Accepted Manuscript

Submission Date: 2022-11-12
Accepted Date: 2023-05-30
Accepted Manuscript online: 2023-05-30

International Journal of Sports Medicine

The genetic association with athlete status, physical performance and injury risk in soccer

Conall F Murtagh, Elliott C Hall, Thomas E Brownlee, Barry Drust, Alun G Williams, Robert M Erskine.

Affiliations below.

DOI: 10.1055/a-2103-0165

Please cite this article as: Murtagh C F, Hall E C, Brownlee T E et al. The genetic association with athlete status, physical performance and injury risk in soccer. International Journal of Sports Medicine 2023. doi: 10.1055/a-2103-0165

Conflict of Interest: The authors declare that they have no conflict of interest.

Abstract:
The aim of this review was to critically appraise the literature concerning the genetic association with player status, physical performance and injury risk in soccer. The objectives were to provide guidance on which genetic markers could potentially be used as part of future practice in soccer; and to provide direction for future research in this area. The most compelling evidence identified six genetic polymorphisms to be associated with soccer athlete status (ACE I/D; ACTN3 rs1815739; AGT rs699; MCT1 rs1049434; NOS3 rs2070744; PPARA rs4253778), six with physical performance (ACTN3 rs1815739; AMPD1 rs17602729; BDNF rs6265; COL2A1 rs2070739; COL5A1 rs12722; NOS3 rs2070744), and seven with injury risk (ACTN3 rs1815739; CCL2 rs2857656; COL1A1 rs1800012; COL5A1 rs12722; EMILIN1 rs2289360; IL6 rs1800795; MMP3 rs679620). As well as replication by independent groups, large-scale genome-wide association studies are required to identify new genetic markers. Future research should also investigate the physiological mechanisms associating these polymorphisms with specific phenotypes. Further, researchers should investigate the above associations in female and non-Caucasian soccer players, as almost all published studies have recruited male participants of European ancestry. Only after robust, independently replicated genetic data have been generated, can genetic testing be considered an additional tool to potentially inform future practice in soccer.

Corresponding Author:
Dr. Robert M Erskine, Liverpool John Moores University, School of Sport and Exercise Sciences, Liverpool, United Kingdom of Great Britain and Northern Ireland, R.M.Erskine@ljmu.ac.uk

Affiliations:
Conall F Murtagh, Liverpool John Moores University, School of Sport and Exercise Sciences, Liverpool, United Kingdom of Great Britain and Northern Ireland
Conall F Murtagh, Liverpool Football Club and Athletic Grounds Ltd, Sports Science Department, Liverpool, United Kingdom of Great Britain and Northern Ireland
Elliott C Hall, Liverpool John Moores University, School of Sport and Exercise Sciences, Liverpool, United Kingdom of Great Britain and Northern Ireland
[...]
Robert M Erskine, University College London, Institute of Sport, Exercise and Health, London, United Kingdom of Great Britain and Northern Ireland

The genetic association with athlete status, physical performance and injury risk in soccer

Conall F. Murtagh[1,2], Elliott C. R. Hall[1,3], Thomas E. Brownlee[4], Barry Drust[4], Alun G. Williams[5,6] and Robert M. Erskine[1,6]

[1]School of Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, UK; [2]Liverpool Football Club, Liverpool, UK; [3]Faculty of Health Sciences and Sport, University of Stirling, Stirling, UK; [4]School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Birmingham, UK; [5]Department of Sport and Exercise Science, Manchester Metropolitan University, Manchester, UK; [6]Institute of Sport, Exercise and Health, University College London, London, UK.

Address for correspondence:

Dr Rob Erskine, School of Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, L3 3AF, United Kingdom.

Email: R.M.Erskine@ljmu.ac.uk;

C. F. Murtagh ORCID: 0000-0002-3924-7452

E. C. R. Hall ORCID: 0000-0003-0540-0735

T. E. Brownlee ORCID: 0000-0002-3355-1867

B. Drust ORCID: 0000-0003-2092-6962

A. G. Williams ORCID: 0000-0002-8052-8184

R. M. Erskine ORCID: 0000-0002-5705-0207

Running title: Genetic variation and football

Abstract

The aim of this review was to critically appraise the literature concerning the genetic association with player status, physical performance and injury risk in soccer. The objectives were to provide guidance on which genetic markers could potentially be used as part of future practice in soccer; and to provide direction for future research in this area. The most compelling evidence identified six genetic polymorphisms to be associated with soccer athlete status (*ACE* I/D; *ACTN3* rs1815739; *AGT* rs699; *MCT1* rs1049434; *NOS3* rs2070744; *PPARA* rs4253778), six with physical performance (*ACTN3* rs1815739; *AMPD1* rs17602729; *BDNF* rs6265; *COL2A1* rs2070739; *COL5A1* rs12722; *NOS3* rs2070744), and seven with injury risk (*ACTN3* rs1815739; *CCL2* rs2857656; *COL1A1* rs1800012; *COL5A1* rs12722; *EMILIN1* rs2289360; *IL6* rs1800795; *MMP3* rs679620). As well as replication by independent groups, large-scale genome-wide association studies are required to identify new genetic markers. Future research should also investigate the physiological mechanisms associating these polymorphisms with specific phenotypes. Further, researchers should investigate the above associations in female and non-Caucasian soccer players, as almost all published studies have recruited male participants of European ancestry. Only after robust, independently replicated genetic data have been generated, can genetic testing be considered an additional tool to potentially inform future practice in soccer.

Key words: physical performance, SNP, polymorphism, DNA, injuries, football

Introduction

Human physical performance and injury risk are influenced by various environmental and genetic factors [1, 2]. The heritability of being a high-level athlete across a range of sports is estimated to be 66% [2], which is itself dependent on the heritability of the other phenotypes, including risk of soft-tissue injury [3, 4] and the physiological and anatomical characteristics known to contribute to competitive sport performance [5-8]. Although such characteristics, e.g. strength, power, speed, agility and aerobic capacity [9-17], and injury risk factors, e.g. previous injury, maturity status, playing position, match intensity, etc. [18-24], have been extensively investigated in soccer players, the genetic contribution to high-level soccer success remains unclear. Variation in the sequence of nucleotides within or adjacent to certain genes has the ability to alter the type/abundance/function of the proteins encoded for by those specific genes [25], thus having the potential to affect tissue characteristics [26, 27] and ultimately physical performance [28] or injury risk [29]. Numerous case-control studies have identified genetic variants (e.g. single nucleotide polymorphisms, or SNPs) that distinguish talented soccer players from non-athletes [28, 30-34], while some studies have investigated the association of specific SNPs with physical performance [28, 35-37] and injury risk [22, 38-47] in soccer players. However, some study designs are limited by small sample sizes and/or cohorts of mixed sports (where soccer is but one), which restrict the ability to detect a genetic association due to a lack of statistical power, or may reveal false positives. Such examples can confuse the perceived importance of genetic variation in high-level soccer. The purpose of this narrative review was therefore to critically appraise the scientific literature concerning the genetic association with high-level athlete status, physical performance and injury risk in soccer. Having identified robust evidence, the main objectives were to provide guidance for researchers and practitioners on which genetic markers could potentially be used to identify talent and/or manage the training/recovery/injury prevention of

46 talented players as part of future practice; and to provide direction for future research in this
47 area.

48

49 **Literature Search**

50 A literature search was conducted for empirical research studies and review articles using
51 PubMed and SPORTDiscus databases from inception to March 2023. Key terms were
52 searched for within the article title, abstract, and keywords using conjunctions "OR" and
53 "AND" with truncation "*." Combinations of the following Boolean's phrases comprised the
54 search terms: 'Soccer'; 'Football'; 'Genetic'; 'Genotype'; 'Polymorphism'; 'Power'; 'Speed';
55 'Sprint'; 'Acceleration'; 'Agility'; 'Strength'; 'Force'; 'Aerobic'; 'Fitness'; 'Endurance';
56 'Injury'. The included publications met the following criteria: (1) contained relevant data
57 concerning genomics in soccer; (2) included soccer players; and (3) were written in English.
58 Studies were excluded on the basis that they: (1) did not contain any relevant genomics data;
59 and (2) were conference abstracts. Studies were not included/excluded on the basis of soccer
60 athlete status (i.e. standard of competitive soccer player) or the biological sex or geographic
61 ancestry of participants.

62

63 **The genetic association with soccer athlete status**

64 Numerous case-control studies have associated specific gene variants with soccer athlete
65 status [28, 30-34]. Logically, and all else being equal, a genotype or allele that is more
66 prevalent in high-level soccer players than non-athletes may have enhanced soccer
67 performance or reduced susceptibility to injury, thereby increasing the likelihood of that
68 soccer player becoming successful. "Success" and "high-level athlete status" are defined here
69 as being selected to represent a professional soccer club (including their academy for youth

athletes) and competing at the highest level in that particular country. Whilst acknowledging the limitations of this broad classification, we will inform the reader of the athlete status in each study, where that information is available.

Perhaps the most investigated gene variant regarding high-level athlete status is the *ACTN3* rs1815739 SNP. *ACTN3* XX homozygotes cannot produce alpha-actinin-3 [48], a structural sarcomeric protein found in type II skeletal muscle fibres that inhibits the transition of larger, stronger, more powerful type II fibres towards a smaller, weaker more oxidative phenotype [26]. This probably explains the association between this SNP and skeletal muscle fibre type composition, with XX homozygotes presenting with a greater proportion of slow twitch muscle fibres [49, 50]. Type I fibres are smaller and contract slower than type II fibres, and therefore produce less force and power than type II fibres [51, 52]. Therefore, it follows that muscle volume (the main determinant of power [53]), maximum strength and power are all lower in XX homozygotes compared to R-allele carriers, which appears to be the case in young, healthy men [27]. Regarding high-level soccer, evidence suggests that professional male players (n = 60) have a higher frequency of the *ACTN3* R allele compared to 123 non-athletes [34]. Hence, the *ACTN3* rs1815739 SNP could be a determinant of high-level soccer playing status, as it appears to predispose players to greater muscular power performance, although it should be noted that the sample size of this relatively early study was relatively small, thus larger-scale studies are required to confirm this finding.

More recently, it has been reported that the frequency distribution of the *ACTN3* rs1815739 R allele (67.5 vs. 60.4%), the angiotensin-I converting enzyme insertion/deletion polymorphism (*ACE* I/D) D allele (65.0 vs. 48.9%), peroxisome proliferator-activated receptor α (*PPARA*) rs4253778 C allele (24.3 vs. 17.3%) and uncoupling protein 2 (*UCP2*) rs66033955 T allele (44.4 vs. 35.8%) were higher in 246 male soccer players representing clubs in the Russian Premier (n = 51), National and Second Division leagues (n = 81) and

academies (114 youth players aged 10.6 ± 0.1 years), compared to 872 non-athletes [30] (Table 1). As the *ACTN3* R, *ACE* D and *PPARA* C alleles have previously been associated with strength/power phenotypes in over 20 case-control studies, and *UCP2* rs66033955 has been associated with endurance performance [54], these results provide evidence to support the hypothesis that, whilst muscular endurance is important to meet the aerobic demands of high-level soccer, strength, power and speed capabilities may play a more important role in determining match and individual player success [10, 13, 55, 56].

In accordance with such findings in Russian soccer players [30], Juffer and colleagues [32] reported that the *ACE* I/D II genotype was less common in 54 professional male soccer players compared to 52 endurance athletes [32], thus suggesting that endurance is not as important as strength and power in professional soccer players. However, similar to the study by Santiago and colleagues [34], the sample size of this relatively early study is modest. Interestingly, these findings are somewhat in contrast to the study by Eynon et al. [57], who reported a higher frequency of the *NOS3* rs2070744 C allele in a similar sample (n = 60) of professional male players compared to 100 non-athletes, while the opposing T allele has previously been found to be more common in high-level power-oriented athletes than non-athletes [58, 59]. Contrary to the study by Eynon et al. [57], the T allele was found to be more common in 536 high-level academy players than non-athletes [28] (Table 1). The reason for this discrepancy between studies is not clear, especially considering the studies by Eynon et al. [57], Santiago et al. [34] and Juffer et al. [32] all appear to have been performed in the same population (i.e. La Liga soccer players). In further support of the association between strength and power genetic profiles and soccer playing status, academy male soccer players from Italian clubs were associated with a higher frequency of the *VDR* rs10735810 FF genotype [33], which has previously been associated with greater maximum isometric quadriceps strength in patients with chronic obstructive pulmonary disease and age-matched

controls [60]. However, a study investigating the role of the *ACE* I/D polymorphism in 199 professional male soccer players representing Lithuanian clubs found a lower frequency of *ACE* DD genotype compared to non-athletes, thus suggesting that endurance capacity may be more important for professional soccer performance in the Lithuanian National League [31]. Such research implies that specific genetic profiles may be advantageous for high-level soccer performance in different countries, perhaps due to between-country differences in playing demands [61] and/or differences in player anthropometry [62].

Indeed, research involving 535 high-level academy male soccer players from multiple clubs across two continents (an English Premier League academy and academies of the highest national category in Uruguay) found that academy players had a higher frequency of the *PPARA* rs4253778 C allele, *AGT* rs699 G allele and *NOS3* rs2070744 T allele than age-matched non-athletes [28] (Table 1). Interestingly, all three of these alleles have previously been associated with elite power athlete status [63-66] and superior power performance [35, 63-65, 67]. However, the physical determinants of high-level youth soccer playing status are known to depend on maturity status [28], so any investigation regarding genetic profiling in youth athletes should account for the confounding effect of maturation. Indeed, when segregated according to peak height velocity (PHV, i.e. the peak rate of somatic growth that is associated with puberty), academy male soccer players have distinct genetic profiles, with post-PHV players demonstrating a greater frequency of the *PPARA* rs4253778 C allele (25.8% *vs.* 19.0% *vs.* 18.5%) and the *AGT* rs699 G allele (53.5% *vs.* 44.8% *vs.* 42.7%) compared to pre-PHV players and non-athletes, respectively [28]. Moreover, the *ACTN3* rs1815739 XX genotype was more common in pre-PHV players than non-athletes (29.0% *vs.* 15.9%, respectively). Thus, the differences in *PPARA* rs4253778, *AGT* rs699 and *ACTN3* rs1815739 allele/genotype frequency distributions between academy players and non-athletes suggest that the polygenic profile of pre-PHV academy players favours fatigue resistance,

rather than power/speed, and that post-PHV academy players would demonstrate a polygenic profile more advantageous for power/speed performance [28] (Table 1).

Recent research has also documented that the G alleles of two SNPs in *ADRB2* (rs1042713 and rs1042714), which have previously been associated with power athlete status in a Polish population [68], were over-represented in a small group (n=48, age 12-18 years) of academy male soccer players compared to non-athletes [36]. However, this study had a limited small sample size and did not account for the confounding factor of maturity status, thus, results should be interpreted with caution. Moreover, this study only included academy soccer players from a single club, which means the findings are not necessarily relevant in other academies, either within the same national league, or between nations.

In another study, Massidda and colleagues [69] recruited 694 professional male soccer players from Italy, Poland, Lithuania and Ukraine, and investigated the association between the *MCT1* (monocarboxylate transporter-1) rs1049434 SNP and playing position (by dividing players into four sub-groups: defenders, midfielders, forwards, and goalkeepers). In this study, the AA genotype was more frequent in forwards (n = 148) than non-athletes (n = 781). As this SNP is thought to be associated with a higher lactate clearance [70-72], it was postulated to contribute to the previously documented better repeated sprint ability (RSA) in forwards [73], who complete more sprints during a match than midfielders and defenders [74]. The results from this study [69], and analyses documenting the unique demands of each position in high-level soccer match-play [75], suggest that future case control genetic studies in soccer should consider playing position. However, research measuring the activity profile (and fitness requirements) of each position recommends that, due to the contrasting match-play demands [75], there should also be a separation between central and lateral midfielders and defenders. It is therefore recommended that future genetics studies recruit a large enough sample to gain sufficient statistical power to perform playing position analyses by stratifying

six positions in total (i.e. goalkeepers, wide defenders, central defenders, wide midfielders, central midfielders, and forwards).

A further two studies have investigated the genetic association with playing position in professional men's soccer [76, 77]. However, regarding the above points and the limited number of soccer players (n=43) of mixed geographic ancestry (and the lack of control group) in the study by Clos et al [76], it is unlikely that this study had sufficient statistical power to detect an association of genotype with playing position. Although a recent study by Petr et al. [77] did include a control group (n=107) and a larger cohort of soccer players (n=99), the latter sample size was still relatively modest, considering this needed to be segregated into much smaller groups for playing position and genetic analyses.

In summary, there is convincing evidence that genetic variation is associated with high-level soccer playing status, and that this association favours strength and power. However, due to physical and technical requirements varying between national leagues [61], and to be representative of high-level soccer on a global scale, case-control genetic association studies should include large sample sizes from multiple clubs in different countries and continents. Further, the fact that physical performance [10] and physiological factors associated with muscular power [78] differ according maturity status, not accounting for maturation is likely to confound any genetic association with playing status in academy soccer players. Studies that have segregated such players according to maturity status have shown that pre-PHV players have a genetic profile that favours endurance performance, while the genetic profile of post-PHV (physically mature) players is more aligned with strength and power (in accordance with that of senior professional players). Finally, when investigating the genetic association with soccer playing position, studies should include a sample size large enough to account for the stratification into multiple positional groups and subsequent genotype groups.

The genetic association with muscular power in soccer players

While it is important to establish if a specific combination of gene variants may characterise high-level soccer in different countries, case-control studies alone do not reveal the relationship between specific genetic polymorphisms and physical performance capabilities. Cross-sectional studies, investigating associations between gene variants and specific soccer performance predictors, may provide more diagnostic information for applied soccer practitioners.

Previous research has found that acceleration (10 m speed) and sprint (20 m speed) capabilities are important determinants of high-level male soccer player status in academy players at all stages of maturity [10, 79], whereas muscular power (horizontal and vertical jump performance) was only a determinant of playing status in post-PHV, *not* pre-PHV players [10]. The genetic association with explosive performance capacities could be due to certain gene variants influencing the protein abundance/expression within key tissues (e.g. skeletal muscle, tendon), which, in turn, affect the mechanical properties of those tissues and, thus, the players' force/power/speed-generating capacity. Of these specific traits, acceleration, sprint and agility performance have been shown to be independent capabilities [80] underpinned by a different combination of physiological factors. While early acceleration is associated with longer ground contact times (0.12-0.20 s) and relies on contractile force capabilities [81], sprinting is associated with shorter ground contact times (0.09-0.12 s) and therefore relies more on the ability of the muscle-tendon unit to perform fast stretch-shortening cycle actions [82]. Agility is associated with unilateral reactive strength capabilities [83] and the main physiological underpinning factors appear to be a combination of neural (inter-muscular co-ordination, neural firing frequency and motor unit recruitment and synchronization) and muscle-tendon unit mechanical properties. Similarly,

vertical and horizontal-forward countermovement jump (CMJ) capabilities are independent qualities [13], and are controlled by different co-ordination strategies [84], with horizontal-forward CMJs requiring greater biceps femoris activation than vertical CMJs [85]. As power assessed in the vertical and horizontal planes are underpinned by independent physiological factors in high-level soccer players, we will discuss the current body of literature documenting the genetic association with each performance capacity separately.

Vertical jump performance

One of the main physiological determinants of vertical jump performance in soccer players is quadriceps femoris muscle volume [11]. Indeed, genetic variants associated with muscle size and hypertrophy have also been associated with favourable vertical jump capabilities [35, 37]. The *ACTN3* rs1815739 R allele has previously been associated with increased serum testosterone [86], greater muscle volume [27] and a greater composition of type II skeletal muscle fibres [49], which are larger (therefore produce more force) and able to shorten quicker, thus able to generate more power, than type I fibres [51, 52]. Pimenta and colleagues [37] documented that the *ACTN3* rs1815739 R allele was associated with greater bilateral vertical CMJ performance in 200 professional male soccer players from the Brazilian premier division (Table 1). Furthermore, in 220 Brazilian youth players aged 14-20 years, *ACTN3* rs1815739 R-allele carriers and *ACE* DD homozygotes jumped higher than their XX and I-allele carrying counterparts [87], although it should be noted that these authors did not account for maturity status in these participants, which may have affected the strength of their associations. These results are partly supported by Ginevičienė et al. [35], who showed that *ACTN3* RR genotype was associated with greater bilateral vertical CMJ performance in 152 male regional or national level athletes, although only 32 were soccer players. The same study also found that *PPARA* rs4253778 (another SNP associated with muscle hypertrophy

and type II skeletal muscle fibre composition [88]) CC homozygotes also had greater estimated total body muscle mass and outperformed G-allele carriers in the vertical jump assessment. However, the results of this study should be treated with caution due to the non-elite standard of athletes, the small sample size of Caucasian only athletes, and the mix of sports the athletes competed in, of which only a small proportion were soccer players. Meanwhile, Massidda and colleagues [89] recruited 90 high-level, professional, male Italian soccer players and performed a total genotype score regression analysis, demonstrating that the combination of the *ACTN3* rs1815739, ACE *I/D* and bradykinin receptor B2 (*BDKRB2*) rs1799722 polymorphisms explained 18–24% of the variance in vertical jump performance. However, there is limited research on the association between polymorphisms of the *BDKRB2* gene and power performance, and the *ACTN3* rs1815739 [28, 90], *ACE* I/D [91] and *PPARA* rs4253778 [28] polymorphisms have not always shown consistent associations with vertical CMJ performance in high-level soccer players, so more research is required to confirm such findings.

While greater quadriceps femoris muscle volume is thought to allow the knee extensors to produce more force when propelling the body in the vertical direction [11], research has also found an association between greater compliance of the *vastus lateralis* tendon–aponeurosis complex and higher vertical CMJ performance [92]. *COL5A1* rs12722 CC genotype has previously been associated with a more extensible patellar tendon [93], thus, it is possible that the tendon of CC homozygotes can store and release more energy, enabling power to be amplified during vertical jumps. More recent research showed that, in 535 academy male soccer players (age 8-23 years), *COL5A1* rs12722 CC homozygotes achieved greater bilateral vertical CMJ height than TT homozygotes [28] (Table 1). Assuming that *COL5A1* rs12722 CC homozygotes have more elastic tendons than TT homozygotes [93], the findings by Murtagh et al. [28] are somewhat in agreement with Kubo

and colleagues [92], who reported that greater compliance of the VL tendon–aponeurosis complex facilitates improved CMJ performance. Interestingly, the association reported by Murtagh et al. [28] was only significant in the academy soccer players (and not in non-athletes), possibly as the soccer players could co-ordinate the vertical CMJ actions better, thus gaining greater utilisation of the stretch shortening cycle and elastic tendon recoil properties. Nevertheless, this has only been investigated in one cohort and the association of *COL5A1* rs12722, and other SNPs related to tendon elasticity, needs to be further investigated. Overall, it appears that studies targeting genetic polymorphisms that are thought to influence the physiological determinants of vertical jump performance could help elucidate if genetic variation can influence such a task.

Horizontal jump performance

Horizontal jump performance is a key determinant of soccer playing status in physically mature academy soccer players [10, 13]. Interestingly, this performance measure was also shown to be a key physical factor at a young age, influencing future contract status and playing minutes after reaching professional status [94]. In a cohort of 535 academy male soccer players and 151 age- and sex-matched non-athletes, *BDNF* rs6265 CC homozygotes demonstrated greater horizontal CMJ distance [28]. Neural function is largely determined by neurotrophins, of which BDNF is one of the most active [95], regulating neuronal survival, growth, maintenance, neurogenesis and synaptic plasticity [96]. As the rs6265 C allele is associated with a greater abundance of exercise-induced serum BDNF concentration [97], the findings by Murtagh and colleagues [28] suggest that individuals with potentially enhanced neuromuscular characteristics perform better in the performance tasks that require the body to be propelled in the horizontal-forward direction. Such findings are in agreement with previous research documenting that high-level soccer players produce greater hamstring

activation during horizontal compared to vertical jumping [13], and the horizontal jump is believed to be a more complex action, requiring more inter-muscular co-ordination [84, 85].

Overall, there is limited research investigating the relationship between genetic variants and horizontal jump performance in soccer players but the study by Murtagh and colleagues [28] suggests the *BDNF* rs6265 SNP plays a role. Considering the importance of horizontal power for determining high-level soccer playing status and future contract status, more research investigating the association between genetic variation and such a key performance variable is of paramount importance if we are to fully understand the genetic contribution to success in soccer.

The genetic association with maximum strength in soccer players

As power is underpinned by strength, it follows that strength is likely an important performance measure in soccer players. Indeed, previous work has shown that high-level academy soccer players are stronger (using the isometric mid-thigh pull strength test) than age-matched non-athletes [9]. Although numerous studies have found genetic associations with strength in different populations [98-102], to our knowledge, there are no published studies that have investigated a genetic association with strength in soccer players. However, work from a doctoral thesis did explore the association of four SNPs (*PPARA* rs4253778, *NOS3* rs2070744, *COLIA1* rs2249492 and *VDR* rs2228570) with isometric mid-thigh pull maximum voluntary force (MVF) in 148 high-level male academy soccer players and 93 age- and sex-matched non-athletes [103]. Although there were no associations with MVF regarding any of these four SNPs, subsequent genotyping for a further six SNPs (*ACTN3* rs1815739; *AGT* rs699; *AMPD1* rs17602729; *BDNF* rs6265; *COL2A1* rs2070739; *COL5A1* rs12722) revealed that *COL2A1* rs2070739 CC homozygotes were stronger than T allele

carriers, regardless of maturity and athlete status (Fig. 1). This association may be due to the SNP influencing RNA stability, potentially altering the structure of the collagen network in and around the muscle (as proposed by Baumert et al. [104]). Increasing the number of lateral attachments within the muscle has the potential to increase the force per muscle cross-sectional area, thus augmenting MVF [53, 105], although this hypothesis remains speculative. It should be noted that the isometric mid-thigh pull test used in the aforementioned study is an applied measure of strength commonly used in the field. However, it is not specific to a particular muscle group and may therefore lack the sensitivity required to detect more than this one genetic association with strength in soccer players. Thus, future studies may wish to incorporate strength tests that isolate key muscle groups, such as the knee extensors, knee flexors or plantar flexors.

Although strength was not measured, a separate study demonstrated that fat-free mass was relatively greater in T-allele carriers of the *MCT1 A1470T* SNP, compared to AA homozygotes in physically mature (~16-18 years of age) academy male soccer players [106]. Due to the association between muscle size and strength [53, 107], it is feasible that this SNP may also be associated with MVF in physically mature soccer players, but this has not yet been investigated. In summary, there is very limited research investigating the genetic association with strength in soccer players and, given the importance of strength in soccer [9], further work is required to elucidate more genetic markers associated with strength in soccer players.

Insert Figure 1 near here.

The genetic association with acceleration and sprint performance

Accelerating and sprinting require soccer players to optimise the production of peak power to propel the body forward in a cyclical motion. In academy soccer players aged ~18 years, it has been shown that peak power during vertical CMJ is correlated with quadriceps femoris muscle volume [11], and running speed has recently been associated independently with greater muscle volume [108]. The *ACE* D [109, 110], *NOS3* rs2070744 T [58] and *AGT* (rs699) G [63, 65] alleles have all been associated with elite power athlete status and are thought to exert their favourable effect on power performance by promoting skeletal muscle hypertrophy [63, 65, 111]. Indeed, in a cohort of 212 high-level Brazilian professional male players, *ACE* DD homozygotes demonstrated faster 20 m sprint speeds compared to players of II genotype but there was no difference in muscular power (vertical CMJ) performance [91]. In a separate study from the same population (200 high-level Brazilian professional male players), *ACTN3* rs1815739 RR players sprinted faster over 10, 20 and 30 m and jumped higher than XX homozygotes [37]. Furthermore, in 220 Brazilian youth players aged 14-20 years, *ACTN3* rs1815739 R-allele carriers and *ACE* DD homozygotes sprinted faster than their XX and I-allele carrying counterparts [87], although it should be noted that this study did not account for the participants' maturity status, which may have affected the strength of the associations. Finally, in 535 high-level male academy soccer players, *NOS3* TT genotype was not associated with muscular power (neither horizontal nor vertical jump performance), but TT homozygotes did exhibit greater acceleration/sprint performance than their C-allele carrier counterparts (regardless of maturity status) [28] (Table 1). It could therefore be postulated that the *ACE* DD, *ACTN3* RR and *NOS3* TT genotypes may stimulate an increase in the number of sarcomeres arranged in series, thus increasing muscle fibre length, which is the main determinant of maximal muscle fibre contraction velocity [53]. Alternatively, simply increasing strength (e.g. through muscle hypertrophy) is known to increase sprint performance [112], thus by influencing muscle size, these polymorphisms

may indirectly affect sprint performance. Furthermore, in a pilot study, comprising 26 male players from the under-19 youth team of a Serie A club in Italy, *MCT1* rs1049434 A-allele carriers performed better than players of TT genotype in the final two sprints of an RSA assessment (6 × 30 m sprints with an active recovery between sprints). This finding may be linked to the contribution of monocarboxylate transporter-1 (MCT1) to increased metabolic use of lactate after repeated sprinting activity [113]. However, given the limited sample size, results should be interpreted with caution. Finally, the *AGT* G allele has been associated with greater 5 and 20 m sprint performance in a small cohort (n = 48) of academy male soccer players aged 12-18 years [36]. Thus, given the age range of players, it is highly likely that there was variation in maturity status, which has been shown to influence the genetic association with athlete status and sprint performance in high-level academy soccer players [28]. Interestingly, no association was found between this SNP and acceleration/sprint performance in a much larger cohort of 535 high-level male academy players, where maturity status was accounted for [28]. Consequently, more large-scale research (including larger sample sizes) is required to confirm the associations between acceleration/sprint speed and the *NOS3* rs2070744 and *AGT* rs699 SNPs, and between RSA and the *MCT1* rs1049434 SNP.

While muscle fibre length (or specifically, the number of serial-aligned sarcomeres) is the main determinant of maximal muscle contraction velocity [53], the extensibility of the tendon is related to speed and horizontal-forward power capabilities in sprinters [114] and post-PHV academy soccer players [12]. Recent research shows that, at all stages of maturity, male academy players and age-matched non-athletes of *COL5A1* rs12722 CC genotype achieved quicker acceleration (10 m) and sprint (20 m) times than CT heterozygotes [28] (Table 1). As the CC genotype has previously been associated with a more extensible knee extensor tendon [93], it is feasible that the tendons of CC homozygotes might store and

release more energy, thus amplifying power during sprint and acceleration. As well as the aforementioned findings in *COL5A1* CC homozygotes, Murtagh et al. [28] showed that *COL2A1* rs2070739 CC homozygotes accelerated and sprinted quicker than their T-allele carrier counterparts, possibly due to having a greater capacity to store and release mechanical energy in the tendon. However, this association was only significant in non-athletes and not academy soccer players. It is possible that the specific training environment and physical interventions in elite soccer clubs, which are often tailored to improve individual athletic weaknesses, could have confounded the association between the *COL2A1* rs2070739 SNP and speed performance in academy soccer players. Overall, the limited research investigating the association between genetic variation (particularly SNPs potentially associated with tendon properties) and speed performance in soccer players is positive but future research in independent groups is required to confirm such findings.

In addition to having a more compliant patellar tendon [114], research showing that sprinters producing more horizontal force during acceleration are able to highly activate their hamstring muscles just before ground contact [115] suggests that neuromuscular activation is also a major factor in determining sprint performance. In the study by Murtagh et al. [28], *BDNF* rs6265 CC homozygotes demonstrated greater horizontal power, acceleration and sprint performance. Therefore, as the C allele is associated with a greater abundance of exercise-induced serum BDNF concentration [97], these findings suggest that individuals with potentially enhanced neuromuscular characteristics perform better in the performance tasks that require the body to be propelled in the horizontal-forward direction. Research in just 48 academy soccer players aligns with these findings, showing that three SNPs thought to be related with neural function were associated with 5 m (acceleration) and 20 m (sprint) speed performance [36]. The G alleles of the *CPNE5* rs3213537 and *CBLN2* rs8093502 SNPs (both genes are thought to play a role in synaptic function [116, 117]), and the A allele of

418 *CNTN4* rs62247016 SNP (the *CNTN4* gene encodes contactin 4, which is thought to play a
419 role in the formation of axon connections in the developing nervous system [118]), were all
420 associated with greater sprint performance [36]. However, the findings of this study should
421 be interpreted with caution, as not only was the sample size limited, but it comprised players
422 from just one academy, and maturity status was not accounted for. Overall, it seems that
423 SNPs associated with neural function/adaptation have a positive impact on sprint
424 performance; however, the research is limited and more evidence is required from large-scale
425 independent studies.

426 Since it is clear that neuromuscular function (e.g. inter- and intra-muscular co-
427 ordination [83]) primarily underpins speed performance prior to maturation, i.e. in pre-PHV
428 athletes [78], it seems logical that certain genetic variants may have a greater influence on
429 speed/power performance in pre-PHV soccer players. While no neural related genetic variant
430 has been shown to exert this association specifically in this population to date, *AMPD1*
431 rs17602729 GG homozygotes did achieve quicker acceleration and sprint times in pre-PHV
432 academy players only [28]. It is possible that pre-PHV *AMPD1* GG homozygotes can repeat
433 high intensity acceleration and sprint efforts more frequently during soccer match-play,
434 which may lead to preferential adaptations in acceleration and sprint performance in GG
435 homozygotes compared to their A allele-carrying counterparts. The *AMPD1* G>A SNP is the
436 only genetic variant to date that appears to influence acceleration and speed in pre-PHV
437 academy players on an individual SNP basis [28]. Such findings support the argument that
438 genetic research in academy soccer players should be specific to maturity status.

439

440 **The genetic association with agility performance**

Only one study to date has investigated the relationship between agility and genetic profile in soccer players [28]. The main physiological factors underpinning agility capacity appear to be a combination of neural (inter-muscular co-ordination, neural firing frequency and motor unit synchronization) and muscle-tendon unit mechanical properties. It is therefore logical that agility performance was associated with a combination of SNPs, rather than one individual genetic variant. Indeed, a total genotype score model derived from four SNPs (*AMPD1* G>A rs17602729, *BDNF* C>T rs6265, *COL5A1* C>T rs12722, *COL2A1* C>T rs2070739) that were individually associated with various power/sprint performance variables was shown to be positively correlated with agility performance in pre-PHV but not post-PHV academy male soccer players. These significant correlations in pre-PHV players could be due to the academy soccer environment increasingly impacting on agility performance (and overall physical performance) as the player matures. It has previously been demonstrated that training exposure in academy soccer increases (and includes more personalized athletic development and resistance training as well as soccer activities) as the players progress to older age group teams [9], thus suggesting environmental influences could increasingly confound any genetic effect. Nevertheless, such a novel polygenic analysis shows that, while one SNP may not have a significant impact on explosive performance, a combination of SNPs could have a favourable effect, especially in pre-PHV individuals.

The genetic association with aerobic capacity in soccer players

Due to the aerobic underpinnings of soccer [119], potential genetic associations with endurance performance could also influence high-level soccer player status and/or endurance capacity in soccer players. Perhaps the most well-known genetic variant associated with prolonged endurance performance is the *ACE* I/D polymorphism, i.e. the II genotype is commonly associated with greater endurance capacity/performance [120, 121], although not

always [122]. However, only two studies have investigated the association of the *ACE* I/D polymorphism with soccer athlete status [31, 32]. The study by Juffer and colleagues [32] found that the II genotype frequency was lower in professional male soccer players than endurance runners, supporting the hypothesis that high-level soccer players tend to have a power/strength oriented genetic profile, as suggested by others [28, 32]. However, the study by Ginevičienė and colleagues [31] found a lower frequency of DD genotype in 199 sub-elite professional male soccer players representing Lithuanian clubs compared to non-athletes. The contrasting results in these two studies [31, 32] suggest the influence of *ACE* genotype on soccer athlete status may depend on the standard of player (e.g. elite [32] *vs.* sub-elite [31]) and the associated differences in training/match demands.

A different study in 200 professional male soccer players incorporated an applied measure of aerobic capacity and found that *ACTN3* XX homozygotes had a greater Yoyo test score than RR and RX genotypes [37], i.e. those players with an 'endurance genotype' exhibited a greater aerobic capacity. In a different study of 212 participants from the same population, higher frequencies of the *ACE* II genotype (compared to I/D or DD genotypes) were observed in "excellent and good performance groups" regarding the YoYo intermittent recovery test [91]. In a separate study in 220 Brazilian youth soccer players, *ACE* II homozygotes were found to perform better in the YoYo test than D-allele carriers, although this was only in the under-16 year age group and thus a limited sample size of 41 [87]. Interestingly, a study in high-level academy male soccer players found that maturity status was associated with allele and genotype frequency distribution, with pre-PHV players having a greater frequency of *ACTN3* XX genotype [28], and post-PHV players demonstrating a greater frequency of the *PPARA* rs4253778 C allele and the *AGT* rs699 G allele compared to pre-PHV players and non-athletes. This suggests that pre-PHV academy players are recruited in part due to their high aerobic capacity. In English Category 1 soccer academies, male

players do not compete 11 vs. 11 on a full-sized pitch until they reach ~13 years of age (approximately when PHV occurs in boys [123]). While 11 vs. 11 match-play requires a greater percentage of time spent at low speeds and more frequent explosive sprint actions, match-play on smaller pitches with fewer players requires a greater distance per minute, more time running at moderate speeds, and a higher work load per minute [124], thus placing a greater physiological demand on aerobic rather than anaerobic capacity. The playing demands of high-level soccer therefore differ according to maturity status and, at pre-PHV level, aerobic capacity appears to be more important than power/ speed, which would explain the apparent endurance-favourable genetic profile of pre-PHV academy soccer players.

Another study combined the genotype score of five other SNPs (*VEGFA* rs2010963, *ADRB2* rs1042713 and rs1042714, *CRP* rs1205 and *PPARGC1A* rs8192678), and found that those male academy players with a favourable combination responded better to eight weeks' aerobic training when measured with the Yo-Yo intermittent recovery test level 1 [125]. Unfortunately, the authors of this study did not specify which alleles/genotypes were classed as "favourable", thus preventing replication by other groups, whether in research or applied practice. Furthermore, it should be noted that this study only included 42 participants, who were not associated with a high-level soccer academy. A combination of small sample size, low sensitivity of field-based performance measure, and small effect size may have limited statistical power in this study, thus these results should be interpreted with caution. In summary, very few studies have investigated the genetic association with endurance capacity in soccer players, and more rigorous research is required to provide more confidence that genetic variation influences these important phenotypes in soccer players.

Insert Table 1 near here.

The genetic association with injury risk in soccer players

Player availability is negatively affected by injury [126]. Consequently, team success is associated with low injury rates [127], having more players available for matches and training [128], and losing fewer days to injury during a season [18]. Despite scientific advances identifying a number of risk factors, including player age, previous injury and joint range of motion [1], as well as maturity status [24], playing position [23] and biological sex [129], inter-individual variation in the frequency and severity of soccer injury suggests not all factors are fully understood. With injury prevention representing a key objective for players and coaches, it is logical that athletes and practitioners believe there may be value in genetic data concerning injury susceptibility [130], not least because previous injury increases future injury risk [1].

The majority of soccer injuries occur in the lower extremities without external contact [20]. Muscle strains commonly occur in the quadriceps, hamstrings and calves [21], whilst the knee and ankle are the most common sites for ligament sprains [21]. Extensive mechanical stress during contractile activity, such as acceleration, deceleration and kicking, are implicated in muscle strain, whilst ligament sprains occur when collagenous fibres are excessively stretched as a joint is forced to go beyond its normative range. Common SNPs have the potential to influence the structure, function or expression of proteins within tissues and affect their phenotypes [25], and the heritability of soft-tissue injury is estimated to be as high as 40% [3]. Therefore, genetic differences between players might alter the mechanical properties of musculoskeletal soft-tissues, such as muscle, ligament and tendon, which are commonly injured in soccer [21, 131]. The complexity of soccer injuries means that any genetic influence is likely to be polygenic, rather than being due to a single SNP. The high cost and complexity of studying multiple SNPs and their interaction means that, to date, most

genetic associations with soccer injuries are from candidate gene studies. This approach involves selecting specific genes or variants known or suspected to contribute to a particular phenotype. For example, certain SNPs can influence the structure and/or function of collagenous tissues [132], and several SNPs in collagen genes have been associated with ligament and tendon injuries in athletes from other sports [133-135]. Whilst there are a growing number of SNPs related to injury in various cohorts [136], this section summarises those SNPs associated with soccer injuries, offering potential mechanisms where possible. It should be noted that a considerable number of SNPs have been associated with injury in cohorts across multiple sports [136], and that the aetiology of soccer injuries is often similar to injuries occurring in those sports. Consequently, it is possible that SNPs associated with injury in other sports, of which some are not discussed in this section, may also be associated with soccer injury.

ACTN3 rs1815739

As previously mentioned, carriers of the *ACTN3* rs1815739 R allele have larger and stronger muscles than XX homozygotes [27, 137], which is important because low muscle strength is a risk factor for soft-tissue injury [138]. Research suggests that professional male soccer players of *ACTN3* XX genotype exhibit greater exercise-induced muscle damage than RR homozygotes [139], which may help explain why this genotype has been associated with a higher rate of non-contact musculoskeletal soft-tissue injuries in professional male soccer players [140], although this study included just 43 participants (Table 2). Another study of 169 professional male players reported a higher rate of non-contact muscle injuries in XX homozygotes than R-allele carriers, with XX homozygotes most likely to suffer severe injury [44] (Table 2). A separate study in professional players found X-allele carriers were most likely to be injured, with XX homozygotes having the longest absence after injury [47].

565 However, this study comprised just 46 participants, of which 22 were male and 24 female. In
566 addition to the well-documented physiological and biomechanical differences between men
567 and women, soccer match demands [141] and injury risk [129, 142] also differ between
568 men's and women's soccer, thus pooling male and female players in the same study is
569 problematic when deciphering the genetic association with soccer injuries. However, it is
570 worth noting that this is one of only two studies to include female participants in a study
571 investigating the genetic association with injury in soccer players.

572 Elsewhere, no association was reported between the *ACTN3* rs1815739 SNP and
573 hamstring injuries in 107 male professional players [40], nor with injury prevalence in 402
574 male academy players [29]. Similarly, a recent study investigating the *ACTN3* rs1815739
575 SNP and injury in 191 female soccer players found no difference in non-contact injury
576 incidence between genotype groups [143]. Nevertheless, the study by Hall and colleagues
577 [29] demonstrated that high-level academy players with at least one X allele missed more
578 days following ankle injuries than RR homozygotes (Table 2), suggesting greater muscle size
579 and strength associated with the R allele [26, 27] may enhance limb stability and favour
580 return to play. It appears that the *ACTN3* rs1815739 SNP may influence soccer injuries,
581 where potential mechanisms include muscle strength enhancement and/or the provision of
582 structural stability to the sarcomere during contraction [144].

583

584 *COL1A1 rs1800012*

585 The rs1800012 SNP is a C>A substitution in the *COL1A1* gene [145], encoding the type I
586 collagen α1 chain. This is important when considering injury risk, as type I collagen is the
587 major fibrillar collagen providing structural stability to ligaments and tendons [146]. The rare
588 AA genotype appears protective against ACL [133] and Achilles tendon [147] injuries in

589 athletes, while professional male soccer players with ACL ruptures are also less likely to be
590 AA homozygotes [39] (Table 2), although others found no associations between this SNP and
591 injury [29, 40, 46]. Differing conclusions in these studies may relate to the specific injury
592 types investigated, as it appears *COL1A1* variants are most likely to influence injuries to
593 ligament and/or tendon. Nevertheless, the rs1800012 A allele has been associated (as part of a
594 haplotype) with significantly higher *COL1A1* gene transcription [148]. If higher *COL1A1*
595 transcriptional activity increases the structural stability of ligaments and tendons, this might
596 explain the apparent protective effect of the rare rs1800012 genotype against ACL and
597 Achilles tendon injuries.

598

599 *COL5A1 rs12722 and rs13946*

600 The rs12722 SNP is a C>T substitution within the *COL5A1* gene [149], encoding the type V
601 collagen α1 chain, a vital structural component of tendons and ligaments during collagen
602 fibrillogenesis [150]. Two studies with relatively small sample sizes (both n < 80) reported
603 severe hamstring injuries in TC heterozygous professional male soccer players [45, 46], with
604 no association reported elsewhere [40, 42]. More severe musculoskeletal injuries appear to be
605 experienced by TT homozygous players than C-allele carriers [42], whilst a case-control
606 study comparing players with and without ACL rupture found that carrying C alleles in two
607 different *COL5A1* SNPs (rs12722 and rs13946) was apparently protective [41] (Table 2). In
608 contrast, a higher prevalence of soft-tissue and ligament injuries has been reported in pre-
609 PHV academy male players carrying the C allele and CC genotype, respectively [29] (Table
610 2). Therefore, more research is required to determine which *COL5A1* rs12722 allele is indeed
611 protective and why, although from this limited evidence, it is possible that it is linked to the
612 maturity status of the player.

EMILIN1 rs2289360

The rs2289360 SNP is a T>C substitution within the *EMILIN1* gene [151], which encodes the elastin microfibril interface-1 (EMILIN-1) protein. This protein assists the fusion of elastin fibres during elastogenesis [152] and provides elasticity to ligaments and tendons [153]. In professional male soccer players, the TT genotype has been associated with more severe MCL injuries, but not with the incidence or severity of muscle or tendon injuries [38, 46] (Table 2). Another study reported no association with hamstring injuries in 107 professional male players [40], whilst in 402 academy male players, CC homozygotes had the highest overall injury prevalence and the 'protective' TT genotype was more frequent in post- than pre-PHV players [29] (Table 2). These findings suggest that fewer 'genetically pre-disposed' (i.e. CC homozygotes) players progress from pre- to post- PHV in a soccer academy. However, in the same study, academy players carrying the T allele experienced the most days missed following ankle injuries [29]. Together with the association of the TT genotype with MCL injury in professional players [38, 46], this suggests that this *EMILIN1* SNP might influence soccer-related ligament injury predisposition and recovery. However, contrasting results exist between and within studies regarding which rs2289360 genotypes and/or alleles are protective, requiring further investigation to understand a potential role in soccer injuries. Hall et al. [29] hypothesised that the C allele might augment collagen tissue stiffness, causing stiffer tendons in CC homozygotes and increasing injury risk as a result of greater strain on muscle fascicles during eccentric activity. In contrast, the T allele might protect against injury in general (but particularly muscle injury) as a product of increased tendon compliance. However, such increased compliance may be insufficient to withstand more excessive external loads, leading to greater tendon damage (and even rupture) requiring a longer period of healing and rehabilitation.

MMP3 rs679620 and rs3025058

The rs679620 SNP is a T>C substitution within the *MMP3* gene [154]. The MMPs, a family of collagen-degrading enzymes, are physiological regulators of extracellular matrix (ECM) remodelling [155] and are important for muscle regeneration [156]. The T allele was associated with hamstring injury in 107 professional male soccer players [40], and with longer absence following knee injuries in academy soccer players [29] (Table 2). As the T allele sits in linkage disequilibrium with the *MMP3* rs3025058 5A allele, which increases MMP3 expression [157], the rs679620 T allele might increase RNA expression and intensify the degradation of proteins, such as collagen and elastin, thereby weakening the structural integrity of ligament and ECM. However, further research is required to confirm this hypothesis.

CCL2 rs2857656

The rs2857656 SNP is a G>C substitution in the *CCL2* gene affecting plasma C-C motif chemokine ligand-2 (CCL2) levels [158]. CCL2 is expressed within the interstitial space between myofibres after damaging activity [159] and mediates systemic changes with chronic exercise [160]. In professional male soccer players (n = 73-74), those with the GG genotype suffered more severe muscle injuries than C-allele carriers [45, 46], although others report no association with hamstring injuries in 107 professional male players [40], nor with injury prevalence or recovery time in male academy players [29] (Table 2). Plasma CCL2 is lower in GG homozygotes [158], who exhibit lower muscle strength than C-allele carriers [161]. Considering the prospective role of CCL2 in satellite cell proliferation [162], higher CCL2

during recovery from muscle injury may be advantageous to players carrying at least one rs2857656 C allele.

IL6 rs1800795

The rs1800795 SNP is a C>G substitution within the *IL6* gene [163]. The myokine IL-6 is produced in skeletal muscle after exercise and is functionally related to growth, atrophy and collagen synthesis [164, 165]. Interestingly, the GG genotype was associated with greater hamstring injury risk in 107 professional male players [40] (Table 2). Contrastingly, CC homozygotes had higher prevalence of any injury and muscle injury in 402 academy male players [29] (Table 2). Having both pro- and anti-inflammatory capacities [166] makes it plausible that IL-6 could increase or reduce injury risk. Increased plasma IL-6 associated with the G allele [163] may heighten soft-tissue injury risk. Specifically, because eccentric contractions induce pro-inflammatory IL-6 expression [167], and because cytokines trigger tenocyte apoptosis and ECM degradation [168], GG homozygotes may experience these unfavourable changes via increased IL-6, thus heightening injury risk. This could explain the association of the GG genotype with hamstring injury risk [40] if eccentric contractions initiate a process of elevated ECM degradation above that which is required for repair and recovery. In contrast, increased injury risk in CC homozygotes reported in academy players [29] could relate to an increased inflammatory response following eccentric activity associated with the C allele [169], where recovery between intense training and/or competition may be impaired by sustained inflammation.

Insert Table 2 near here.

Future directions for genetics research in soccer

In addition to the SNPs highlighted above, several other SNPs have been identified as potentially important but require further investigation, e.g. due to low sample size or lack of replication by independent research groups. For example, there are currently no published studies that have investigated the genetic association with strength in soccer players, although previously unpublished work from our group suggests the *COL2A1* rs2070739 SNP is associated with maximum strength in 148 high-level male academy soccer players and 93 age- and sex-matched non-athletes, regardless of maturity status and athlete status (Fig. 1). Further, in a study with a limited sample size, the combined genotype score of five SNPs (*VEGF* rs2010963, *ADRB2* rs1042713 and rs1042714, *CRP* rs1205 and *PPARGC1A* rs8192678) was associated with a greater adaptation to eight weeks' aerobic training in academy players [125]. Thus, replication by independent groups is necessary to provide more confidence in the results from these studies.

Regarding the genetic association with injury risk, numerous studies have found individual SNP associations that require replication by independent groups. For example, *IGF2* rs3213221 GC genotype, *GEFT* rs11613457 GG genotype, *HGF* rs5745678 T allele and *LIF* rs9290271 T allele were suggested to be protective against severe non-contact injuries, whilst *HGF* rs5745697 CC and rs1011694 AA homozygotes suffered fewer injuries, albeit in relatively small sample sizes [45, 46]. Male academy players with at least one *MYLK* rs28497577 T allele required longer recovery time following knee injury [29], whilst professional male players with the *SOX15* GG genotype were injured more often than T-allele carriers [45], although not in another study by the same group [46]. Further, hamstring injury risk was associated with the *TNC* rs2104772 A allele, and with the *HIF1A* CC and *NOS3* rs1799983 GG genotypes, in 107 professional male players [40]. However, others have reported no association between the *TNC* rs2104772 SNP and injury [46], which could be due

to the latter study having a smaller sample size (n = 73, potentially reducing statistical power to detect an association). Recently, the *ACE* I/D D allele was associated with reduced muscle injury risk in 710 professional male players from Italy and Japan [43], and in 402 male academy players, those homozygous for the *VEGFA* rs2010963 C allele had a higher prevalence of ligament and tendon injuries [29], while the *MCT1* rs1049434 AA genotype has been associated with higher muscle injury incidence in 173 high-level soccer players [170]. It is pertinent to mention additional SNPs, which were not associated with soccer injury. These include no association between muscle injury and three *VDR* SNPs [171] or *COL5A1* rs12722 [42], between *COL12A1* rs970547 and ACL injury [39], between *TTN* rs2742327 and non-contact musculoskeletal injuries [46], or between hamstring injuries and 32 of 37 investigated SNPs [40].

As well as replication of robust studies by independent research groups, further investigation is required to conduct large-scale genome-wide association studies to identify new genetic markers associated with athlete status, physical performance and injury risk in soccer. To identify a substantial proportion of the gene variants with small effect sizes that collectively contribute to the genetic component of variability in a complex trait, such as athlete status, physical performance and injury, a sample of at least 2,000 is probably required [172]. Furthermore, only a handful of genetic markers have been investigated to the extent whereby their function is known to impact physiology/biomechanics, which is crucial if we are to truly understand the potential impact of genetic testing in soccer. Moreover, to date, all genetic studies in soccer players have recruited almost exclusively male participants. Given the differences in match demands between women's and men's soccer [141], as well as the physiological and biomechanical sex differences, and variability in injury risk between male and female soccer players [129, 142], we cannot assume that the genetic associations found with athlete status, physical performance and injury risk in male soccer players will be

735 the same for female players. Therefore, future research should investigate the genetic
736 association with athlete status, physical performance characteristics and injury risk in
737 women's soccer. Finally, almost all studies published to date have recruited participants of
738 European ancestry due to the evidence that genetic variation differs between geographic
739 ancestries [173]. To provide a more comprehensive understanding of the genetic contribution
740 to success as a soccer player, however, future studies should investigate the genetic
741 association with athlete status, physical performance characteristics and injury risk in soccer
742 players from different geographic ancestries.

743

744 **Summary**

745 Throughout this review, we have critically appraised the literature pertaining to the genetic
746 association with physical performance and injury risk in soccer players. Using the most
747 robust evidence, we have suggested directions for future research in this area, and provided
748 guidance on the most compelling evidence to help practitioners understand the importance of
749 genetic variation and its influence on soccer athlete status, physical performance (Table 1)
750 and injury risk (Table 2). Gaining a more detailed insight into the genetic contribution to
751 soccer-related physical performance and injury risk could potentially help *future* talent
752 selection and physical development procedures in elite soccer clubs. However, it should be
753 stressed that genetics research in soccer is still in its infancy, and future studies should
754 attempt to replicate those more robust studies highlighted in this review by recruiting large
755 cohorts from various leagues around the world, and distinguishing between maturity status
756 when recruiting academy players. Further, very few studies have investigated genetic
757 associations with athlete status, physical performance or injury in female players, or players
758 of non-European geographic ancestry, so future studies should include these populations.
759 Regarding the genetic association with soccer playing position, studies should include a

sample size large enough to account for the stratification into six positional groups (i.e. goalkeepers, wide defenders, central defenders, wide midfielders, central midfielders, and forwards) and subsequent genotype groups. Furthermore, the candidate gene approach, although practical, relies on a degree of prior knowledge and is restricted to the specific SNPs selected for analysis. However, reduced costs and improved accessibility to more powerful approaches, such as genome-wide association studies, may aid the discovery of new SNPs to be investigated in larger cohorts. Studies with broader scope in terms of sample size, injury frequency and the number SNPs investigated are required to advance the existing evidence for a genetic contribution to soccer injuries. If large-scale studies following our recommended criteria could be initiated through collaborative analyses (and replicated by independent groups), genetics research could have a positive impact in elite soccer clubs by improving the level and detail of individual physical programming aimed to overcome and/or develop potential genetic deficiencies and/or strengths, respectively.

References

1. Arnason A, Sigurdsson SB, Gudmundsson A et al. Risk factors for injuries in football. Am J Sports Med 2004; 32: 5S-16S. doi:10.1177/0363546503258912
2. De Moor MH, Spector TD, Cherkas LF et al. Genome-wide linkage scan for athlete status in 700 British female DZ twin pairs. Twin Res Hum Genet 2007; 10: 812-820. doi:10.1375/twin.10.6.812
3. Hakim AJ, Cherkas LF, Spector TD et al. Genetic associations between frozen shoulder and tennis elbow: a female twin study. Rheumatology (Oxford) 2003; 42: 739-742. doi:10.1093/rheumatology/keg159
4. Magnusson K, Turkiewicz A, Hughes V et al. High genetic contribution to anterior cruciate ligament rupture: Heritability ~69. Br J Sports Med 2020. doi:10.1136/bjsports-2020-102392. doi:10.1136/bjsports-2020-102392
5. Bouchard C, An P, Rice T et al. Familial aggregation of VO(2max) response to exercise training: results from the HERITAGE Family Study. J Appl Physiol (1985) 1999; 87: 1003-1008. doi:10.1152/jappl.1999.87.3.1003
6. Peeters MW, Thomis MA, Loos RJ et al. Heritability of somatotype components: a multivariate analysis. Int J Obes (Lond) 2007; 31: 1295-1301. doi:10.1038/sj.ijo.0803575
7. Simoneau JA, Bouchard C. Genetic determinism of fiber type proportion in human skeletal muscle. FASEB J 1995; 9: 1091-1095. doi:10.1096/fasebj.9.11.7649409
8. Zhai G, Ding C, Stankovich J et al. The genetic contribution to longitudinal changes in knee structure and muscle strength: a sibpair study. Arthritis Rheum 2005; 52: 2830-2834. doi:10.1002/art.21267
9. Brownlee TE, Murtagh CF, Naughton RJ et al. Isometric maximal voluntary force evaluated using an isometric mid-thigh pull differentiates English Premier League

youth soccer players from a maturity-matched control group. Sci Med Football 2018; 2: 209-215. doi:10.1080/24733938.2018.1432886

10. Murtagh CF, Brownlee TE, O'Boyle A et al. Importance of Speed and Power in Elite Youth Soccer Depends on Maturation Status. J Strength Cond Res 2018; 32: 297-303. doi:10.1519/jsc.0000000000002367

11. Murtagh CF, Nulty C, Vanrenterghem J et al. The Neuromuscular Determinants of Unilateral Jump Performance in Soccer Players Are Direction-Specific. Int J Sports Physiol Perform 2018; 13: 604-611. doi:10.1123/ijspp.2017-0589

12. Murtagh CF, Stubbs M, Vanrenterghem J et al. Patellar tendon properties distinguish elite from non-elite soccer players and are related to peak horizontal but not vertical power. Eur J Appl Physiol 2018; 118: 1737-1749. doi:10.1007/s00421-018-3905-0

13. Murtagh CF, Vanrenterghem J, O'Boyle A et al. Unilateral jumps in different directions: a novel assessment of soccer-associated power? J Sci Med Sport 2017; 20: 1018-1023. doi:10.1016/j.jsams.2017.03.016

14. Abbott W, Clifford T. The influence of muscle strength and aerobic fitness on functional recovery in professional soccer. J Sports Med Phys Fitness 2021. doi:10.23736/s0022-4707.21.13401-2. doi:10.23736/s0022-4707.21.13401-2

15. Buchheit M, Simpson BM, Hader K et al. Occurrences of near-to-maximal speed-running bouts in elite soccer: insights for training prescription and injury mitigation. Sci Med Footb 2021; 5: 105-110. doi:10.1080/24733938.2020.1802058

16. Dos'Santos T, Cowling I, Challoner M et al. What are the significant turning demands of match play of an English Premier League soccer team? J Sports Sci 2022; 40: 1750-1759. doi:10.1080/02640414.2022.2109355

17. Manzi V, Annino G, Savoia C et al. Relationship between aerobic fitness and metabolic power metrics in elite male soccer players. Biol Sport 2022; 39: 599-606

18. Arnason A, Sigurdsson SB, Gudmundsson A et al. Physical fitness, injuries, and team performance in soccer. Med Sci Sports Exerc 2004; 36: 278-285. doi:10.1249/01.Mss.0000113478.92945.Ca

19. Carling C, Orhant E, LeGall F. Match injuries in professional soccer: inter-seasonal variation and effects of competition type, match congestion and positional role. Int J Sports Med 2010; 31: 271-276. doi:10.1055/s-0029-1243646

20. Ekstrand J, Hägglund M, Waldén M. Injury incidence and injury patterns in professional football: the UEFA injury study. Br J Sports Med 2011; 45: 553-558. doi:10.1136/bjsm.2009.060582

21. Hawkins RD, Hulse MA, Wilkinson C et al. The association football medical research programme: an audit of injuries in professional football. Br J Sports Med 2001; 35: 43-47

22. Hall ECR, Larruskain J, Gil SM et al. An injury audit in high-level male youth soccer players from English, Spanish, Uruguayan and Brazilian academies. Phys Ther Sport 2020; 44: 53-60. doi:10.1016/j.ptsp.2020.04.033

23. Hall ECR, Larruskain J, Gil SM et al. Playing Position and the Injury Incidence Rate in Male Academy Soccer Players. J Athl Train 2022; 57: 696-703. doi:10.4085/1062-6050-0346.21

24. Hall ECR, Larruskain J, Gil SM et al. Injury risk is greater in physically mature versus biologically younger male soccer players from academies in different countries. Phys Ther Sport 2022; 55: 111-118. doi:10.1016/j.ptsp.2022.03.006

25. Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. Nat Rev Genet 2002; 3: 391-397. doi:10.1038/nrg796

26. Seto JT, Quinlan KG, Lek M et al. ACTN3 genotype influences muscle performance through the regulation of calcineurin signaling. J Clin Invest 2013; 123: 4255-4263

27. Erskine RM, Williams AG, Jones DA et al. The individual and combined influence of ACE and ACTN3 genotypes on muscle phenotypes before and after strength training. Scand J Med Sci Sports 2014; 24: 642-648. doi:10.1111/sms.12055

28. Murtagh CF, Brownlee TE, Rienzi E et al. The genetic profile of elite youth soccer players and its association with power and speed depends on maturity status. PLoS One 2020; 15: e0234458. doi:10.1371/journal.pone.0234458

29. Hall ECR, Baumert P, Larruskain J et al. The genetic association with injury risk in male academy soccer players depends on maturity status. Scand J Med Sci Sports 2022; 32: 338-350. doi:10.1111/sms.14077

30. Egorova ES, Borisova AV, Mustafina LJ et al. The polygenic profile of Russian football players. J Sports Sci 2014; 32: 1286-1293. doi:10.1080/02640414.2014.898853

31. Gineviciene V, Jakaitiene A, Tubelis L et al. Variation in the ACE, PPARGC1A and PPARA genes in Lithuanian football players. Eur J Sport Sci 2014; 14 Suppl 1: S289-295. doi:10.1080/17461391.2012.691117

32. Juffer P, Furrer R, González-Freire M et al. Genotype distributions in top-level soccer players: a role for ACE? Int J Sports Med 2009; 30: 387-392. doi:10.1055/s-0028-1105931

33. Micheli ML, Gulisano M, Morucci G et al. Angiotensin-converting enzyme/vitamin D receptor gene polymorphisms and bioelectrical impedance analysis in predicting athletic performances of Italian young soccer players. J Strength Cond Res 2011; 25: 2084-2091. doi:10.1519/JSC.0b013e31820238aa

34. Santiago C, González-Freire M, Serratosa L et al. ACTN3 genotype in professional soccer players. Br J Sports Med 2008; 42: 71-73. doi:10.1136/bjsm.2007.039172

35. Ginevičienė V, Pranckevičienė E, Milašius K et al. Relating fitness phenotypes to genotypes in Lithuanian elite athletes. . Acta Medica Lituanica 2010; 17: 1-10. doi:10.2478/v10140-010-0001-0

36. Pickering C, Suraci B, Semenova EA et al. A Genome-Wide Association Study of Sprint Performance in Elite Youth Football Players. J Strength Cond Res 2019; 33: 2344-2351. doi:10.1519/jsc.0000000000003259

37. Pimenta EM, Coelho DB, Veneroso CE et al. Effect of ACTN3 gene on strength and endurance in soccer players. J Strength Cond Res 2013; 27: 3286-3292. doi:10.1519/JSC.0b013e3182915e66

38. Artells R, Pruna R, Dellal A et al. Elastin: a possible genetic biomarker for more severe ligament injuries in elite soccer. A pilot study. Muscles Ligaments Tendons J 2016; 6: 188-192

39. 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. doi:10.1016/j.jsams.2012.10.004

40. Larruskain J, Celorrio D, Barrio I et al. Genetic Variants and Hamstring Injury in Soccer: An Association and Validation Study. Med Sci Sports Exerc 2018; 50: 361-368. doi:10.1249/mss.0000000000001434

41. Lulińska-Kuklik E, Rahim M, Domańska-Senderowska D et al. Interactions between COL5A1 Gene and Risk of the Anterior Cruciate Ligament Rupture. J Hum Kinet 2018; 62: 65-71

42. Massidda M, Bachis V, Corrias L et al. Influence of the COL5A1 rs12722 on musculoskeletal injuries in professional soccer players. J Sports Med Phys Fitness 2015; 55: 1348-1353

43. Massidda M, Miyamoto-Mikami E, Kumagai H et al. Association between the ACE I/D polymorphism and muscle injuries in Italian and Japanese elite football players. J Sports Sci 2020; 38: 2423-2429. doi:10.1080/02640414.2020.1787683

44. Massidda M, Voisin S, Culigioni C et al. ACTN3 R577X Polymorphism Is Associated With the Incidence and Severity of Injuries in Professional Football Players. Clin J Sport Med 2019; 29: 57-61. doi:10.1097/jsm.0000000000000487

45. Pruna R, Artells R, Lundblad M et al. Genetic biomarkers in non-contact muscle injuries in elite soccer players. Knee Surg Sports Traumatol Arthrosc 2017; 25: 3311-3318. doi:10.1007/s00167-016-4081-6

46. 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

47. Rodas G, Moreno-Pérez V, Del Coso J et al. Alpha-Actinin-3 Deficiency Might Affect Recovery from Non-Contact Muscle Injuries: Preliminary Findings in a Top-Level Soccer Team. Genes (Basel) 2021; 12

48. Beggs AH, Byers TJ, Knoll JH et al. Cloning and characterization of two human skeletal muscle alpha-actinin genes located on chromosomes 1 and 11. J Biol Chem 1992; 267: 9281-9288

49. Ahmetov II, Druzhevskaya AM, Lyubaeva EV et al. The dependence of preferred competitive racing distance on muscle fibre type composition and ACTN3 genotype in speed skaters. Exp Physiol 2011; 96: 1302-1310. doi:10.1113/expphysiol.2011.060293

50. Vincent B, De Bock K, Ramaekers M et al. ACTN3 (R577X) genotype is associated with fiber type distribution. Physiol Genomics 2007; 32: 58-63. doi:10.1152/physiolgenomics.00173.2007

51. Bottinelli R, Canepari M, Pellegrino MA et al. Force-velocity properties of human skeletal muscle fibres: myosin heavy chain isoform and temperature dependence. J Physiol 1996; 495 (Pt 2): 573-586

52. Gilliver SF, Degens H, Rittweger J et al. Variation in the determinants of power of chemically skinned human muscle fibres. Exp Physiol 2009; 94: 1070-1078. doi:10.1113/expphysiol.2009.048314

53. Jones DA, Rutherford OM, Parker DF. Physiological changes in skeletal muscle as a result of strength training. Q J Exp Physiol 1989; 74: 233-256. doi:10.1113/expphysiol.1989.sp003268

54. Ahmetov II, Fedotovskaya ON. Current Progress in Sports Genomics. Adv Clin Chem 2015; 70: 247-314. doi:10.1016/bs.acc.2015.03.003

55. Faude O, Koch T, Meyer T. Straight sprinting is the most frequent action in goal situations in professional football. J Sports Sci 2012; 30: 625-631. doi:10.1080/02640414.2012.665940

56. Murtagh CF, Naughton RJ, McRobert AP et al. A Coding System to Quantify Powerful Actions in Soccer Match Play: A Pilot Study. Res Q Exerc Sport 2019; 90: 234-243. doi:10.1080/02701367.2019.1576838

57. Eynon N, Ruiz JR, Yvert T et al. The C allele in NOS3 -786 T/C polymorphism is associated with elite soccer player's status. Int J Sports Med 2012; 33: 521-524. doi:10.1055/s-0032-1306337

58. Gómez-Gallego F, Ruiz JR, Buxens A et al. The -786 T/C polymorphism of the NOS3 gene is associated with elite performance in power sports. Eur J Appl Physiol 2009; 107: 565-569. doi:10.1007/s00421-009-1166-7

59. Sessa F, Chetta M, Petito A et al. Gene polymorphisms and sport attitude in Italian athletes. Genet Test Mol Biomarkers 2011; 15: 285-290. doi:10.1089/gtmb.2010.0179

60. Hopkinson NS, Li KW, Kehoe A et al. Vitamin D receptor genotypes influence quadriceps strength in chronic obstructive pulmonary disease. Am J Clin Nutr 2008; 87: 385-390. doi:10.1093/ajcn/87.2.385

61. Dellal A, Chamari K, Wong dP et al. Comparison of physical and technical performance in European soccer match-play: FA Premier League and La Liga. . Eur J Sport Sci 2011; 11: 51-59

62. Bloomfield J, Polman R, Butterly R et al. Analysis of age, stature, body mass, BMI and quality of elite soccer players from 4 European Leagues. J Sports Med Phys Fitness 2005; 45: 58-67

63. Gomez-Gallego F, Santiago C, González-Freire M et al. The C allele of the AGT Met235Thr polymorphism is associated with power sports performance. Appl Physiol Nutr Metab 2009; 34: 1108-1111. doi:10.1139/h09-108

64. Ruiz JR, Arteta D, Buxens A et al. Can we identify a power-oriented polygenic profile? J Appl Physiol (1985) 2010; 108: 561-566. doi:10.1152/japplphysiol.01242.2009

65. Zarębska A, Sawczyn S, Kaczmarczyk M et al. Association of rs699 (M235T) polymorphism in the AGT gene with power but not endurance athlete status. J Strength Cond Res 2013; 27: 2898-2903. doi:10.1519/JSC.0b013e31828155b5

965 66. Petr M, Maciejewska-Skrendo A, Zajac A et al. Association of Elite Sports Status
966 with Gene Variants of Peroxisome Proliferator Activated Receptors and Their
967 Transcriptional Coactivator. Int J Mol Sci 2019; 21
968 67. Petr M, Stastny P, Pecha O et al. PPARA intron polymorphism associated with power
969 performance in 30-s anaerobic Wingate Test. PLoS One 2014; 9: e107171
970 68. Sawczuk M, Maciejewska-Karlowska A, Cieszczyk P et al. Association of the
971 ADRB2 Gly16Arg and Glu27Gln polymorphisms with athlete status. J Sports Sci
972 2013; 31: 1535-1544. doi:10.1080/02640414.2013.786184
973 69. Massidda M, Mendez-Villanueva A, Ginevičienė V et al. Association of
974 Monocarboxylate Transporter-1 (MCT1) A1470T Polymorphism (rs1049434) with
975 Forward Football Player Status. Int J Sports Med 2018; 39: 1028-1034.
976 doi:10.1055/a-0634-6387
977 70. Merezhinskaya N, Fishbein WN, Davis JI et al. Mutations in MCT1 cDNA in patients
978 with symptomatic deficiency in lactate transport. Muscle Nerve 2000; 23: 90-97.
979 doi:10.1002/(sici)1097-4598(200001)23:1<90::aid-mus12>3.0.co;2-m
980 71. Cupeiro R, Benito PJ, Maffulli N et al. MCT1 genetic polymorphism influence in
981 high intensity circuit training: a pilot study. J Sci Med Sport 2010; 13: 526-530.
982 doi:10.1016/j.jsams.2009.07.004
983 72. Cupeiro R, Pérez-Prieto R, Amigo T et al. Role of the monocarboxylate transporter
984 MCT1 in the uptake of lactate during active recovery. Eur J Appl Physiol 2016; 116:
985 1005-1010. doi:10.1007/s00421-016-3365-3
986 73. Aziz AR, Mukherjee S, Chia MY et al. Validity of the running repeated sprint ability
987 test among playing positions and level of competitiveness in trained soccer players.
988 Int J Sports Med 2008; 29: 833-838. doi:10.1055/s-2008-1038410

74. Di Salvo V, Baron R, Tschan H et al. Performance characteristics according to playing position in elite soccer. Int J Sports Med 2007; 28: 222-227. doi:10.1055/s-2006-924294

75. Carling C, Bloomfield J, Nelsen L et al. The role of motion analysis in elite soccer: contemporary performance measurement techniques and work rate data. Sports Med 2008; 38: 839-862. doi:10.2165/00007256-200838100-00004

76. Clos E, Pruna R, Lundblad M et al. ACTN3's R577X Single Nucleotide Polymorphism Allele Distribution Differs Significantly in Professional Football Players according to Their Field Position. Med Princ Pract 2021; 30: 92-97

77. Petr M, Thiel D, Kateřina K et al. Speed and power-related gene polymorphisms associated with playing position in elite soccer players. Biol Sport 2022; 39: 355-366

78. Viru A, Loko J, Harro M et al. Critical periods in the development of performance capacity during childhood and adolescence. Eur J Phys Edu 1999; 4: 75–119

79. Devismes M, Aeles J, Philips J et al. Sprint force-velocity profiles in soccer players: impact of sex and playing level. Sports Biomech 2021; 20: 947-957. doi:10.1080/14763141.2019.1618900

80. Little T, Williams AG. Specificity of acceleration, maximum speed, and agility in professional soccer players. J Strength Cond Res 2005; 19: 76-78. doi:10.1519/14253.1

81. Mero A. Force-Time Characteristics and Running Velocity of Male Sprinters during the Acceleration Phase of Sprinting. Res Q Exerc Sport 1988; 59: 94-98. doi:10.1080/02701367.1988.10605484

82. Waldron M, Murphy A. A comparison of physical abilities and match performance characteristics among elite and subelite under-14 soccer players. Pediatr Exerc Sci 2013; 25: 423-434. doi:10.1123/pes.25.3.423

83. Venturelli M, Bishop D, Pettene L. Sprint training in preadolescent soccer players. Int J Sports Physiol Perform 2008; 3: 558-562. doi:10.1123/ijspp.3.4.558

84. Nagano A, Komura T, Fukashiro S. Optimal coordination of maximal-effort horizontal and vertical jump motions--a computer simulation study. Biomed Eng Online 2007; 6: 20

85. Fukashiro S, Besier TF, Barrett R et al. Direction control in standing horizontal and vertical jumps. . Int J Sport Health Sci 2005; 3: 272–279

86. Ahmetov II, Donnikov AE, Trofimov DY. Actn3 genotype is associated with testosterone levels of athletes. Biol Sport 2014; 31: 105-108. doi:10.5604/20831862.1096046

87. Dionísio TJ, Thiengo CR, Brozoski DT et al. The influence of genetic polymorphisms on performance and cardiac and hemodynamic parameters among Brazilian soccer players. Appl Physiol Nutr Metab 2017; 42: 596-604. doi:10.1139/apnm-2016-0608

88. Ahmetov, II, Williams AG, Popov DV et al. The combined impact of metabolic gene polymorphisms on elite endurance athlete status and related phenotypes. Hum Genet 2009; 126: 751-761. doi:10.1007/s00439-009-0728-4

89. Massidda M, Scorcu M, Calò CM. New genetic model for predicting phenotype traits in sports. Int J Sports Physiol Perform 2014; 9: 554-560. doi:10.1123/ijspp.2012-0339

90. Coelho DB, Pimenta E, Rosse IC et al. The alpha-actinin-3 r577x polymorphism and physical performance in soccer players. J Sports Med Phys Fitness 2016; 56: 241-248

91. Coelho DB, Pimenta EM, Rosse IC et al. Polymorphism of the angiotensin converting enzyme gene (ACE-I/D) differentiates the aerobic and speed performance of football players. J Sports Med Phys Fitness 2022; 62: 192-198. doi:10.23736/s0022-4707.21.12060-2

92. Kubo K, Kawakami Y, Fukunaga T. Influence of elastic properties of tendon structures on jump performance in humans. J Appl Physiol (1985) 1999; 87: 2090-2096. doi:10.1152/jappl.1999.87.6.2090

93. Kubo K, Yata H, Tsunoda N. Effect of gene polymorphisms on the mechanical properties of human tendon structures. Springerplus 2013; 2: 343

94. Deprez DN, Fransen J, Lenoir M et al. A retrospective study on anthropometrical, physical fitness, and motor coordination characteristics that influence dropout, contract status, and first-team playing time in high-level soccer players aged eight to eighteen years. J Strength Cond Res 2015; 29: 1692-1704. doi:10.1519/jsc.0000000000000806

95. Zigova T, Pencea V, Wiegand SJ et al. Intraventricular administration of BDNF increases the number of newly generated neurons in the adult olfactory bulb. Mol Cell Neurosci 1998; 11: 234-245. doi:10.1006/mcne.1998.0684

96. McAllister AK, Katz LC, Lo DC. Neurotrophins and synaptic plasticity. Annu Rev Neurosci 1999; 22: 295-318. doi:10.1146/annurev.neuro.22.1.295

97. Egan MF, Kojima M, Callicott JH et al. The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. Cell 2003; 112: 257-269. doi:10.1016/s0092-8674(03)00035-7

98. Charlier R, Caspers M, Knaeps S et al. Limited potential of genetic predisposition scores to predict muscle mass and strength performance in Flemish Caucasians between 19 and 73 years of age. Physiol Genomics 2017; 49: 160-166. doi:10.1152/physiolgenomics.00085.2016

99. Dong B, Li Q, Zhang T et al. Population Genetic Polymorphism of Skeletal Muscle Strength Related Genes in Five Ethnic Minorities in North China. Front Genet 2021; 12: 756802

100. He L, Khanal P, Morse CI et al. Associations of combined genetic and epigenetic scores with muscle size and muscle strength: a pilot study in older women. J Cachexia Sarcopenia Muscle 2020; 11: 1548-1561

101. He L, Van Roie E, Bogaerts A et al. Genetic predisposition score predicts the increases of knee strength and muscle mass after one-year exercise in healthy elderly. Exp Gerontol 2018; 111: 17-26. doi:10.1016/j.exger.2018.06.030

102. Maciejewska-Skrendo A, Leźnicka K, Leońska-Duniec A et al. Genetics of Muscle Stiffness, Muscle Elasticity and Explosive Strength. J Hum Kinet 2020; 74: 143-159

103. Brownlee TE. An investigation into the capabilities and affecting factors of isometric mid-thigh pull force production in elite youth soccer players. [Doctoral Thesis]: Liverpool John Moores University; 2016. doi:10.24377/LJMU.t.00006072

104. Baumert P, Stewart CE, Lake MJ et al. Variations of collagen-encoding genes are associated with exercise-induced muscle damage. Physiol Genomics 2018; 50: 691-693. doi:10.1152/physiolgenomics.00145.2017

105. Erskine RM, Williams AG, Jones DA et al. Do PTK2 gene polymorphisms contribute to the interindividual variability in muscle strength and the response to resistance training? A preliminary report. J Appl Physiol (1985) 2012; 112: 1329-1334. doi:10.1152/japplphysiol.01137.2011

106. Massidda M, Eynon N, Bachis V et al. Association Between MCT1 A1470T Polymorphism and Fat-Free Mass in Well-Trained Young Soccer Players. J Strength Cond Res 2016; 30: 1171-1176. doi:10.1519/jsc.0000000000001176

107. Erskine RM, Fletcher G, Folland JP. The contribution of muscle hypertrophy to strength changes following resistance training. Eur J Appl Physiol 2014; 114: 1239-1249. doi:10.1007/s00421-014-2855-4

108. Nuell S, Illera-Domínguez V, Carmona G et al. Hamstring Muscle Volume as an Indicator of Sprint Performance. J Strength Cond Res 2021; 35: 902-909. doi:10.1519/jsc.0000000000003976

109. Boraita A, de la Rosa A, Heras ME et al. Cardiovascular adaptation, functional capacity and Angiotensin-converting enzyme I/D polymorphism in elite athletes. Rev Esp Cardiol 2010; 63: 810-819. doi:10.1016/s1885-5857(10)70166-3

110. Papadimitriou ID, Papadopoulos C, Kouvatsi A et al. The ACTN3 gene in elite Greek track and field athletes. Int J Sports Med 2008; 29: 352-355. doi:10.1055/s-2007-965339

111. Pescatello LS, Kostek MA, Gordish-Dressman H et al. ACE ID genotype and the muscle strength and size response to unilateral resistance training. Med Sci Sports Exerc 2006; 38: 1074-1081. doi:10.1249/01.mss.0000222835.28273.80

112. Delecluse C. Influence of strength training on sprint running performance. Current findings and implications for training. Sports Med 1997; 24: 147-156. doi:10.2165/00007256-199724030-00001

113. Dubouchaud H, Butterfield GE, Wolfel EE et al. Endurance training, expression, and physiology of LDH, MCT1, and MCT4 in human skeletal muscle. Am J Physiol Endocrinol Metab 2000; 278: E571-579. doi:10.1152/ajpendo.2000.278.4.E571

114. Stafilidis S, Arampatzis A. Muscle - tendon unit mechanical and morphological properties and sprint performance. J Sports Sci 2007; 25: 1035-1046. doi:10.1080/02640410600951589

115. Morin JB, Gimenez P, Edouard P et al. Sprint Acceleration Mechanics: The Major Role of Hamstrings in Horizontal Force Production. Front Physiol 2015; 6: 404

116. Ding X, Jin Y, Wu Y et al. Localization and cellular distribution of CPNE5 in embryonic mouse brain. Brain Res 2008; 1224: 20-28. doi:10.1016/j.brainres.2008.05.051

117. Reiner A, Yang M, Cagle MC et al. Localization of cerebellin-2 in late embryonic chicken brain: implications for a role in synapse formation and for brain evolution. J Comp Neurol 2011; 519: 2225-2251

118. Fernandez T, Morgan T, Davis N et al. Disruption of contactin 4 (CNTN4) results in developmental delay and other features of 3p deletion syndrome. Am J Hum Genet 2004; 74: 1286-1293

119. Stølen T, Chamari K, Castagna C et al. Physiology of soccer: an update. Sports Med 2005; 35: 501-536. doi:10.2165/00007256-200535060-00004

120. Bueno S, Pasqua LA, de Araújo G et al. The Association of ACE Genotypes on Cardiorespiratory Variables Related to Physical Fitness in Healthy Men. PLoS One 2016; 11: e0165310

121. Konopka MJ, van den Bunder J, Rietjens G et al. Genetics of long-distance runners and road cyclists-A systematic review with meta-analysis. Scand J Med Sci Sports 2022; 32: 1414-1429

122. Papadimitriou ID, Lockey SJ, Voisin S et al. No association between ACTN3 R577X and ACE I/D polymorphisms and endurance running times in 698 Caucasian athletes. BMC Genomics 2018; 19: 13. doi:10.1186/s12864-017-4412-0

123. Mirwald RL, Baxter-Jones AD, Bailey DA et al. An assessment of maturity from anthropometric measurements. Med Sci Sports Exerc 2002; 34: 689-694. doi:10.1097/00005768-200204000-00020

124. Casamichana D, Castellano J, Castagna C. Comparing the physical demands of friendly matches and small-sided games in semiprofessional soccer players. J Strength Cond Res 2012; 26: 837-843. doi:10.1519/JSC.0b013e31822a61cf

125. Pickering C, Kiely J, Suraci B et al. The magnitude of Yo-Yo test improvements following an aerobic training intervention are associated with total genotype score. PLoS One 2018; 13: e0207597

126. Parry L, Drust B. Is injury the major cause of elite soccer players being unavailable to train and play during the competitive season? Phys Ther Sport 2006; 7: 58-64

127. Eirale C, Tol JL, Farooq A et al. Low injury rate strongly correlates with team success in Qatari professional football. Br J Sports Med 2013; 47: 807-808

128. Carling C, Le Gall F, McCall A et al. Squad management, injury and match performance in a professional soccer team over a championship-winning season. Eur J Sport Sci 2015; 15: 573-582. doi:10.1080/17461391.2014.955885

129. Larruskain J, Lekue JA, Diaz N et al. A comparison of injuries in elite male and female football players: A five-season prospective study. Scand J Med Sci Sports 2018; 28: 237-245. doi:10.1111/sms.12860

130. Varley I, Patel S, Williams AG et al. The current use, and opinions of elite athletes and support staff in relation to genetic testing in elite sport within the UK. Biol Sport 2018; 35: 13-19

131. Tucker AM. Common soccer injuries. Diagnosis, treatment and rehabilitation. Sports Med 1997; 23: 21-32. doi:10.2165/00007256-199723010-00003

132. Laguette MJ, Abrahams Y, Prince S et al. Sequence variants within the 3'-UTR of the COL5A1 gene alters mRNA stability: implications for musculoskeletal soft tissue injuries. Matrix Biol 2011; 30: 338-345. doi:10.1016/j.matbio.2011.05.001

133. 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. doi:10.1136/bjsm.2008.056150

134. Posthumus M, September AV, O'Cuinneagain D et al. The COL5A1 gene is associated with increased risk of anterior cruciate ligament ruptures in female participants. Am J Sports Med 2009; 37: 2234-2240. doi:10.1177/0363546509338266

135. September AV, Cook J, Handley CJ et al. Variants within the COL5A1 gene are associated with Achilles tendinopathy in two populations. Br J Sports Med 2009; 43: 357-365. doi:10.1136/bjsm.2008.048793

136. Ahmetov II, Hall ECR, Semenova EA et al. Advances in sports genomics. Adv Clin Chem 2022; 107: 215-263. doi:10.1016/bs.acc.2021.07.004

137. Broos S, Malisoux L, Theisen D et al. Evidence for ACTN3 as a Speed Gene in Isolated Human Muscle Fibers. PLoS One 2016; 11: e0150594

138. Timmins RG, Bourne MN, Shield AJ et al. Short biceps femoris fascicles and eccentric knee flexor weakness increase the risk of hamstring injury in elite football (soccer): a prospective cohort study. Br J Sports Med 2016; 50: 1524-1535. doi:10.1136/bjsports-2015-095362

139. Pimenta EM, Coelho DB, Cruz IR et al. The ACTN3 genotype in soccer players in response to acute eccentric training. Eur J Appl Physiol 2012; 112: 1495-1503. doi:10.1007/s00421-011-2109-7

140. Clos E, Pruna R, Lundblad M et al. ACTN3 single nucleotide polymorphism is associated with non-contact musculoskeletal soft-tissue injury incidence in elite professional football players. Knee Surg Sports Traumatol Arthrosc 2019; 27: 4055-4061. doi:10.1007/s00167-019-05381-x

141. Bradley PS, Dellal A, Mohr M et al. Gender differences in match performance characteristics of soccer players competing in the UEFA Champions League. Hum Mov Sci 2014; 33: 159-171. doi:10.1016/j.humov.2013.07.024

142. Waldén M, Hägglund M, Werner J et al. The epidemiology of anterior cruciate ligament injury in football (soccer): a review of the literature from a gender-related perspective. Knee Surg Sports Traumatol Arthrosc 2011; 19: 3-10. doi:10.1007/s00167-010-1172-7

143. Del Coso J, Rodas G, Buil M et al. Association of the ACTN3 rs1815739 Polymorphism with Physical Performance and Injury Incidence in Professional Women Football Players. Genes (Basel) 2022; 13

144. Baumert P, Lake MJ, Stewart CE et al. Genetic variation and exercise-induced muscle damage: implications for athletic performance, injury and ageing. Eur J Appl Physiol 2016; 116: 1595-1625

145. 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

146. Boot-Handford RP, Tuckwell DS. Fibrillar collagen: the key to vertebrate evolution? A tale of molecular incest. Bioessays 2003; 25: 142-151. doi:10.1002/bies.10230

147. Posthumus M, September AV, Schwellnus MP et al. 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. doi:10.1016/j.jsams.2007.12.006

148. Jin H, van't Hof RJ, Albagha OM et al. Promoter and intron 1 polymorphisms of COL1A1 interact to regulate transcription and susceptibility to osteoporosis. Hum Mol Genet 2009; 18: 2729-2738. doi:10.1093/hmg/ddp205

149. Collins M, Posthumus M. Type V collagen genotype and exercise-related phenotype relationships: a novel hypothesis. Exerc Sport Sci Rev 2011; 39: 191-198. doi:10.1097/JES.0b013e318224e853

150. Wenstrup RJ, Florer JB, Brunskill EW et al. Type V collagen controls the initiation of collagen fibril assembly. J Biol Chem 2004; 279: 53331-53337. doi:10.1074/jbc.M409622200

151. Zacchigna L, Vecchione C, Notte A et al. Emilin1 links TGF-beta maturation to blood pressure homeostasis. Cell 2006; 124: 929-942. doi:10.1016/j.cell.2005.12.035

152. Randell A, Daneshtalab N. Elastin microfibril interface-located protein 1, transforming growth factor beta, and implications on cardiovascular complications. J Am Soc Hypertens 2017; 11: 437-448. doi:10.1016/j.jash.2017.04.010

153. Kannus P. Structure of the tendon connective tissue. Scand J Med Sci Sports 2000; 10: 312-320. doi:10.1034/j.1600-0838.2000.010006312.x

154. Posthumus M, Collins M, van der Merwe L et al. Matrix metalloproteinase genes on chromosome 11q22 and the risk of anterior cruciate ligament (ACL) rupture. Scand J Med Sci Sports 2012; 22: 523-533. doi:10.1111/j.1600-0838.2010.01270.x

155. Birkedal-Hansen H, Moore WG, Bodden MK et al. Matrix metalloproteinases: a review. Crit Rev Oral Biol Med 1993; 4: 197-250. doi:10.1177/10454411930040020401

156. Chen X, Li Y. Role of matrix metalloproteinases in skeletal muscle: migration, differentiation, regeneration and fibrosis. Cell Adh Migr 2009; 3: 337-341

157. Medley TL, Kingwell BA, Gatzka CD et al. Matrix metalloproteinase-3 genotype contributes to age-related aortic stiffening through modulation of gene and protein expression. Circ Res 2003; 92: 1254-1261. doi:10.1161/01.Res.0000076891.24317.Ca

158. Guo YQ, Zheng LN, Wei JF et al. Expression of CCL2 and CCR2 in the hippocampus and the interventional roles of propofol in rat cerebral ischemia/reperfusion. Exp Ther Med 2014; 8: 657-661

159. Hubal MJ, Chen TC, Thompson PD et al. Inflammatory gene changes associated with the repeated-bout effect. Am J Physiol Regul Integr Comp Physiol 2008; 294: R1628-1637. doi:10.1152/ajpregu.00853.2007

160. Catoire M, Kersten S. The search for exercise factors in humans. FASEB J 2015; 29: 1615-1628. doi:10.1096/fj.14-263699

161. Hubal MJ, Devaney JM, Hoffman EP et al. CCL2 and CCR2 polymorphisms are associated with markers of exercise-induced skeletal muscle damage. J Appl Physiol (1985) 2010; 108: 1651-1658. doi:10.1152/japplphysiol.00361.2009

162. Yahiaoui L, Gvozdic D, Danialou G et al. CC family chemokines directly regulate myoblast responses to skeletal muscle injury. J Physiol 2008; 586: 3991-4004

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

164. Andersen MB, Pingel J, Kjær M et al. Interleukin-6: a growth factor stimulating collagen synthesis in human tendon. J Appl Physiol (1985) 2011; 110: 1549-1554. doi:10.1152/japplphysiol.00037.2010

165. Muñoz-Cánoves P, Scheele C, Pedersen BK et al. Interleukin-6 myokine signaling in skeletal muscle: a double-edged sword? FEBS J 2013; 280: 4131-4148

166. Pedersen BK, Febbraio MA. Muscle as an endocrine organ: focus on muscle-derived interleukin-6. Physiol Rev 2008; 88: 1379-1406. doi:10.1152/physrev.90100.2007

167. Nieman DC, Nehlsen-Cannarella SL, Fagoaga OR et al. Influence of mode and carbohydrate on the cytokine response to heavy exertion. Med Sci Sports Exerc 1998; 30: 671-678. doi:10.1097/00005768-199805000-00005

168. Millar NL, Wei AQ, Molloy TJ et al. Cytokines and apoptosis in supraspinatus tendinopathy. J Bone Joint Surg Br 2009; 91: 417-424. doi:10.1302/0301-620x.91b3.21652

169. Yamin C, Duarte JA, Oliveira JM et al. IL6 (-174) and TNFA (-308) promoter polymorphisms are associated with systemic creatine kinase response to eccentric exercise. Eur J Appl Physiol 2008; 104: 579-586. doi:10.1007/s00421-008-0728-4

170. Massidda M, Eynon N, Bachis V et al. Influence of the MCT1 rs1049434 on Indirect Muscle Disorders/Injuries in Elite Football Players. Sports Med Open 2015; 1: 33

171. Massidda M, Corrias L, Bachis V et al. Vitamin D receptor gene polymorphisms and musculoskeletal injuries in professional football players. Exp Ther Med 2015; 9: 1974-1978

172. Bouchard C. Overcoming barriers to progress in exercise genomics. Exerc Sport Sci Rev 2011; 39: 212-217

173. Mills M, Yang N, Weinberger R et al. Differential expression of the actin-binding proteins, alpha-actinin-2 and -3, in different species: implications for the evolution of functional redundancy. Hum Mol Genet 2001; 10: 1335-1346. doi:10.1093/hmg/10.13.1335

174. Reneland R, Lithell H. Angiotensin-converting enzyme in human skeletal muscle. A simple in vitro assay of activity in needle biopsy specimens. Scand J Clin Lab Invest 1994; 54: 105-111. doi:10.3109/00365519409086516

175. Gordon SE, Davis BS, Carlson CJ et al. ANG II is required for optimal overload-induced skeletal muscle hypertrophy. Am J Physiol Endocrinol Metab 2001; 280: E150-159. doi:10.1152/ajpendo.2001.280.1.E150

176. Rigat B, Hubert C, Alhenc-Gelas F et al. An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. J Clin Invest 1990; 86: 1343-1346

177. Charbonneau DE, Hanson ED, Ludlow AT et al. ACE genotype and the muscle hypertrophic and strength responses to strength training. Med Sci Sports Exerc 2008; 40: 677-683

178. North KN, Beggs AH. Deficiency of a skeletal muscle isoform of alpha-actinin (alpha-actinin-3) in merosin-positive congenital muscular dystrophy. Neuromuscul Disord 1996; 6: 229-235. doi:10.1016/0960-8966(96)00361-6

179. Sethi AA, Nordestgaard BG, Tybjaerg-Hansen A. Angiotensinogen gene polymorphism, plasma angiotensinogen, and risk of hypertension and ischemic heart disease: a meta-analysis. Arterioscler Thromb Vasc Biol 2003; 23: 1269-1275. doi:10.1161/01.Atv.0000079007.40884.5c

180. Morisaki T, Gross M, Morisaki H et al. Molecular basis of AMP deaminase deficiency in skeletal muscle. Proc Natl Acad Sci U S A 1992; 89: 6457-6461

181. Kuntz LA, Rossetti L, Kunold E et al. Biomarkers for tissue engineering of the tendon-bone interface. PLoS One 2018; 13: e0189668

182. Birk DE, Fitch JM, Babiarz JP et al. Collagen fibrillogenesis in vitro: interaction of types I and V collagen regulates fibril diameter. J Cell Sci 1990; 95 (Pt 4): 649-657. doi:10.1242/jcs.95.4.649

183. Halestrap AP, Meredith D. The SLC16 gene family-from monocarboxylate transporters (MCTs) to aromatic amino acid transporters and beyond. Pflugers Arch 2004; 447: 619-628. doi:10.1007/s00424-003-1067-2

184. Massidda M, Flore L, Kikuchi N et al. Influence of the MCT1-T1470A polymorphism (rs1049434) on repeated sprint ability and blood lactate accumulation in elite football players: a pilot study. Eur J Appl Physiol 2021; 121: 3399-3408. doi:10.1007/s00421-021-04797-z

185. Nakayama M, Yasue H, Yoshimura M et al. T(-786)--> C mutation in the 5'-flanking region of the endothelial nitric oxide synthase gene is associated with myocardial infarction, especially without coronary organic stenosis. Am J Cardiol 2000; 86: 628-634. doi:10.1016/s0002-9149(00)01041-9

186. Kersten S. Integrated physiology and systems biology of PPARα. Mol Metab 2014; 3: 354-371

187. Ahmetov, II, Mozhayskaya IA, Flavell DM et al. PPARalpha gene variation and physical performance in Russian athletes. Eur J Appl Physiol 2006; 97: 103-108. doi:10.1007/s00421-006-0154-4

188. Auton A, Brooks LD, Durbin RM et al. A global reference for human genetic variation. Nature 2015; 526: 68-74

189. McKee TJ, Perlman G, Morris M et al. Extracellular matrix composition of connective tissues: a systematic review and meta-analysis. Sci Rep 2019; 9: 10542

190. Oh VM, Chua BM, Heng CK et al. Association of intronic single-nucleotide polymorphisms in the EMILIN1 gene with essential hypertension in a Chinese population. J Hum Hypertens 2012; 26: 553-561. doi:10.1038/jhh.2011.68

191. Nie G, Wen X, Liang X et al. Additional evidence supports association of common genetic variants in MMP3 and TIMP2 with increased risk of chronic Achilles

tendinopathy susceptibility. J Sci Med Sport 2019; 22: 1074-1078. doi:10.1016/j.jsams.2019.05.021

Figure legend

Figure 1. Association between the *COL2A1* rs2070739 SNP and isometric mid-thigh pull maximum voluntary force (MVF) normalised to body mass (BM) in 148 high-level male academy soccer players (ASP) and 93 age- and sex-matched non-athletes (NA). * Main effect for genotype between CC (black bars) and CT/TT (chequered bars) genotypes ($P = 0.033$), ^ main effect for cohort and maturity status ($P < 0.001$).

Table headings

Table 1. Key polymorphisms associated with athlete status and physical performance phenotypes in soccer players. Alleles for each polymorphism are reported according to the forward DNA strand.

Table 2. Key polymorphisms associated with injury risk in soccer players. Alleles for each polymorphism are reported according to the forward DNA strand.

Table 1. Key polymorphisms associated with athlete status and physical performance phenotypes in soccer players. Alleles for each polymorphism are reported according to the forward DNA strand.

Gene	Encoded protein	Protein function	Alleles and rs number	Polymorphism function	Associations with phenotypes of interest in soccer
ACE	Angiotensin-I converting enzyme	ACE converts angiotensin-I (Ang I) to Ang II and is expressed in SkM [174], where Ang II modulates SkM hypertrophy after mechanical loading [175].	Insertion (I)/ Deletion (D) of 287 base pairs rs1799752 / rs4340 / rs13447447 / rs4646994	D-allele carriers express higher ACE activity than II homozygotes [176].	***Phenotype: Athlete status*** • D allele more frequent in pro players vs. NA [30, 32]. • Possibly due to D allele influencing circulating ACE and Ang II in SkM, potentially leading to larger [177], more powerful muscles. ***Phenotype: Sprint Performance*** • DD genotype associated with faster 20 m sprint speed compared to II genotype but no difference in vertical CMJ performance in 212 high-level pro players [91].
ACTN3	α-actinin-3	Binds actin to the Z-line in type II SkM fibres [48]; blocks calcineurin, inhibiting fast-to-slow myofibre transition [26].	C(R)>T(X) rs1815739	X allele leads to a stop codon at amino acid 577, preventing protein production [178]. Consequently, the R allele is associated with greater type II SkM fibre composition [49].	***Phenotype: Athlete status*** • X allele (generally associated with endurance phenotypes) more frequent in pre-PHV academy players *vs.* non-athletes (NA) [28]; • R allele (generally associated with strength/power phenotypes) more frequent in pro players *vs.* NA [30, 34]; • Above discrepancy probably due to variability in match demands between pre- (endurance-based) and post-PHV/adult (power/speed-based) players [28]; • Combination of 'favourable' alleles from *PPARA* rs4253778 (C), *NOS3* rs2070744 (T), *ACTN3* rs1815739 (R), and *AGT* rs699 (G) SNPs correlated with agility performance in pre-PHV academy players [28]. ***Phenotype: CMJ (Vertical Power) Performance*** • R allele associated with greater bilateral vertical CMJ height in pro players [37]. ***Phenotype: Aerobic Capacity*** XX genotype associated with greater YoYo score in pro players [37].
AGT	Angiotensinogen	Precursor of Ang II.	A>G rs699	G allele associated with higher blood AGT concentration and higher	***Phenotype: Athlete status*** • D allele more frequent in pro players vs. NA [30, 32]. • G allele more frequent in academy players *vs.* NA [28];

				blood pressure [179].	• Combination of 'favourable' alleles from *PPARA* rs4253778 (C), *NOS3* rs2070744 (T), *ACTN3* rs1815739 (R), and *AGT* rs699 (G) SNPs correlated with agility performance in pre-PHV academy players [28].
AMPD1	Adenosine Monophosphate Deaminase 1	Enzyme catalysing the deamination of AMP to IMP in SkM	G>A rs17602729	A allele results in a premature stop codon and non-functional enzyme, causing impaired AMP metabolism, which can lead to muscle fatigue, weakness and cramping [180].	***Phenotype: Horizontal Acceleration/ Sprint/ Agility Performance*** • GG genotype associated with faster 10 m acceleration and 20 m sprint times than pre-PHV A-allele carriers in pre-PHV academy players [28]. ***Phenotype: Agility (Change of Direction) Performance*** • Combination of 'favourable' alleles from *AMPD1* rs17602729 (G), *BDNF* rs6265 (C), *COL5A1* rs12722 (C), *COL2A1* rs2070739 (C) SNPs correlated with agility performance in pre-PHV academy players [28].
BDNF	Brain-derived neurotrophic factor	Promotes neurone growth, differentiation, and maintenance [95].	C>T rs6265	C allele results in valine rather than methionine (T allele) at amino acid 66, leading to greater abundance of exercise-induced serum BDNF concentration [97].	***Phenotype: Horizontal Acceleration/ Sprint/ Agility Performance*** • CC genotype associated with greater 10 m acceleration and 20 m sprint performance in academy players [28], possibly linked to C allele's relation with higher exercise-induced serum BDNF concentration [97], potentially enhancing neuromuscular activation/control during acceleration/sprint performance. ***Phenotype: Agility (Change of Direction) Performance*** • Combination of 'favourable' alleles from *AMPD1* rs17602729 (G), *BDNF* rs6265 (C), *COL5A1* rs12722 (C), *COL2A1* rs2070739 (C) SNPs correlated with agility performance in pre-PHV academy players [28].
COL2A1	Pro-α1 (II) chain	Major component of type II collagen, providing structure and strength to connective tissue, e.g. at the enthesis [181].	C>T rs2070739	Not yet known but 3'-UTR SNPs have the potential to alter the level, location, or timing of gene expression [25].	***Phenotype: Horizontal Acceleration/ Sprint/ Agility Performance*** • CC genotype associated with faster 10 m acceleration times than T allele carriers in academy players [28]. ***Phenotype: Agility (Change of Direction) Performance*** • Combination of 'favourable' alleles from *AMPD1* rs17602729 (G), *BDNF* rs6265 (C), *COL5A1* rs12722 (C), *COL2A1* rs2070739 (C) SNPs correlated with agility performance in pre-PHV academy players [28].
COL5A1	Pro-α1 (V) chain	Major component of type V collagen, which regulates the diameter of collagen fibrils [182].	C>T rs12722	Not yet known but 3'-UTR SNPs have the potential to alter the level, location, or timing of gene expression [25].	***Phenotype: CMJ (Vertical Power) Performance*** • CC genotype associated with greater bilateral vertical CMJ height in academy players [28], possibly linked to more extensible knee extensor tendon in CC homozygotes [93], enabling greater storage and release of energy, thus amplifying power during vertical CMJ.

					Phenotype: Horizontal Acceleration/ Sprint/ Agility Performance • CC genotype associated with quicker acceleration, sprint and horizontal CMJ performance in academy players [28], possibly linked to CC genotype associated with more extensible knee extensor tendon [93], enabling greater storage and release of energy, thus amplifying power during acceleration, sprint and horizontal CMJ. ***Phenotype: Agility (Change of Direction) Performance*** • Combination of 'favourable' alleles from *AMPD1* rs17602729 (G), *BDNF* rs6265 (C), *COL5A1* rs12722 (C), *COL2A1* rs2070739 (C) SNPs correlated with agility performance in pre-PHV academy players [28].
MCT1	Monocarboxylate transporter-1	Transports lactate and H^+ across SkM sarcolemma, found predominantly in oxidative fibres [183].	A>T rs1049434	Not yet known but intronic SNPs have the potential to influence gene expression and mRNA stability [25]. T allele results in aspartic acid replacing glutamic acid at position 490, potentially causing a defect in MCT1's ability to transport lactate.	***Phenotype: Athlete status*** • A allele and AA genotype more frequent in 148 pro forwards than 781 NA [69]. ***Phenotype: Horizontal Acceleration/ Sprint/ Agility Performance*** • A-allele associated with faster time in final two sprints of RSA assessment compared to TT homozygotes in 26 academy players [184]. • The above associations may be linked to increased blood lactate accumulation after high-intensity exercise in T-allele carriers vs. AA homozygotes [71].
NOS3	Nitric oxide synthase-3	Potentially stimulating muscle hypertrophy through NO-mediated vasodilatation [58].	C>T rs2070744	T allele increases gene promoter activity, thus increasing eNOS and NO synthesis [185].	***Phenotype: Athlete status*** • T allele greater in academy players vs. NA [28]. • Combination of 'favourable' alleles from *PPARA* rs4253778 (C), *NOS3* rs2070744 (T), *ACTN3* rs1815739 (R), and *AGT* rs699 (G) SNPs correlated with agility performance in pre-PHV academy players [28]. ***Phenotype: Horizontal Acceleration/ Sprint/ Agility Performance*** • TT genotype associated with greater acceleration and sprint performance in academy players [28], possibly linked to greater number of sarcomeres in series, thus increasing SkM contraction velocity [53].
PPARA	Peroxisome proliferator-activated receptor-α	Present in SkM and promotes uptake, utilization, and catabolism of fatty	G>C rs4253778	Not yet known but intronic SNPs have the potential to influence gene expression	***Phenotype: Athlete status*** • C allele more frequently distributed in post-PHV academy [28] and pro players [30] *vs.* NA, possibly due to C allele's association with greater SkM type II fibre composition [187].

	acids (FAs) by upregulation of genes involved in FA transport, FA binding activation, and peroxisomal and mitochondrial FA β-oxidation [186].	and mRNA stability [25].	• Combination of 'favourable' alleles from *PPARA* rs4253778 (C), *NOS3* rs2070744 (T), *ACTN3* rs1815739 (R), and *AGT* rs699 (G) SNPs correlated with agility performance in pre-PHV academy players [28].

SNP, single nucleotide polymorphism; *MAF*, minor allele frequency (according to European ancestry [188]); *SkM*, skeletal muscle; *NA*, non-athletes; *PHV*, peak height velocity; *RSA*, repeated sprint ability.

Table 2. Key polymorphisms associated with injury risk in soccer players. Alleles for each polymorphism are reported according to the forward DNA strand.

Gene	Encoded protein	Protein function	Alleles and rs number	Polymorphism function	Associations with injury/ poor recovery
ACTN3	α-actinin-3	Binds actin to the Z-line in type II SkM fibres [48]; blocks calcineurin, inhibiting fast-to-slow myofibre transition [26].	C(R)>T(X) rs1815739	X allele leads to a stop codon at amino acid 577, preventing protein production [178]. Consequently, the R allele is associated with greater type II SkM fibre composition [49].	• XX genotype associated with greater indices of exercise-induced muscle damage than RR genotype in pro players [139], possibly due to less structural stability of the sarcomere during contraction [144] and/or having smaller, weaker and less powerful muscles [27], thus potentially experiencing relatively greater ground reaction forces. • XX genotype associated with higher rate of non-contact musculoskeletal soft-tissue injuries in 42 pro players [140]. • XX genotype associated with higher rate of non-contact muscle injuries in XX homozygotes than R-allele carriers in pro players, with XX homozygotes most likely to suffer severe injury [44]. • X-allele carriers most likely to be injured in pro players, with XX homozygotes having the longest absence after injury [47]. • X-allele carriers missed more days following ankle injuries than RR homozygotes in academy players [29], suggesting greater muscle size and strength associated with the R allele [27] may enhance limb stability and favour return to play. • No association with hamstring injuries in pro players [40].
CCL2	C-C motif chemokine ligand-2 (CCL2)	CCL2 is expressed within the interstitial space between myofibres after damaging activity [159] and mediates systemic changes with chronic exercise [160].	G>C rs2857656	Affects plasma CCL2 levels [158]. Plasma CCL2 is lower in GG homozygotes [158], who exhibit lower muscle strength than C-allele carriers [161].	• GG genotype associated with more severe muscle injuries than C-allele carriers in pro players [45, 46]. • Not associated with hamstring injuries in pro players [40]. • Not associated with injury prevalence or recovery time in academy players [29]. • Considering the prospective role of CCL2 in satellite cell proliferation [162], higher CCL2 during recovery from muscle injury may be advantageous to soccer players carrying at least one C allele.
COL1A1	Pro-α1 (I) chain	Major component of type I collagen, a structural protein found in most connective tissues,	C>A rs1800012	Affects binding site for the *COL1A1* Sp1 transcription factor. A allele associated with greater α1 (I) chain	• Rare AA genotype appears protective against ACL [133] and Achilles tendon [147] injuries in athletes. • Pro soccer players with ACL ruptures are less likely to be AA homozygotes [39].

		including ligament and tendon [189].		mRNA and protein [145].	• However, others found no associations between this SNP and injury [29, 40, 46].
COL5A1	Pro-α1 (V) chain	Major component of type V collagen, which regulates the diameter of collagen fibrils [182].	C>T rs12722	Not yet known but 3'-UTR SNPs have the potential to alter the level, location, or timing of gene expression [25].	• TC genotype associated with severe hamstring injuries in pro players [45, 46] but sample sizes were low in these studies (n=74). • More severe musculoskeletal injuries are exhibited by TT homozygous players than C-allele carriers [42]. • C-allele carriers in two different *COL5A1* SNPs (rs12722 and rs13946) was protective against ACL rupture in pro players [41]. • Higher prevalence of soft-tissue and ligament injuries in pre-PHV academy players with the C allele and CC genotype, respectively [29]. • Considering rs12722 may influence mRNA stability and affect the tensile strength and stiffness of collagen fibrils [149], the *COL5A1* rs12722 SNP has a plausible association with musculoskeletal injuries in soccer.
EMILIN1	Elastin microfibril interfacer-1	ECM glycoprotein that assists the fusion of elastin fibres during elastogenesis [152] and provides elasticity to ligaments and tendons [153].	C>T rs2289360	Currently unknown. CC genotype associated with hypertension, possibly by reducing arterial compliance [190].	• CC genotype associated with the highest overall injury prevalence, while T-allele carriers missed the most days following ankle injuries in academy players [29]. • TT genotype associated with more severe MCL injuries, but not with the incidence or severity of muscle or tendon injuries in pro players [38, 45]. • Not associated with hamstring injuries in pro players [40].
IL6	Interleukin-6 (IL-6)	IL-6 is a myokine produced in SkM after exercise and is functionally related growth, atrophy and collagen synthesis [164, 165]. Potential role as signaling molecule associated with post-exercise satellite cell proliferation.	C>G rs1800795	Possibly affects glucocorticoid receptor and transcription. G allele associated with increased plasma IL6 [163].	• GG genotype associated with greater hamstring injury risk in pro players [40]. • Contrastingly, CC homozygotes associated with higher prevalence of all injury types and specifically muscle injury in post-PHV academy players [29]. • Having both pro- and anti- [166] inflammatory capacities, it is plausible that IL-6 could increase or reduce injury risk. • Increased plasma IL-6 associated with the G allele [163] may heighten soft-tissue injury risk. • Because eccentric contractions induce pro-inflammatory IL-6 expression [167], and because cytokines trigger tenocyte apoptosis and ECM degradation [168], GG homozygotes may experience these unfavourable changes via increased IL-6, heightening injury risk.
MMP3	Matrix metalloproteinase	One of >20 MMPs that can catalytically	C>T	Currently unknown. C>T substitution replaces	• rs679620 T allele associated with hamstring injury in pro players [40]. • rs679620 T allele associated with longer absence following knee injuries in

- 3	degrade collagens and ECM substrates [155]. Also activates other MMPs and important for SkM regeneration [156].	rs679620	glutamate residue with lysin residue but does not affect MMP3 protein function [191].	academy soccer players [29]. • rs679620 T allele is in linkage disequilibrium with *MMP3* rs3025058 5A allele, which increases MMP3 expression [157]. • rs679620 T allele might increase RNA expression and intensify the degradation of proteins, such as collagen and elastin, weakening the structural integrity of ligament and ECM. However, further research is required.

SNP, single nucleotide polymorphism; *MAF*, minor allele frequency (according to European ancestry [188]); *SkM*, skeletal muscle; *PHV*, peak height velocity.