related
muscle injury in a Japanese population. Furthermore, we provide evidence suggesting that the polymorphism is not related to passive muscle stiffness or joint flexibility.

Abbreviations
ANCOVA: analysis of covariance, BF: biceps femoris long head, CI: confidence intervals, COL5A1: collagen type V α1 chain, OR: odds ratios, PSLR: passive straight-leg-raise, ROM: range of motion, SM: semimembranosus, SR: sit-and-reach, ST: semitendinosus, UTR: untranslated region

Declarations

Ethics approval and consent to participate
The present study was approved by the Ethics Committees of the Juntendo University, Nippon Sport Science University, Tenri University, and National Institute of Fitness and Sports in Kanoya. All participants provided written informed consent.

Consent for publication
Not applicable.

Availability of data and material
The datasets analyzed during the present study are available from the corresponding author on reasonable request.

Competing interests
The authors declare that they have no competing interests.

Funding
This study was supported in part by grants from the Japan Society for Promotion of Science (JSPS) KAKENHI Scientific Research (B) (18H03155 to N. F. and 16H03233 to N. M.), Young Scientists (A) (17H04752 to E. M.), Young Scientists (18K17863 to H. K.), MEXT-Supported Program for the Strategic Research Foundation at Private Universities (to Juntendo University), and Tenri University Research Grant (N. Kamiya.). H. K. and K. H. were recipients of a Grant-in-Aid for JSPS fellows from the JSPS.

Authors' contributions
EM, NM, HN, and NF designed the work; EM, NM, HKumagai, KH, NKikuchi, HZ, NKimura, NKamiya and NF collected the samples and the data, and performed the analyses; EM, NM, HKanehisa, and NF
interpreted the data and have drafted the work or substantively revised it. All authors read and approved the final manuscript.

**Acknowledgements**

We would like to thank Editage (www.editage.jp) for English language editing.

**References**

1. Ekstrand J, Walden M, Hagglund M. Hamstring injuries have increased by 4% annually in men's professional football, since 2001: a 13-year longitudinal analysis of the UEFA Elite Club injury study. Br J Sports Med. 2016;50(12):731-7.

2. Witvrouw E, Danneels L, Asselman P, D'Have T, Cambier D. Muscle flexibility as a risk factor for developing muscle injuries in male professional soccer players. A prospective study. Am J Sports Med. 2003;31(1):41-6.

3. Nakamura M, Ikezoe T, Takeno Y, Ichihashi N. Effects of a 4-week static stretch training program on passive stiffness of human gastrocnemius muscle-tendon unit in vivo. Eur J Appl Physiol. 2012;112(7):2749-55.

4. Weppler CH, Magnusson SP. Increasing muscle extensibility: a matter of increasing length or modifying sensation? Phys Ther. 2010;90(3):438-49.

5. Morton SK, Whitehead JR, Brinkert RH, Caine DJ. Resistance training vs. static stretching: effects on flexibility and strength. J Strength Cond Res. 2011;25(12):3391-8.

6. Schutte NM, Nederend I, Hudziak JJ, de Geus EJ, Bartels M. Differences in Adolescent Physical Fitness: A Multivariate Approach and Meta-analysis. Behav Genet. 2016;46(2):217-27.

7. Larruskain J, Celorio D, Barrio I, Odriozola A, Gil SM, Fernandez-Lopez JR, et al. Genetic Variants and Hamstring Injury in Soccer: an Association and Validation Study. Med Sci Sports Exerc. 2017.
8. Kumagai H, Miyamoto-Mikami E, Hirata K, Kikuchi N, Kamiya N, Hoshikawa S, et al. ESR1 rs2234693 Polymorphism Is Associated with Muscle Injury and Muscle Stiffness. Med Sci Sports Exerc. 2019;51(1):19-26.

9. Collins M, Posthumus M. Type V collagen genotype and exercise-related phenotype relationships: a novel hypothesis. Exerc Sport Sci Rev. 2011;39(4):191-8.

10. Brown JC, Miller CJ, Schwellnus MP, Collins M. Range of motion measurements diverge with increasing age for COL5A1 genotypes. Scand J Med Sci Sports. 2011;21(6):e266-72.

11. Collins M, Mokone GG, September AV, van der Merwe L, Schwellnus MP. The COL5A1 genotype is associated with range of motion measurements. Scand J Med Sci Sports. 2009;19(6):803-10.

12. Lim ST, Kim CS, Kim WN, Min SK. The COL5A1 genotype is associated with range of motion. J Exerc Nutrition Biochem. 2015;19(2):49-53.

13. Magnusson SP, Simonsen EB, Aagaard P, Boesen J, Johannsen F, Kjaer M. Determinants of musculoskeletal flexibility: viscoelastic properties, cross-sectional area, EMG and stretch tolerance. Scand J Med Sci Sports. 1997;7(4):195-202.

14. Miyamoto N, Hirata K, Kimura N, Miyamoto-Mikami E. Contributions of Hamstring Stiffness to Straight-Leg-Raise and Sit-and-Reach Test Scores. Int J Sports Med. 2018;39(2):110-4.

15. Gajdosik RL. Passive extensibility of skeletal muscle: review of the literature with clinical implications. Clin Biomech (Bristol, Avon). 2001;16(2):87-101.

16. Light N, Champion AE. Characterization of muscle epimysium, perimysium and endomysium collagens. Biochem J. 1984;219(3):1017-26.

17. Ono H, Ogasawara O, Okubo K, Bono H. RefEx, a reference gene expression dataset as a web tool for the functional analysis of genes. Sci Data. 2017;4:170105.
18. Massidda M, Bachis V, Corrias L, Piras F, Scorcu M, Calo CM. Influence of the COL5A1 rs12722 on musculoskeletal injuries in professional soccer players. J Sports Med Phys Fitness, 2015;55(11):1348-53.

19. Pruna R, Artells R, Ribas J, Montoro B, Cos F, Munoz C, 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.

20. Laguette MJ, Abrahams Y, Prince S, Collins M. Sequence variants within the 3'-UTR of the COL5A1 gene alters mRNA stability: implications for musculoskeletal soft tissue injuries. Matrix Biol. 2011;30(5-6):338-45.

21. Pitsiladis YP, Tanaka M, Eynon N, Bouchard C, North KN, Williams AG, et al. Athlome Project Consortium: a concerted effort to discover genomic and other "omic" markers of athletic performance. Physiol Genomics. 2016;48(3):183-90.

22. Fuller CW, Ekstrand J, Junge A, Andersen TE, Bahr R, Dvorak J, et al. Consensus statement on injury definitions and data collection procedures in studies of football (soccer) injuries. Br J Sports Med. 2006;40(3):193-201.

23. Abrahams Y, Laguette MJ, Prince S, Collins M. Polymorphisms within the COL5A1 3'-UTR that alters mRNA structure and the MIR608 gene are associated with Achilles tendinopathy. Ann Hum Genet. 2013;77(3):204-14.

24. Longo UG, Margiotti K, Petrillo S, Rizzello G, Fusilli C, Maffulli N, et al. Genetics of rotator cuff tears: no association of col5a1 gene in a case-control study. BMC Med Genet. 2018;19(1):217.

25. Lv ZT, Gao ST, Cheng P, Liang S, Yu SY, Yang Q, et al. Association between polymorphism rs12722 in COL5A1 and musculoskeletal soft tissue injuries: a systematic review and meta-analysis. Oncotarget. 2018;9(20):15365-74.
26. Mokone GG, Schwellnus MP, Noakes TD, Collins M. The COL5A1 gene and Achilles tendon pathology. Scand J Med Sci Sports. 2006;16(1):19-26.

27. Posthumus M, September AV, Keegan M, O’Cuinneagain D, Van der Merwe W, Schwellnus MP, et al. Genetic risk factors for anterior cruciate ligament ruptures: COL1A1 gene variant. Br J Sports Med. 2009;43(5):352-6.

28. 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(11):2234-40.

29. September AV, Cook J, Handley CJ, van der Merwe L, Schwellnus MP, Collins M. Variants within the COL5A1 gene are associated with Achilles tendinopathy in two populations. Br J Sports Med. 2009;43(5):357-65.

30. Neagoe C, Opitz CA, Makarenko I, Linke WA. Gigantic variety: expression patterns of titin isoforms in striated muscles and consequences for myofibrillar passive stiffness. J Muscle Res Cell Motil. 2003;24(2-3):175-89.

31. Granzier HL, Irving TC. Passive tension in cardiac muscle: contribution of collagen, titin, microtubules, and intermediate filaments. Biophys J. 1995;68(3):1027-44.

32. Miyamoto N, Hirata K, Kanehisa H. Effects of hamstring stretching on passive muscle stiffness vary between hip flexion and knee extension maneuvers. Scand J Med Sci Sports. 2017;27(1):99-106.

33. Kim JH, Jung ES, Kim CH, Youn H, Kim HR. Genetic associations of body composition, flexibility and injury risk with ACE, ACTN3 and COL5A1 polymorphisms in Korean ballerinas. J Exerc Nutrition Biochem. 2014;18(2):205-14.

34. dbSNP (Short Genetic Variation): National Center for Biotechnology Information. 2018 [cited 2019 15 May 2019]; Available from: https://www.ncbi.nlm.nih.gov/snp/rs12722#frequency_tab.
35. Integrative Japanese Genome Variation Database: Tohoku Medical Megabank
Organization. 2018 [cited 2019 15 May 2019]; Available from:
https://ijgvd.megabank.tohoku.ac.jp/search/.

Tables
Table 1. Participant characteristics in Study 1 (n = 363)

|                     | All       | Men       | Women     |
|---------------------|-----------|-----------|-----------|
| n                   | 363       | 231       | 132       |
| Age, year (year)    | 20.5 ± 2.0| 20.7 ± 2.0| 20.2 ± 1.9|
| Height, cm (cm)     | 169.3 ± 9.0| 174.0 ± 6.6 | 161.2 ± 6.5 |
| Body mass, kg       | 64.2 ± 11.6| 68.4 ± 11.5 | 56.8 ± 7.2 |
| PSLR ROM, degree    | 82.3 ± 17.3| 77.1 ± 14.4 | 91.5 ± 18.1 |
| SR ROM, cm          | 10.43 ± 9.5| 8.3 ± 9.6 | 13.9 ± 8.2 |
| Shear modulus (kPa) |           |           |           |
| Biceps femoris      | 19.9 ± 7.0| 19.8 ± 7.3 | 20.1 ± 6.4 |
| Semitendinosus      | 25.8 ± 8.6| 27.3 ± 9.1 | 22.9 ± 6.4 |
| Semimembranosus     | 40.4 ± 16.7| 42.5 ± 17.4 | 36.2 ± 14.5 |
| Stretch, % yes      | 78.2      | 73.9      | 81.1      |

Values are presented as the mean ± SD unless noted otherwise.

aData are available 353, 353, 331 participants for the biceps femoris, semitendinosus, and semimembranosus, respectively.
Table 2. Participant characteristics by *COL5A1* genotype in Study 1 (n = 363)

|                  | CC    | TC    | TT    | T-dominant (TT+TC vs. CC) |
|------------------|-------|-------|-------|---------------------------|
| n                | 253   | 100   | 10    |                           |
| Sex, % men       | 60.8  | 71.0  | 60.0  | 0.097                     |
| Age, year        | 20.5 ± 1.9 | 20.7 ± 2.3 | 19.5 ± 1.5 | 0.597                     |
| Height, cm       | 169.0 ± 9.2  | 170.6 ± 8.7  | 166.4 ± 7.7  | 0.227                     |
| Body mass, kg    | 63.5 ± 10.6   | 66.2 ± 13.8   | 61.3 ± 7.2   | 0.084                     |
| PSLR ROM, degree | 83.4 ± 18.0   | 81.0 ± 17.2   | 83.4 ± 12.2  | 0.274/0.7                 |
| SR ROM, cm       | 10.4 ± 9.3    | 9.8 ± 10.2    | 13.4 ± 7.1   | 0.781/0.6                 |
| Stretch, % yes   | 78.2  | 72.0  | 80.0  | 0.261                     |

Values are presented as the mean ± SD unless noted otherwise.
*For PSLR and SR, adjusted for sex and regular stretching.*
Table 3. Characteristics of subjects in muscle injury and no muscle injury group in Study 2

|                          | Muscle injury  | No muscle injury |          |
|--------------------------|----------------|-----------------|----------|
|                          | (n = 190)      | (n = 1369)      |          |
| Sex                      |                |                 |          |
| Men, n (%)               | 134 (70.5)     | 929 (67.9)      |          |
| Women, n (%)             | 56 (29.5)      | 440 (32.1)      |          |
| Age, years               | 20.1 ± 1.7     | 20.5 ± 2.8      |          |
| Height, cm               | 169.9 ± 7.9    | 169.5 ± 8.3     |          |
| Body mass, kg            | 64.4 ± 10.4    | 63.4 ± 10.1     |          |
| Playing years, years     | 10.4 ± 3.8     | 11.2 ± 3.9      |          |
| Competitive level        |                |                 |          |
| International n (%)      | 17 (9.0)       | 153 (11.2)      |          |
| National n (%)           | 112 (59.0)     | 756 (55.2)      |          |
| Regional n (%)           | 30 (15.8)      | 286 (20.9)      |          |
| Other n (%)              | 31 (16.3)      | 174 (12.7)      |          |
| Main sport               |                |                 |          |
| Track & field            | 91 (47.9)      | 341 (24.9)      |          |
| Soccer                   | 62 (32.6)      | 599 (43.8)      |          |
| Other                    | 37 (19.5)      | 429 (31.3)      |          |

Values are presented as the mean ± SD unless noted otherwise.
## Table 4. Associations of COL5A1 rs12722 genotype with muscle injury

| Genotype | n (%)         | Muscle injury | No muscle injury | OR [95% CI]   |
|----------|---------------|---------------|------------------|---------------|
|          |               |               |                  |               |
| All      |               |               |                  |               |
| CC       | 135 (71.1)    | 935 (68.6)    | 0.87 [0.62-1.22] |
| TC       | 52 (27.4)     | 400 (29.2)    |                  |               |
| TT       | 3 (1.6)       | 34 (2.5)      |                  |               |
| Men      |               |               |                  |               |
| CC       | 95 (70.9)     | 628 (67.6)    | 0.88 [0.57-1.30] |
| TC       | 37 (27.6)     | 277 (29.8)    |                  |               |
| TT       | 2 (1.5)       | 24 (2.6)      |                  |               |
| Women    |               |               |                  |               |
| CC       | 40 (71.4)     | 307 (69.8)    | 0.86 [0.45-1.59] |
| TC       | 15 (26.8)     | 123 (28.0)    |                  |               |
| TT       | 1 (1.8)       | 10 (2.3)      |                  |               |

CI: confidence intervals, OR: odds ratio. Values were adjusted by playing years and main sport.

CC genotype was considered as the reference (OR = 1.00).

**Figures**
Figure 1

Stiffness of the biceps femoris long head (n=353), semitendinosus (n=353), and semimembranosus (n=331) muscles by the COL5A1 rs12722 genotype. Data are expressed as the mean ± standard deviation.
Figure 2
Proportion of severity of muscle injury by the COL5A1 rs12722 genotype.

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.
COL5A1_Supplemental Table_BMC_Final.xlsx