St Mary's University Twickenham London

Evaluation of Selected Genetic, Physiological and Environmental Factors to Support the Development of Elite Male Youth Football Players

A thesis submitted by:

Edward Ryan-Moore

MSc, BSc (Hons), CSci

In partial fulfilment of the requirements for the award of Doctor of Philosophy

Faculty of Sport, Allied Health and Performance Science

St Mary's University, London

June 2022

St Mary's University Twickenham London

DECLARATION OF ORIGINALITY

Students are reminded that the work that they submit for assessment must be theirown.

Please read the following statements and sign and date at the bottom of this form toshow that you have complied:

1. This thesis and the work to which it refers are the results of your own efforts. Any ideas, data or text resulting from the work of others (whetherpublished or unpublished) are fully identified as such within the work andattributed to the originator in the text, bibliography or footnotes.

2. This thesis has not been submitted in whole or in part for any other academic degree or professional qualification at this or any other institution.

3. Any chapters that describe the outcomes of joint research should be clearly identified as such with a statement inserted as a footnote on the first page and contributors named. Significant data, images or text resulting from the input of other researchers should be identified as such and attributed to the persons concerned by means of a footnote within thechapter.

4. It is usual to acknowledge the help and guidance of others who have assisted you during your research and preparation of your thesis. Such acknowledgements do not replace or obviate the need for individual attribution as discussed in points 1 and 3.

5. The University reserves the right to submit electronic versions of your draft documents for assessment of plagiarism using electronic detection software such as 'turnitin'. In addition, whether or not drafts have been so assessed, the University reserves the right to require an electronic versionof the final document (as submitted) for assessment.

SIGNED:  ……

EDWARD RYAN-MOORE

PRINT NAME:……..……

23RD OF JUNE 2022

DATE:……

Thesis Abstract

Title: Evaluation of Selected Genetic, Physiological and Environmental Factors to Support the Development of Elite Male Youth Football Players. **Author**: Edward Ryan-Moore. **Degree:** Doctor of Philosophy, St Mary's University, London. **Date:** 23rd of June 2022

Time-loss injury incidence disrupts access to developmental opportunities for elite male youth footballers. Interindividual genetic variations influence musculoskeletal injury susceptibility along with environmental factors and biological processes of growth and maturation. Greater understanding of these factors and how they may influence each other could inform targeted and individualised training programmes to support the long-term development of elite footballers. The aims of this research were to evaluate and explore the current applicability of selected genetic, physiological and environmental factors to inform support for the long-term development of elite male youth footballers.

Initial scoping of literature highlighted injury prevention as the most ethically appropriate and practically meaningful area to apply genetic information to inform support of long-term elite footballer development at this time. A systematic review and meta-analysis of candidate gene association studies with fracture risk in physically active participants was completed to identify candidate genes for subsequent analysis along with existent literature. The influence of growth and maturation on injuries in elite male football were then explored to identify physiological factors which could mediate the influence of genetic variants on injury susceptibility. The association between candidate genes (*COL1A1*, *COL1A2*, *COL5A1*, *ACTN3*, *ESR1*, *MMP3*, *ACE*, *VDR*, and *GDF5*) and tissue-specific injury incidence in elite male youth footballers were then explored to identify genotypes which influence injury risk. Finally, growth, maturation, loading exposure, and tissue-specific total genotype injury risk score were combined to explore how different risk categories influence injury incidence, and which could inform targeted interventions to reduce injury risk in elite male youth football.

The research contributed to current knowledge with the following findings: identifying a potential sex-specific influence of the *COL1A1* (rs1800012) T allele on fracture risk in young physically active females from the extant literature; highlighting that growth rate alone appears unable to sensitively identify individuals at increased risk of injury when measured every 3-12 weeks, unlike previous research using longer measurements, which may have observed spurious relationships with maturational processes occurring alongside growth; detecting significant genetic associations between the *MMP3* (rs679620) and *VDR* (rs2228570) SNPs with non-contact injury, and the *COL1A2* (rs412777) variant with fracture risk ($p \leq 0.01$); observing that when incorporating growth rate and loading exposure into a risk model, *COL1A1* (rs1800012) T allele carriers and those experiencing rapid growth (>0.6 cm.m^{-1}) were associated with significantly greater risk of apophysitis injury ($p \leq 0.03$). These findings contribute towards the long-term goal of supporting elite male youth footballers to achieve their potential with genetically informed individualised training programmes.

Acknowledgements

I would first like to thank my partner Valentina for all her love and support throughout my PhD research. Without her countless hours of advice, reassurance and proof reading the thesis would not be as well written and she has directly contributed to getting the research to where it is today. However, more than anything I am forever grateful for Vali's help in celebrating successes, patient care through moments of stress, and understanding of time we have had to sacrifice together so that I could complete the work. I am not sure I would have made it through without her. I love you and I will be forever grateful. I would also like to thank my wider support network of friends and family for always being there to pick me up when I was down, and keep me going on by helping me to see the light at the end of the tunnel. I am incredibly grateful and fortunate for all the people around me who have, in no uncertain terms, enabled me to complete the research through their support, love and guidance.

I would also like to thank my supervisors, Yiannis Mavrommatis, Mark Waldron, Tom Sawczuk and Jamie North for their exceptional ability to balance the necessity for the research to keep moving forward, despite competing demands, while showing genuine care and compassion for the challenges I have faced during the journey. Their faith and support in my ability to complete the work has always been a source of strength and motivation for me. The overwhelming quality of my supervisory team spans many areas of humanity and academic expertise. However, I am most grateful for their humility, which has allowed me to be supported throughout all the challenges we have faced with little in return. My supervisors are outstanding research professionals and people, without their support and patience I would not have completed the PhD. Thank you.

I would also like to thank Fulham Football Club and St Mary's University for fundings the research. In addition to the support of many within both organisations, I would like to specifically thank; James Dunkin and Jon Goodwin for their continuous support, stimulating debates, and roles in the collection of anthropometric measurements, and progression of the physical development programme at Fulham Football Club enabling me to protect time for the research; Laura Bettinelli for her practical insight, continuous support and guidance, and expertise to scrutinise the injury classification records; Liam Beale, Michael English, Alissa Dibley and Natalie Creek for their diligence and detail in the assessment and documentation of injuries, complaints, and treatments at the club along with their time in reviewing historical notes. In addition to those above I would like to thank all the staff, players and parents at Fulham Football Club who have supported me and allowed me to complete my PhD research over the years. To everyone, thank you.

Table of Contents

Final Thesis Word Count: 82381

List of Tables

List of Figures

Research Outputs

Peer Reviewed Journal Articles

Ryan-Moore, E., Mavrommatis, Y., & Waldron, M. (2020). Systematic Review and Meta-Analysis of Candidate Gene Association Studies With Fracture Risk in Physically Active Participants. *Frontiers in Genetics*, 11(June). https://doi.org/10.3389/fgene.2020.00551

CHAPTER 1: Thesis Introduction

1.1 Thesis summary

Chapter 1 outlines the thesis structure, research context and rationale. Ethical considerations and the potential value of genetically informed training programmes are discussed to support long-term development for every individual in an elite male football academy. The evolution and formation of the thesis aims are explained as initial scoping of literature highlights that the current understanding of genetic factors affecting performance is insufficient to inform applied practice. Protecting individuals from harm with targeted interventions, based on variants which mediate susceptibility to musculoskeletal injury risk emerges as a promising area of exploration.

Chapter 2 explores literature on: genetics, long-term development, football performance, injury, and training theory. The findings of genetic association studies with candidate genes are described and potential mechanisms of influence on injury risk explored. Finally, the research aims are outlined: to systematically explore existing literature to identify candidate genetic variants associated with injury risk in healthy physically active males, and to evaluate the potential applicability of selected genetic, physiological and environmental factors to reduced injury and support the long-term development of elite male youth football players.

Chapter 3 details the systematic review and meta-analysis of candidate gene association studies with fracture risk in physically active participants. The 5,699 identified records were refined to 10 genetic variants from six different genes in 10 studies for the final quantitative meta-analysis. The *COL1A1* (rs1800012) T allele was associated with a significant trivial reduction of fracture risk in physically active females only and may not be ubiquitously detrimental to bone strength, as has been previously suggested. The genetic penetrance of this variant appears to be influenced by sex and age. No other significant overall effect was observed but further investigation of the *COL1A1* (rs1800012), *COL1A2* (rs412777) and *VDR* (rs2228570) variants with fracture risk was warranted.

Chapter 4 examines the influence of growth and maturation measured every 3-12 weeks on the incidence of injuries in an elite male English Premier League football academy. Data from 80 players from the Under-9 to Under-23 age groups, across six seasons, indicated that growth rate alone was not associated with injury risk. However, injury incidence significantly increased with maturation for all, non-contact, and non-contact muscle injuries, but decreased for apophysitis injury. It is mechanistically unlikely that growth rate is related to all injuries but a causal link with apophysitis injury remains unclear, due to the potential contribution of temporal disturbances in bone strength around puberty. Nevertheless, stature growth alone appears to be insensitive to detect periods of increased injury risk for individuals when measured at a practically meaningful frequency.

Chapter 5 investigates the genetic association between previously identified candidate variants and all, non-contact, non-contact muscle, fracture, ligament, tendon and apophysitis injuries. Injury surveillance data between July 2013 and June 2021 in 112 elite male footballers between 10 and 33 years of age was included. Significant genetic associations were observed between *COL1A2* (rs412777) and fracture risk, *MMP3* (rs679620) and *VDR* (rs2228570) with non-contact injury, and *COL5A1* (rs12722) with non-contact muscle injury ($p \leq 0.01$). The *GDF5* (rs143383), *COL5A1*

(rs12722) and *COL1A1* (rs1800012) variants were also indicated to influence fracture, tendon and apophysitis injury risk, respectively ($p \leq 0.03$), but were considered with greater caution due to the number of independent hypothesis-driven tests.

Chapter 6 examines the potential interaction of tissue-specific total genotype injury risk scores with growth or maturation and loading exposure on non-contact, non-contact muscle, tendon, joint, fracture and apophysitis injury incidence. No significant injury risk model was observed ($p > 0.16$), except for apophysitis injury ($p < 0.04$). Possession of the *COL1A1* (rs1800012) T allele ($p = 0.03$) and rapid growth (>0.6 $cm.m^{-1}$) ($p = 0.02$) significantly increased apophysitis injury risk. Lower bone mineral density and formation relative to resorption associated with the T allele may interact with a temporal reduction in bone strength around puberty to significantly increase the risk of apophysitis injury during rapid growth. Current literature indicates increased dietary calcium supplementation could mitigate and even reverse this increased susceptibility.

Chapter 7 concludes by summarising and discussing the overall findings of the research and its implications. The research identified a novel association between the *COL1A1* (rs1800012) variant, rapid growth (>0.6 $cm.m^{-1}$) and apophysitis injury. If this can be replicated, it is possible that nutritional interventions could protect adolescent male youth footballers who are *COL1A1* (rs1800012) T allele carriers from apophysitis injury. Nevertheless, the overall conclusion of the research is that our current understanding of the dynamic and complex interactions between environmental, physiological and genetic factors remains insufficient to meaningfully inform development in elite male football. Theoretically, genetically individualised training programmes remain a valid area of future investigation. Future research should integrate numerous factors to develop complex holistic tissue-specific injury risk models to protect players from harm and support long-term development.

1.2 Research context

As a category one football academy, Fulham Football Club is required to meet and exceed a variety of quality assurance criteria set out in the Elite Player Performance Plan by the Premier League (Premier League, 2012). Specifically, a category one academy sports science department is expected to be actively engaged in research and development (Premier League, 2012). This was identified as an area for development and the opportunity arose for me to complete a PhD research project alongside my work as a sports scientist in the academy. However, because the desire for research was not instigated by the need to answer a specific question, I was able to explore potential options for innovation, which could support player development. This meant the preparation for, and start of, the PhD project was very exploratory as I began gaining greater understanding of the applied context and potential research topics.

The formation of the present research project began following an approach from a direct-to-consumer genetic testing company to Fulham Football Club. This company claimed that they could generate percentile estimates of adult physical performance capabilities using genetic samples of academy players. At this time, I believed that, if accurate, such estimates could support individualisation of player development and retention of late-maturing players with unseen potential. Observable differences in performance in elite male youth football development are substantially confounded by physical differences attributable to growth and maturation (Philippaerts et al., 2006). This creates a false, and difficult to control, conceptual framework of player potential as elite performers during development ages may not be the elite players in adulthood when physical differences are determined by genetics and training (Davids & Baker, 2007). Therefore, I drafted a research proposal seeking to investigate the potential applications of genetic information to support long-term athletic development in elite male youth football players. However, when exploring the literature on genetic determinants of physical performance, and considering the contextual considerations of football performance, I realised that this avenue of investigation would yield little to no applied innovations. Firstly, although the ethical implications of utilising genetic information are complex and without consensus, there is general agreement that it should not be used to determine future potential in youth athletes (Vlahovich et al., 2016; Wackerhage et al., 2009; Williams et al., 2016). However, it also became clear that our understanding of the genetic determinants of physical capability lacked the confidence for any predictions of future physical performance potential to be accurate. Therefore, the research project needed to evolve and pivot towards a more valuable area of exploration for identification of applied intervention innovations.

The Fulham Football Club Academy mission statement seeks to support every individual to achieve their potential. This axiom permeates everything we do and, consequently, also helped to shape the evolution of the research question. As a practising sports scientist I was motivated to conduct research that was feasible alongside, and which could improve, my applied support within an elite male youth football development organisation. Therefore, the research shifted to focus on exploring how genetic information could support player development through more efficient, targeted and bespoke interventions at the moment of measurement.

Completing research whilst being embedded within the applied sport context provides opportunities to innovate practise from research-practitioner perspective. This is commonplace in

healthcare services with Clinical research-practitioners recognised as an occupational group by the Professional Standards Authority. An application gap can be conceptualised, whereby research and innovation is carried out by those who may lack contextual understanding of the elite sport environment within which a new system or process may be applied (Burden et al., 2022). Conversely, those within elite sport may be resistant to change and reluctant to adopt a new system or process, which represents uncertainty and movement away from a tried and trusted previously winning formula (Burden et al., 2022). However, research-practitioners may be better placed to combine contextual expertise with research in elite sport to develop impactful applied innovations (Burden et al., 2022). Adopting this research perspective allowed me to understand what others within the Fulham Football Club Academy considered to be the determinants of successful football performance. Combining this understanding with a continuous review of genetics and physical performance literature I concluded that injury prevention was a more viable topic of exploration to support long-term player development.

Football performance is not determined by a singular exceptional physical performance capability but rather an ability to utilise the dynamic interaction of physical, technical, tactical and psychological factors to successfully execute football actions within the temporal and spatial constraints of the game moment (Reilly et al., 2000). Therefore, the ability to practise how to be effective, based on individual characteristics, is arguably of greater importance to football performance, than physical performance alone as players can compensate for weaknesses in one area by strengths in another (Davids & Baker, 2007; Reilly et al., 2000). Consequently, although physical performance is undoubtedly important to football performance, and a potential performance limiter, it is not solely deterministic to success (Reilly et al., 2000). However, injury prevents access to football specific developmental opportunities (Larruskain et al., 2021). Therefore, genetically individualised injury prevention strategies were considered to have more potential applicability in supporting long-term holistic football development than future physical performance estimation. Furthermore, the use of genetic information to prevent individuals from harm is broadly acknowledged as a more ethically appropriate use of the information (Vlahovich et al., 2016; Wackerhage et al., 2009; Williams et al., 2016).

1.3 Research introduction

Genetics represent access to information underpinning factors inherent to an individual, which were previously unavailable and, therefore, the potential to identify novel opportunities to support applied practise in elite sport (Pickering & Kiely, 2019). The progression of knowledge initially imagined with the discovery of genetics has not yet been realised, as the dynamic and complex reality of genetic expression became apparent (Pickering & Kiely, 2019). Nevertheless, research continuing to progress towards a greater understanding of genetic factors which contribute to injury risk could improve our ability to support the development of every individual. Indeed, some promising research exists which appears to be able to differentiate individuals at different levels of risk (Pickering & Kiely, 2019). Confidence in asserting the effectiveness of a novel intervention, based on rigorous scientific research, can help to overcome the initial resistance to innovations, which can occur in elite sport (Burden et al., 2022). Therefore, the present research sought to identify the

strongest and most consistent genetic associations with injury risk observed in the literature in an attempt to identify what, if any, genetically individualised interventions could be utilised to support long-term elite male youth player development.

Evaluation in the present research was considered based on the American Evaluation Association definition as: “the systematic process to determine merit, worth, value or significance.” (American Evaluation Association, 2014, p.1). Within this systematic process the research will explore the current applicability of selected genetic, physiological, and environmental factors to support the long-term development of elite male youth football players. As custodians of development and wellbeing, football academies continuously monitor players’ progress in an attempt to forecast long-term potential and design the most effective training programme to maximise it. However, a one-size-fits-all approach to long-term development fails to be flexible enough to match the temporal training requirements of individuals within a group at any given time. Indeed, significant interindividual differences are seen in the response to homogenous training interventions (Bouchard et al., 1999, 2011; Gonzalez et al., 2016) and the importance of individualisation in sports science and exercise medicine is well documented (Buckthorpe et al., 2019; Lahti et al., 2020; Pickering & Kiely, 2019). Inherent genetic differences influence the differences in training responsiveness (Battié et al., 2008; Bouchard et al., 1998; De Moor et al., 2007; Silventoinen et al., 2008) and genetically informed training programmes have been considered as an emerging strategy for individualised developmental interventions (Jones et al., 2016; Pickering & Kiely, 2019). Nevertheless, our understanding of the heritable factors of sport and exercise related traits remain limited, despite rapid growth in the research (Ahmetov et al., 2016; Rahim et al., 2019).

The nature *vs*. nurture debate is a popular (false) dichotomy, which often divides opinion. Heritable differences clearly provide performance advantages and disadvantages between individuals (Battié et al., 2008; Bouchard et al., 1998; De Moor et al., 2007; Silventoinen et al., 2008). However, the complex nature of football performance means that innate weaknesses can, to some degree, be overcome or managed by compensating with strengths in other areas, which can be learnt through an appropriate training environment (Davids & Baker, 2007; Reilly et al., 2000). This applies to both an individual in their own performance, minimising weaknesses while maximising strengths, and against opponents, targeting weaknesses and nullifying strengths, both of which can be tactically designed in a team / individual playing strategy. Therefore, although elite footballers require a fundamental baseline competency of physical, technical, tactical, and psychological abilities, the opportunity to develop strategies for effective performance may be the most important factor for successful long-term development. Consequently, time-loss injury could have the most disruptive impact on long-term football development, as it precludes access to training and competition exposures whereby such strategies for success are learnt and refined (Davids & Baker, 2007). Indeed, time-loss injury has recently been shown to hamper the progression of elite male youth football players (Jones et al., 2019; Larruskain et al., 2021). Consequently, understanding how genetic variations influence injury susceptibility could allow interventions to be implemented, which aim to mitigate individual injury and increase players’ exposure to football development opportunities. Therefore, the aim of this research project was to try and investigate which, if any, genetic variants could be used to support elite football player development in an applied context via individual injury risk mitigation strategies with our current knowledge.

CHAPTER 2: Literature Review

This chapter aims to describe background concepts, such as trait heritability, fundamentals of genetics and genetic variation. An overview of the structure and objectives of elite youth football academies is described, along with musculoskeletal injury aetiology and common injuries in elite male youth and adult football. The influence of physiological and environmental factors, such as growth, maturation, training and progressive overload on injury risk in football is discussed and potential applications of genetic information on the long-term development of elite footballers explored. The findings of previous genetic association studies with musculoskeletal injuries are discussed to identify viable candidates for evaluation in the present research. The research aims are then explicitly outlined based on the literature review to systematically explore existing literature to identify candidate genetic variants associated with injury risk in healthy physically active males and to evaluate the potential applicability of selected genetic, physiological and environmental factors to reduced injury and support the long-term development of elite male youth football players.

2.1 Genetics and heritability of sport and exercise phenotypes

Genes represent the basic physical and functional units of heredity by encoding for proteins, and other components of the body, which affect the interindividual variability of observable traits known as phenotypes (Gibson, 2016; Wackerhage, 2014). Genes are sections of deoxyribonucleic acid (DNA) formed from polynucleotide molecules using four different single nucleotide bases; cytosine (C), guanine (G), adenine (A), or thymine (T) (Watson & Crick, 1953). Genes can range in size from a few hundred to more than two million nucleotide bases (Pertea & Salzberg, 2010). The sequence of G, C, T and A nucleotide monomer units form an alternating sugar-phosphate backbone and provide the information encoded in the gene (Watson & Crick, 1953). There is estimated to be three billion nucleotide bases and 22,333 genes in the human genome, which is organised into 23 chromosomal pairs (Pertea & Salzberg, 2010). An example of the structure of DNA from the cell to nucleotides can be seen in Figure 1. One chromosome of a chromosomal pair is inherited from the paternal genome, whilst the other is inherited from the maternal genome and, consequently, every individual possesses two copies of each gene (Gibson, 2009). Nevertheless, all humans are 99% identical (Levy et al., 2007) and, therefore, every observable difference between individuals results from variation in less than 1% of the genome (Davids & Baker, 2007).

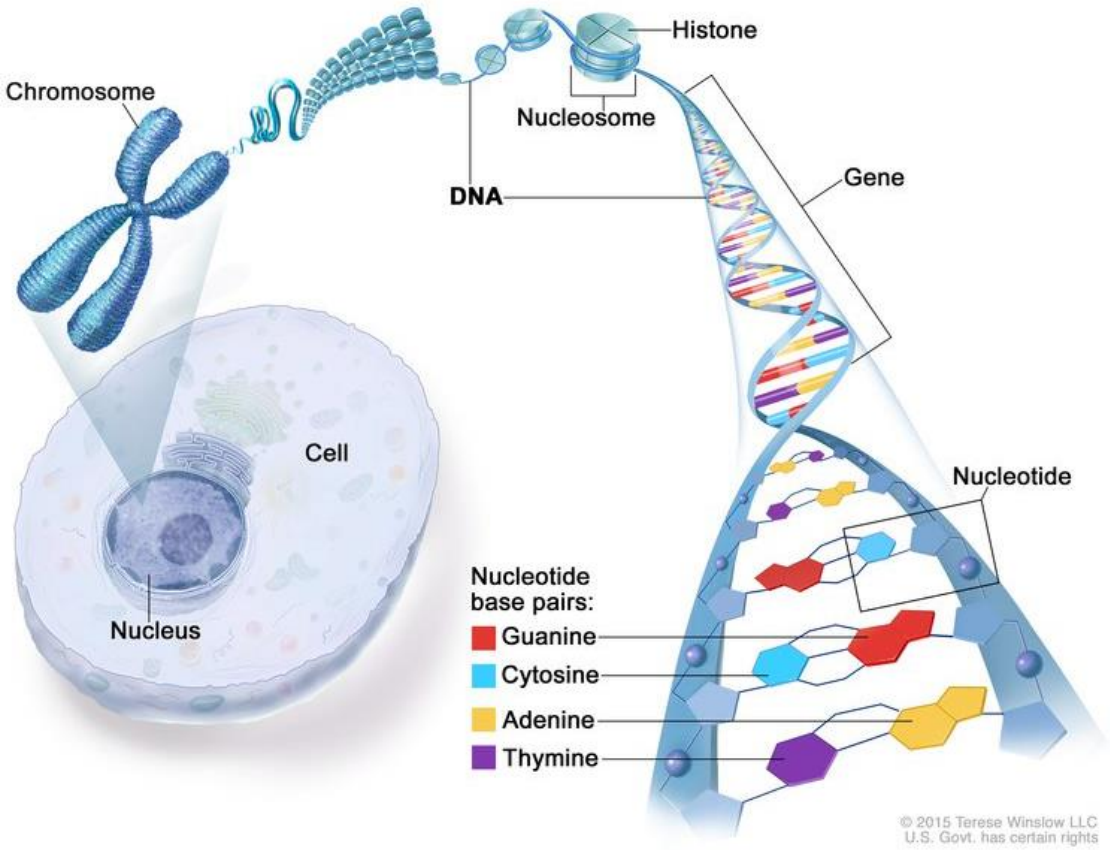

Figure 1. Structure of DNA from Cell to Nucleotide. Reused with permission from Terese Winslow © (2015) Terese Winslow LLC, U.S. Govt. has certain rights.

Heritability estimates provide an approximate quantification of the relative contribution of hereditary factors to the interindividual variability of a trait (Falconer & MacKay, 2011; Visscher et al., 2008). Heritability estimates for sports and exercise related traits appear substantial for overall athletic status (66%) (De Moor et al., 2007), maximal aerobic capacity (~50%) (Bouchard et al., 1998), strength (60%) (Silventoinen et al., 2008), flexibility (47%) (Battié et al., 2008) and injury risk (40-69%) (Andrew et al., 2004; Hakim et al., 2003; Magnusson et al., 2020). These phenotypes have a genetic architecture which describes; the contributory genetic variants, the size of their effects, frequency in the population and interactions with other variants and the environment (Timpson et al., 2018). Sport and exercise related phenotypes are complex traits, influenced by large numbers of genetic and environmental variables (Bouchard et al., 1999; Wackerhage, 2014). This complexity quickly increases when considering the interaction of multiple traits to influence injury risk in football performance. Despite this complexity, the genetic architecture of heritable traits generally results from the accumulation of a finite number of both; large effects from a few rare, and small effects from numerous common, genetic variations between individuals (Gibson, 2016). Therefore, understanding how genetic variants influence injury risk could inform training strategies targeting individual susceptibilities to help protect players from harm.

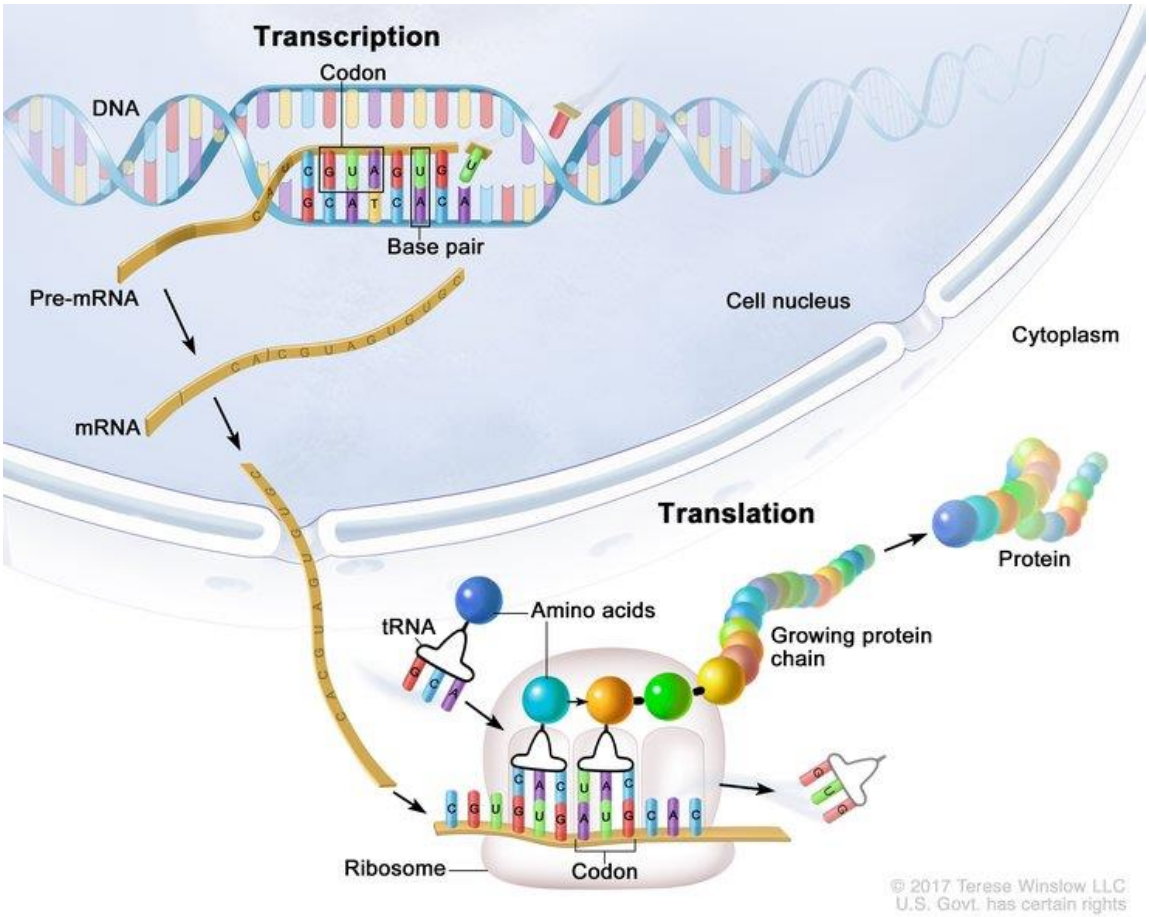

Figure 2. Genetic transcription and translation to produce protein. Reused with permission from Terese Winslow © (2017) Terese Winslow LLC, U.S. Govt. has certain rights.

There are different types of genetic variation between individuals but the most common, frequently studied and influential on the expression of phenotypes are single nucleotide polymorphisms (e.g., C in place of T at the same genomic location), known as SNPs (International HapMap Consortium, 2005; Stranger et al., 2009). The DNA nucleotide sequence in genes encode for proteins and proteins affect physiological function (Gibson, 2016). Gene expression represents the activation of the genetic machinery to produce proteins by exposing the tightly bound DNA and transcription of DNA to Ribonucleic acid (RNA) in the nucleus, followed by the translation of RNA to protein by Ribosomes, as outlined in Figures 1 and 2. Initiation of gene expression occurs as tissues detect the physical and chemical changes experienced during exercise, which result in signal transduction pathways specific to the stimulus (Wackerhage, 2014). These signals cause the DNA, which is tightly bound around histone molecules, to become accessible to transcription factors at targeted locations and specific genes (Wackerhage, 2014). The RNA polymerase enzyme then unzips the double stranded DNA and transcribes a sequence of complimentary base pairs from the DNA template strand (3' to 5' direction) into Pre-mature RNA (Pre-mRNA) (Wackerhage, 2014). In this way the Pre-mRNA nucleotide sequence matches that of the DNA coding strand (5' to 3' direction), with the exception of Uracil in place of Thymine due to the different chemical structure of RNA to DNA (Wackerhage, 2014). Pre-mRNA is then spliced by spliceosomes as introns – sections of the gene that do not encode amino acids – are cut out, and exons – sections of the gene encoding the amino acid sequence of the expressed protein – joined together to form mature messenger-RNA (mRNA) (Wackerhage, 2014). The mRNA then travels outside the nucleus to the ribosomes so that

the encoded protein can be translated into a polypeptide chain of amino acids (Wackerhage, 2014). The triplet sequence of nucleotide bases on the mRNA form a codon, which bind to specific transfer RNAs (tRNAs) carrying specific amino acids (Wackerhage, 2014). The polypeptide chain then forms as the mRNA, ribosome and tRNA molecules encode, combine and supply the amino acids in the specific order which, in turn, determines the structure and function of the resultant protein (Wackerhage, 2014).Therefore, a SNP located in an exon can completely alter the amino acid sequence, structure and physiological function of the protein encoded by the gene (North et al., 1999). Alternatively, SNPs located in introns may alter the physiological activity of proteins by altering gene expression (Jo & Choi, 2015). Consequently, SNPs can have substantial physiological consequences to mediate the observable variation in phenotypes between individuals.

A form of a genetic variation which could result in differential phenotype expression is referred to as an allele (Wackerhage, 2014). Alleles can be denoted by the different nucleotides present at the variant location (e.g., C allele or T allele) or by the genetic consequence of each allele (e.g., R for the arginine codon allele and X for the stop codon allele). Numerous genetic variants can be present in a single gene and proximal portions of the DNA are frequently inherited together (Lewontin, 1964). Therefore, the frequency of alleles can vary within the population and significantly differ between ethnicities (Gibson, 2016). If different alleles demonstrate non-random inheritance patterns, they are considered to be in linkage disequilibrium and common groups of alleles, known as haplotypes, can occur (Lewontin, 1964). The most common allele in a group is referred to as the wildtype, or major, allele with other variants referred to as minor alleles (Gibson, 2016). As we each possess two copies of each gene, if a genetic variant has two possible forms, or alleles, then three combinations, known as genotypes, of these alleles are possible for each individual (e.g., CC, CT or TT) (Gibson, 2016). Individuals who possess two copies of the same allele (e.g., CC & TT) are termed homozygotes and those who have two different alleles (e.g., CT) heterozygotes (Gibson, 2016). Expression of a particular phenotype may only occur in certain genotypes based on numerous factors (Gibson, 2016). For example, the risk of injury may only significantly increase for homozygotes who possess two copies of the risk allele (recessive) or heterozygotes who only carry one risk allele (dominant), in addition to other patterns which will be discussed later (Gibson, 2016). Furthermore, the same genetic variant and genotype may not affect everyone equally and genetic penetrance describes the probability that an individual who possesses an allele associated with a trait will express it (Cooper, Krawczak, Polychronakos, Tyler-Smith, & Kehrer-Sawatzki, 2013). An allele with complete penetrance will always result in the carrier expressing the phenotype but reduced or incomplete penetrance does occur and can be dependent up age, sex or ethnicity (Cooper et al., 2013).

Understanding the genetic architecture of physical performance traits, such as strength, endurance, flexibility and injury risk, in combination with the genetic profile of an individual would dramatically improve our understanding of an athlete. Theoretically, this information could then be used to better inform the complex decision-making process required to support long-term development with targeted interventions designed to protect the athlete from injury based on particular inherent susceptibilities or sensitivities. This ideal is driving an increasing effort towards personalised training and treatments in other areas, such as medicine, to optimise individual outcomes (Abt & Lovell, 2009; Jackson & Chester, 2015). Indeed, companies have begun marketing

tests to identify specific genetic variants associated with sports and exercise related traits, with genetically informed training programmes already being trialled (Jones et al., 2016; Pickering et al., 2018). Whilst such strategies are theoretically possible, current understanding of the complex interactions between environmental and genetic factors, which also vary by physiological processes of maturation and aging, is still limited for effective application (Kraemer & Ratamess, 2004). Advancing the understanding of genetic variants which influence injury risk in elite male football has the potential to improve the personalisation of interventions to optimise individual performance development opportunities.

2.2 Football academies, development, injury, growth and maturation

Football academies provide multi-disciplinary support and organise developmental training programmes for those in the local area who appear to have the greatest potential for long-term football performance success. The aim of these academies is to support the holistic development of elite professional football players to compete for club and country, and both national governing bodies and individual clubs invest substantially in achieving this goal (Ernst & Young LLP, 2022). In England, the academy system organises players into competitive groups based on chronological age (Under-9 to Under-23) between September 1st and August 31st the following year. Therefore, players may join a club age eight and remain there until they are aged twenty-three, and older if they sign as professional adults. This represents a substantial amount of time during which dramatic physical, personal, social, emotional, technical and tactical development can occur (Davids & Baker, 2007; Lloyd et al., 2016; Ribeiro et al., 2021; Ryan et al., 2018). Exposure to deliberate practice during this time can facilitate the development of elite performance, as players learn to become effective footballers (Davids & Baker, 2007; Ryan et al., 2018). The physical activity required for football participation has repeatedly been shown to provide long-term health and wellbeing benefits (García-Hermoso et al., 2019; Granger et al., 2017; Lubans et al., 2016; Torres-Costoso et al., 2020; Utesch et al., 2019). Nevertheless, a structured elite training and competition environment also represents exposure to potentially injurious forces (Bowen et al., 2017, 2020).

Injuries occur when the repetitive and / or acute loading of musculoskeletal tissue exceeds the threshold tolerance of an individual at that time (Kalkhoven et al., 2021; Meeuwisse et al., 2007; Nielsen et al., 2018). The subsequent injury rehabilitation time disrupts access to developmental experiences for elite youth footballers and restricts opportunities to realise their long-term potential (Jones et al., 2019; Larruskain et al., 2021). Non-contact injuries of muscle, tendon, ligament and bone are particularly prevalent in football, with muscle strains and ligament sprains amongst the most frequent injuries, accounting for approximately 20-30% of all injuries each (Le Gall et al., 2006; Pfirrmann et al., 2016; Price et al., 2004; Read et al., 2018b; Rumpf & Cronin, 2012). The risk of injury is three to five times higher in matches than training, although, the risk of injury in training appears higher for youth than adult players (Brito et al., 2012; Le Gall et al., 2006; Pfirrmann et al., 2016; Rumpf & Cronin, 2012). Players experiencing rapid pubertal growth develop their physical capabilities alongside increased muscle mass, strength, power and endurance, which translates into their athletic abilities, resulting in greater jump height, sprint speed and shot velocity (Hansen et al., 1999; Malina et al., 2005a; Philippaerts et al., 2006; Saavedra et al., 1991). This increases the

strength of forces an individual experiences both, within their own musculoskeletal tissues, and those imposed upon them from other players in training and competition. Consequently, the incidence and severity of injuries in football increase with age, except for growth-related apophysitis overuse injuries, which appear to peak in the Under-13 & Under-14 groups (Le Gall et al., 2006; Price et al., 2004; Read et al., 2018b; Rumpf & Cronin, 2012). Apophysitis injuries account for 13-25% of all injuries between the ages of 11 and 14 years and each occurrence can result in over 28 days lost from training and matches (de Loës, 1995; Le Gall et al., 2006; Light et al., 2021; Materne et al., 2020; Wik et al., 2020a). Apophysitis injuries frequently affect the knee (Osgood-Schlatter's Disease) and ankle (Sever's Disease), which are also the most common sites of injury overall in football, each accounting for roughly 20% of all injuries (Le Gall et al., 2006; Price et al., 2004; Read et al., 2018b).

Apophysitis injuries are thought to result from repeated microfractures of bone at the apophyseal attachment site due to an inability to withstand strong shear forces exerted by the tendon during exercise (Arnold et al., 2017; Gholve et al., 2007; Holden & Rathleff, 2020). Although generally considered to resolve after six months, some have observed potential long-term implications of apophysitis injury on pain and physical performance (Guldhammer et al., 2019; Kaya et al., 2013; Rathleff et al., 2020). Apophysitis injuries appear more common during maturational phases of rapid growth and may be related to a transient reduction in cortical bone mineral density (BMD) (Faulkner et al., 2006; Wang et al., 2010; Wang & Seeman, 2009). Estimates suggest that BMD accounts for approximately 70% of bone strength and heritability estimates range between 50-85% (Tinkle & Wenstrup, 2005). Bone fracture susceptibility is predominantly determined by bone geometry and mineral density (Fonseca et al., 2014; Hernandez & van der Meulen, 2017) and peak incidence of fracture also appear to coincide with pubertal maturation (13-14 years in boys) (Cooper et al., 2004; Parfitt, 1994). Indeed, fractures are the most frequent cause of hospitalisation for children, accounting for up to 25% of admissions due to injury (Cooper et al., 2004) and some estimates have suggested that up to 50% of all children will have experienced a fracture before adulthood (Tinkle & Wenstrup, 2005). Therefore, it is likely that genetic variants that influence apophysitis and fracture injury risk in elite male youth footballers could be particularly informative to reduce injury occurrence in youth footballers. Furthermore, targeted nutritional and / or training interventions have been shown to reduce fracture risk (Anderson et al., 2017; Clark et al., 2011), indicating the potential of applied support to players' continued development by participation in a training programme.

The association between incidence of apophysitis injury in the Under-13 & Under-14 groups in elite youth football, which coincide with pubertal growth, has encouraged clubs to assess the growth and maturation status to understand when interventions could reduce the risk of injury (Materne et al., 2021; Parfitt, 1994; Read et al., 2018a; Towlson et al., 2021). The peak in adolescent stature growth is referred to as peak height velocity (PHV) and is followed by a slowing and eventual cessation of growth as players mature into adulthood (Abbassi, 1998; Tanner & Whitehouse, 1976). Males enter puberty between the ages of 11 and 13 years as growth rate begins to increase with PHV occurring between 13 and 16 years of age (Tanner & Whitehouse, 1976) with most experiencing PHV around 13 to 14 years as shown in Figure 3 (Abbassi, 1998). The age at take-off (when the adolescent growth spurt begins) has been shown to vary between age 8 to 13 years in boys (Malina et al., 2016). Indeed, substantial differences in the timing, duration, and tempo of physical maturation is observed between individuals, as shown in Figure 4 (Tanner et al., 1966). Between three months

to a year after PHV a period peak of weight velocity (PWV) occurs as circulating androgens stimulate muscle growth and significant improvements in physical performance occur (Lloyd et al., 2015; Tanner & Whitehouse, 1976; Van Hooren & De Ste Croix, 2020). These physiological changes mean that maturational stages are frequently separated into Pre-, Circa- and Post-PHV phases in long-term athletic developmental programmes, as shown in Figure 5 below (Van Hooren & De Ste Croix, 2020). This allows youth strength and conditioning coaches to design training with consideration of sensitive periods of adaptation which are adjusted to account for individual maturation status.

The authors of long-term athletic development models often acknowledge the need to account for various stages of pubertal maturation to reduce injury risk (Faigenbaum et al., 2009; Lloyd et al., 2015; Ryan et al., 2018). Typically, treatment for apophysitis injury involves rest from physical activity, therefore, some have recommended that training load should be reduced during PHV to reduce injury risk (Arnold et al., 2017; DiFiori et al., 2014; Faigenbaum et al., 2009). However, the PHV phase can span entire seasons, during which time the reduction in training load may actually exceed the time which would have been lost from injury if no reduction occurred. Anecdotally, some players suffer apophysitis injuries prior to, or after, PHV when growth rate may or may not be rapid and the one-size-fits-all approach to training load modifications during PHV may be too broad to result in an overall net decrease in time lost from football development opportunity. Indeed, although several studies have observed associations between pubertal growth and injury risk (Kemper et al., 2015; Wik et al., 2020b), a systematic review concluded that insufficient evidence existed to indicate that physical growth and / or maturational status during adolescence are associated with injury and further investigation is required (Swain et al., 2018). Furthermore, about half of peak adult bone mass is accumulated during adolescence (Parfitt, 1994), which can be augmented with appropriate mechanical loading and is much greater than that possible in adulthood (Abrams, 2003; Khan et al., 2000). That withstanding, fracture and apophysitis injury incidence appear to peak during adolescence, as a temporary reduction in bone strength occurs around pubertal growth (Parfitt, 1994; Wang et al., 2010). Consequently, a challenging dichotomy occurs during adolescence, whereby bones are more fragile, yet possess greater ability to accumulate lifelong strength (Parfitt, 1994). Therefore, a blanket reduction in training load for circa-PHV individuals could potentially be detrimental to long-term development and a more precise identification of individuals at increased risk is required. Furthermore, it may be that a change in training type and / or nutritional interventions may be targeted for individuals in need of support to reduce the risk of apophysitis injury but maintain long-term strength.

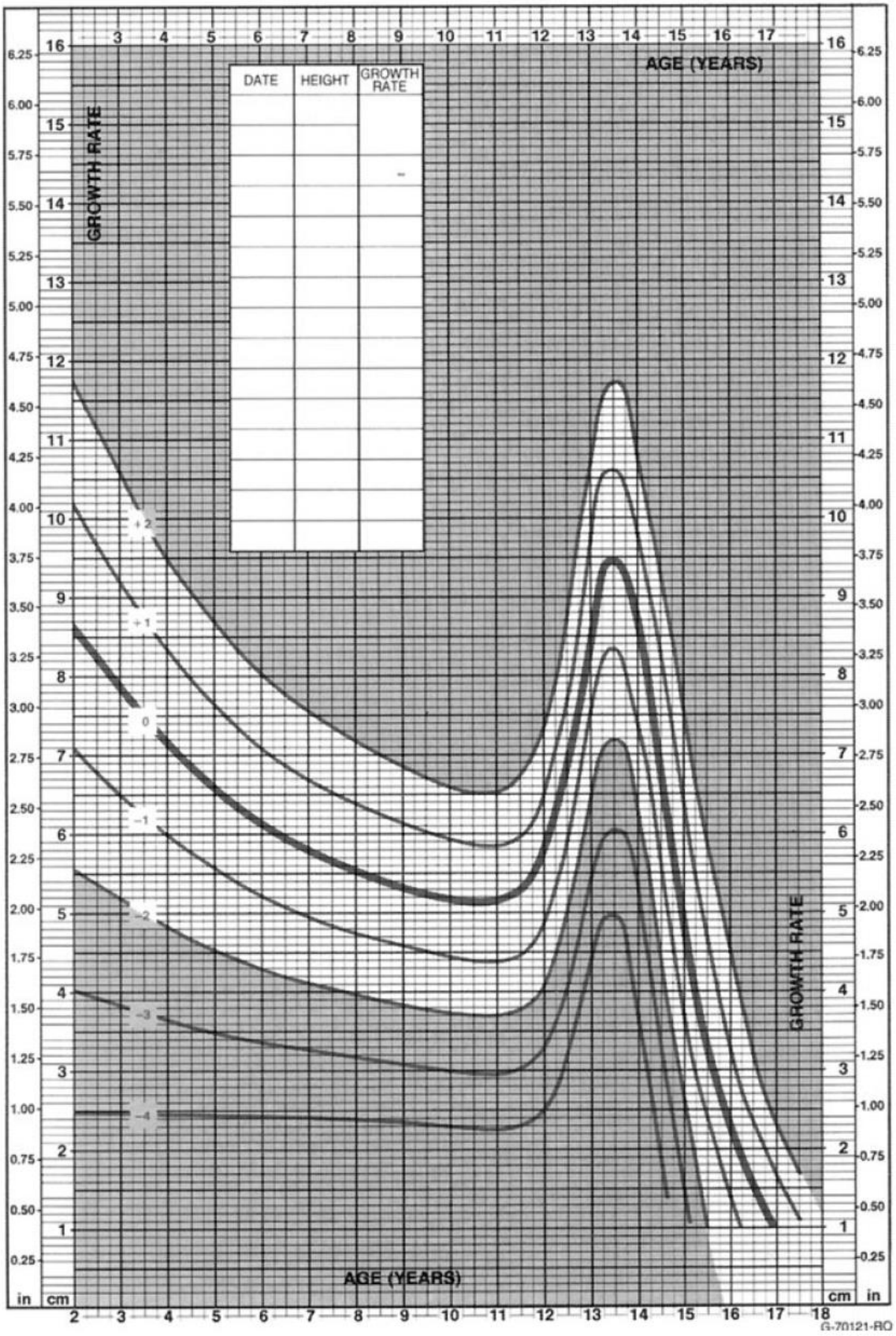

Figure 3. Mean and standard deviations of stature growth velocities in boys aged 2 to 18 years. Note: From Abbassi, V. (1998). Growth and normal puberty. Reproduced with permission from Journal Pediatrics, Vol. 102, Page(s) 507-511, Copyright © 1998 by the AAP.

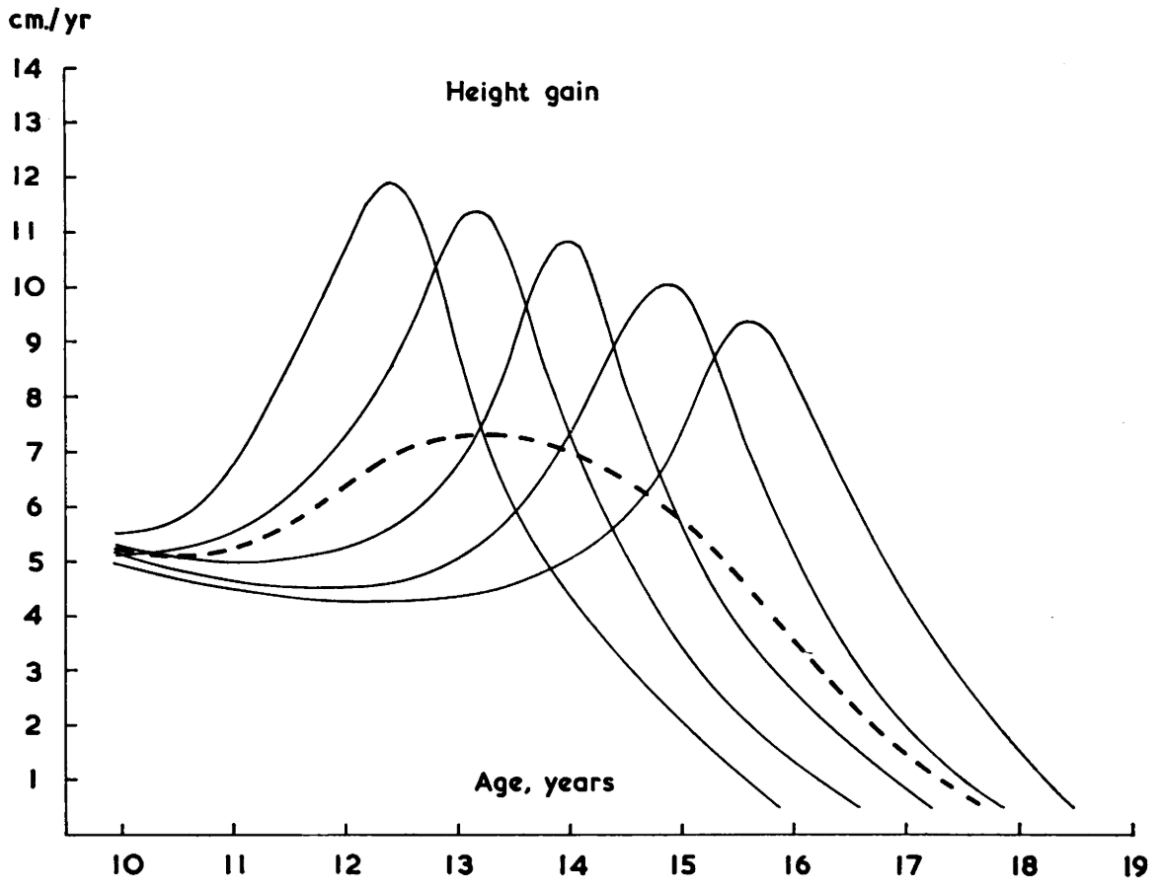

Figure 4. Example of individual and average height velocity growth curves. Note: solid lines represent individual growth curves, dashed line the average growth velocity curve by averaging their values at each age. Reproduced from Standards from birth to maturity for height, weight, height velocity, and weight velocity: British children, 1965. I., Tanner, J. M., Whitehouse, R. H., and Takaishi, M., 41, 454-471, 1966 with permission from BMJ Publishing Group Ltd.

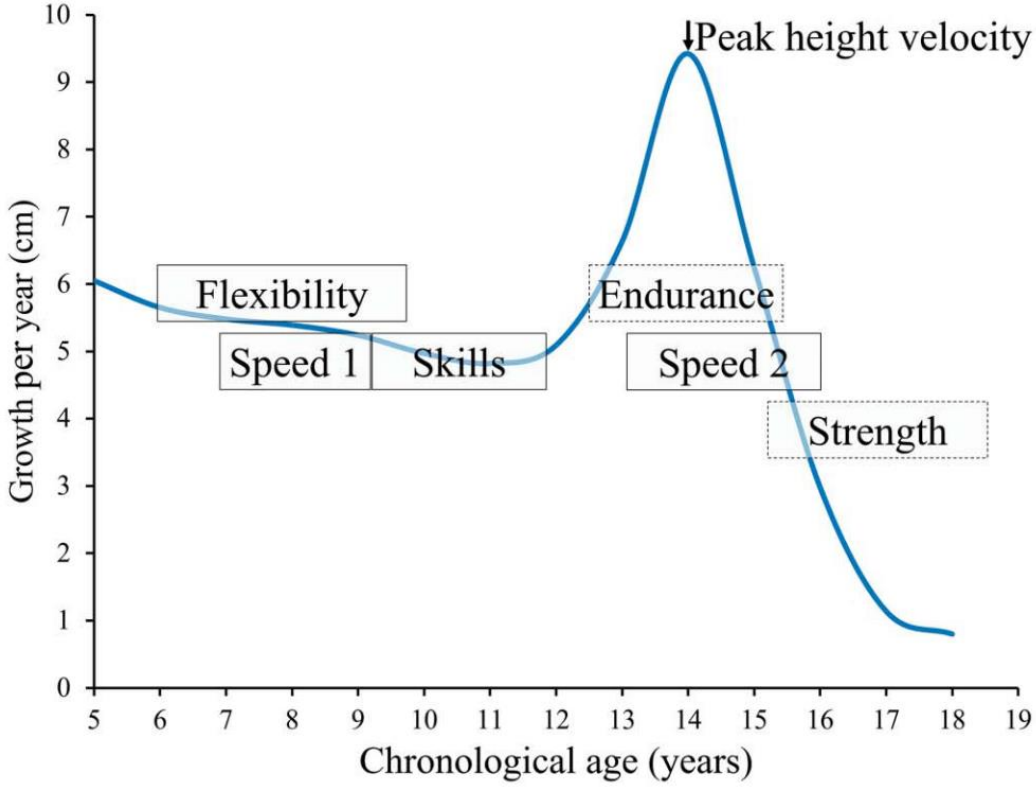

Figure 5. Sensitive periods to training according to the long-term athletic development model in boys. Note: From Van Hooren, B., & De Ste Croix, M., Sensitive Periods to Train General Motor Abilities in Children and Adolescents., *Strength and Conditioning Journal*, 42(6), 7-14. https://doi.org/10.1519/SSC.

As players mature, the overall risk of injury appears to increase with age and several distinctions have been observed in the injury risk profile of pre-, circa- and post-PHV players (Light et al., 2021; Materne et al., 2020; Monasterio et al., 2021b; Read et al., 2018b; Wik et al., 2020a). Specifically, the incidence of muscle, ligament and tendon injuries in male footballers increase into adulthood and with age (Light et al., 2021; Materne et al., 2020; Monasterio et al., 2021b; Wik et al., 2020a). Therefore, an individual's injury risk susceptibility may only become apparent once they are matured and understanding how genetic variants associated with injury vary with age and maturation may be important to appropriately apply the information. The majority (34-72%) of injuries in football occur via non-contact mechanisms (Faude et al., 2013; Materne et al., 2020; Price et al., 2004; Read et al., 2018b; Renshaw & Goodwin, 2016; Rumpf & Cronin, 2012; Wik et al., 2020a), which are considered to be relatively preventable with appropriate progression in training and competition load (Bowen et al., 2017, 2020; Kalkhoven et al., 2021; Nielsen et al., 2018). Injury may occur acutely from an impact force or without a single identifiable event, due to repeated micro-trauma with insufficient recovery (Bennell et al., 1999; Fuller, 2006; Meeuwisse, 1994). Therefore, the risk of musculoskeletal injury is dependent upon the dynamic interaction between an individual's predisposition at that time and environmental exposures (Kalkhoven et al., 2021; Meeuwisse et al., 2007; Nielsen et al., 2018). Indeed, spikes in injury incidence are observed in all age groups around September and January, as players return from breaks or following periods of increased musculoskeletal loading due to match congestion (Le Gall et al., 2006; Price et al., 2004; Read et al., 2018b; Rumpf & Cronin, 2012). Physical training is designed to overload a targeted physiological system and stimulate a specific adaptive response with the aim of performance enhancement (Cunanan et al., 2018). A temporal reduction in performance resulting from fatiguing physical training is followed, given sufficient recovery time, by a supercompensation of performance above the pre-trained state, as conceptually outlined in Figures 6 (Cunanan et al., 2018) and 7. However, if insufficient recovery or tissue adaptation is achieved between loading exposures, tissue damage can accumulate resulting in overuse injury or increased risk of acute injury due to tissue weakness (DiFiori et al., 2014; Kalkhoven et al., 2021; Nielsen et al., 2018). The adaptive response of an individual is mediated by the activation of genetic processes specific to the training stimulus (Cunanan et al., 2018; Wackerhage, 2014). Therefore, genetic variations affecting the adaptive response and / or mechanical integrity of muscle, tendon, ligament and bone tissues will mediate the risk of injury in football and interindividual differences in recovery, magnitude of adaptation, and tissue strength, have all been associated with genetic variants (Baumert et al., 2016; Ma et al., 2013; Strandberg et al., 2003).

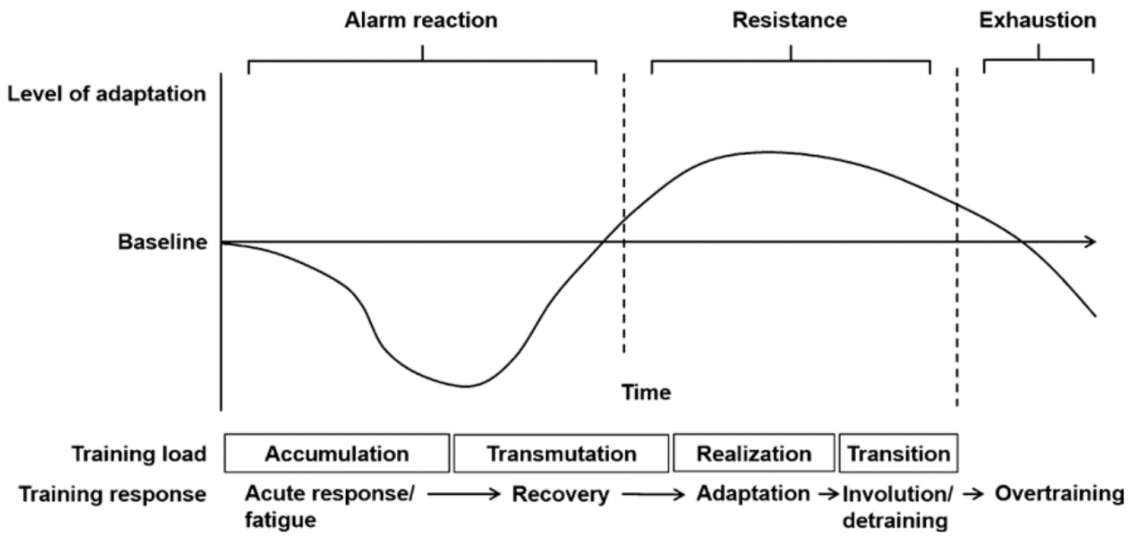

Figure 6. The General Adaptive Syndrome response to training stimulus. Note: Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature Sports Medicine (The General Adaptation Syndrome: A Foundation for the Concept of Periodization, Cunanan, A. J., DeWeese, B. H., Wagle, J. P., Carroll, K. M., Sausaman, R., Hornsby, W. G., Haff, G. G., Triplett, N. T., Pierce, K. C., & Stone, M. H.), Springer Nature and Copyright Clearance Center (2018). https://doi.org/10.1007/s40279-017-0855-3.

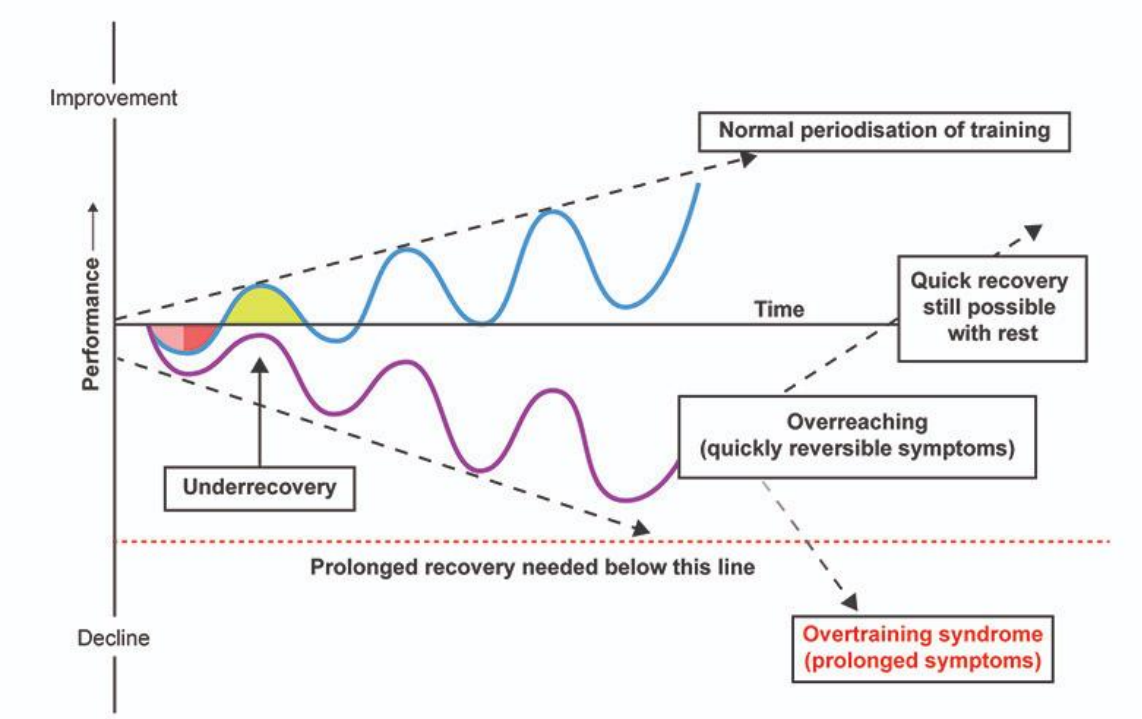

Figure 7. Functional and non-functional overreaching leading to overtraining. Note: From learning from sport burnout and overtraining. An OpenLearn chunk used/reworked by permission of The Open University copyright © (2022).

Football is a predominantly aerobic sport, which is intermittent in nature with short accelerations, tackles, jumps and shots fuelled by anaerobic processes often determine performance success (Bangsbo et al., 2006; Mohr et al., 2003). Consequently, high levels of aerobic fitness are associated with football performance (Waldron & Murphy, 2013; Williams & Reilly, 2000) and professional senior players cover on average 9-12 km, including 1.5-3.3 km of high-speed running,

in a 90-min match (Bradley et al., 2009; Mohr et al., 2003). The ability to perform high-intensity exercise is thought to differentiate between playing level in both youth and adult players (Mohr et al., 2003; Waldron & Murphy, 2013). Senior international players cover ~28% more high-intensity (>18.0 km/h or >5 m/s) and 58% more sprint-intensity (>30.0 km/h or >8.3 m/s) distance than lower standard professional players (Mohr et al., 2003). Similar results are observed in youth players with total distance (~9%), high-intensity distance (~21%), and 10-30m (~16%) sprint time performance all better in elite than sub-elite under-14 players (Waldron & Murphy, 2013). The distance covered during elite youth football matches gradually increases towards the senior demands from the Under-11 (~5.8 km) to Under-15 (~7.7-8.1 km) and Under-18 (~8.8 km) age groups as players physically develop and mature (Buchheit et al., 2010a; Goto et al., 2015). Aerobic performance appears similar between elite and non-elite players at the beginning of puberty but by the end of puberty elite players show significantly greater aerobic fitness than non-elite peers (Strøyer et al., 2004), which also appears to translate into greater distance covered and ball interactions in under-14 players (Waldron & Murphy, 2013). The physical performance developments associated with growth and maturation are also indicated with a steady increase in maximal sprint velocity, jump, acceleration, and repeated sprint ability of football players, with each age group (Atan et al., 2016; Buchheit et al., 2010a, 2010b; Deprez et al., 2015; Malina et al., 2004b; Mendez-Villanueva et al., 2011; Philippaerts et al., 2006). The physiological demands of football vary significantly depending on playing formation and position (Bradley et al., 2009, 2011; Bush et al., 2015; Deprez et al., 2015; Tierney et al., 2016) but distance covered at different speeds provide valuable insight into the physical demands and exposure on players from football training and competition.

Elements of external training load (operationally defined as intensity x volume) experienced by players can be quantified using training time (~ volume), and distance covered at different speeds (~ intensity zones) using microtechnology systems (Gabbett, 2016; Hulin et al., 2014). Strength and conditioning coaches and sports scientists aim to ensure footballers are sufficiently prepared for the physical demands of competitive games to reduce injury risk and support successful performance (Bowen et al., 2017; Gabbett, 2018). Therefore, because targeted overload is necessary for adaptation and physical performance enhancement, but excessive overload may lead to injury, safe increments of training have been investigated (Bowen et al., 2020; Gabbett, 2018; Hulin et al., 2014, 2016). The acute to chronic workload ratio (ACWR) has been suggested as a potential tool to examine the relationship between acute exposure of the current period relative to that of the chronic accumulated workload in the recent past (Hulin et al., 2014). A high chronic workload appears to reduce injury risk as athletes are conditioned to tolerate the competitive and training demands of the sport (Bowen et al., 2017). However, incremental overload must be achieved gradually (Bowen et al., 2017), with acute exposure between 0.8 – 1.3 of the chronic workload suggested to be the ‘sweet spot’ ratio to minimise injury risk (Gabbett, 2018). The validity of this ‘sweet spot’ ratio has recently been questioned due to limitations of the original figure and implications of mathematical coupling when the acute period is included in the chronic calculation (Impellizzeri et al., 2019, 2020; Lolli et al., 2019). Logically, it is unclear how an ACWR < 0.8 would increase injury risk as the exposure of the current period is lower than that previously experienced. Furthermore, the upper tolerance limit may differ depending on individual differences in physical fitness (Buchheit, 2017; Cook & Docking, 2015) and potentially genetics. Although issues may exist in the use of ACWRs, the fundamental principle of progressive overload aligns with current considerations on injury occurrence, as

determined by loading of musculoskeletal tissue (acute exposure) that exceeds the threshold (chronic tolerance) of an individual at that time (Kalkhoven et al., 2021; Meeuwisse et al., 2007; Nielsen et al., 2018). Therefore, the influence of genetic variants, which may affect the inherent tissue loading capacity of an individual, likely interact with the loading exposure of football to influence injury risk. Despite the criticisms of using ACWRs for monitoring workload (Impellizzeri et al., 2019, 2020; Lolli et al., 2019), few alternatives have been proposed (Wang et al., 2020). Wang et al. (2020) highlight the limitation of using a simple ratio to represent meaningful changes in load and the need to explore alternative methods of evaluating the relationship between loading exposure and injury risk.

Qualitative analysis of athletes and coaches perpetual understanding of injuries shows that hereditary factors are already considered to affect injury susceptibility (van Wilgen & Verhagen, 2012; Varley et al., 2017). Despite being frequently acknowledged the genetic contribution of injury risk in elite sport is poorly understood. A large body of work shows that exercise-based injury prevention programmes are beneficial and can result in a significant reduction in injury (Faude et al., 2017; Roessler et al., 2014; Soomro et al., 2016). Furthermore, professional athletes and support staff appear willing to engage in genetic testing to support the individualisation of training programmes to support performance and reduce injury risk (Varley et al., 2017). Therefore, a greater understanding of an individual's genetic susceptibility to injury may allow targeted prevention strategies to be implemented to mitigate risk and prevent potential future elite players being released during adolescence due to injury. Ensuring that players remain injury free and continue to progress during these important developmental years could help to maintain some of the potential elite performers of the future who could become de-motivated and released from the elite youth development system if not supported correctly (Helsen et al., 1998; Musch & Grondin, 2001). An integrated understanding of genetics, physical exposure, growth and maturation may improve the development of elite male footballers by reducing injury risk and supporting opportunities for players to achieve their potential. Muscle, bone, ligament and tendon injuries are particularly common and potentially disruptive in elite male football development. Therefore, genetic variants which influence the strength or physiological regulation of these tissues, and their mechanical properties, are candidates for investigation of applicability to mediate injury risk and support development in elite male football.

2.3 Candidate genes

There are several types of genetic variant that can alter the DNA sequence with potentially substantial consequences on biological function (Tabor et al., 2002). Even one SNP in a gene which encodes for an important regulatory system can have far reaching effects on whole-body physiology to significantly influence musculoskeletal injury risk (Romero et al., 2002). A polymorphism located in a protein coding region of a gene may be synonymous, resulting in a change to the codon without a change to the specific amino acid sequence, or non-synonymous, resulting in an altered codon signal, which may alter the structure and function of the resultant protein (Chakravarti, 1999; Romero et al., 2002; Tabor et al., 2002). Noncoding polymorphisms can occur; in promoter regions which control transcription rates; within introns which are transcribed but not present in fully matured messenger RNA (mRNA); or in 5´- and 3´- untranslated regions (UTR) present in mature mRNA but

which do not code for protein (Romero et al., 2002). Although, the base pair sequence of a gene is unchanged by noncoding polymorphisms the rate of transcription, processing, and stability of mRNA encoding the genetic product may be drastically altered, thus affecting the quantity of protein generated during translation and its physiological activity in the body (Romero et al., 2002).

Genetic variants that demonstrate complete genetic penetrance with a particular trait and have clear mendelian inheritance patterns can be identified with relative ease (Zondervan & Cardon, 2007). In such cases, the functional consequences of a genetic variant alone are sufficient to result in expression of the trait (Zondervan & Cardon, 2007). However, complex traits such as injury risk result from the cumulative influence of numerous genetic variants (Dionne et al., 1991). Consequently, identifying genetic variants that affect complex traits are much more challenging, as numerus genetic and environmental factors interact to influence the interindividual variability of the trait (Zondervan & Cardon, 2007). Genetic case-control association study designs can be utilised to examine the relationship between genetic variants and a particular complex trait (Romero et al., 2002; Zondervan & Cardon, 2007). Genetic case-control studies can be hypothesis-driven candidate gene association studies or genome wide association studies (GWAS) conducted without prior hypothesis (Zondervan & Cardon, 2007). The selection of genetic variants investigated in candidate gene association studies must be underpinned by a biologically valid rationale and / or supported by genomic positional linkage data (Romero et al., 2002). Therefore, the following genes have been included as candidates for association based on the potential physiological influence, they may exert on the structural integrity of tissues which are frequently injured in elite male football.

2.3.1 Structural genes

Genes that encode for the physical constituents of structural tissue in the body are natural candidates for investigation with injury risk. Therefore, SNPs in genes which encode for constituent proteins of structural tissues such as bone, ligament, tendon and muscle have been targeted as candidate genes in genetic association studies of musculoskeletal injury (Collins & Raleigh, 2009; Mann & Ralston, 2003). The collage type I α1 chain (*COL1A1*), collage type I α2 chain (*COL1A2*) and collage type V α1 chain (*COL5A1*) genes all encode for procollagen sub-units, which form and / or organise the fundamental structure of bone, ligament and tendon (Birk et al., 1990; Ghosh, 2002; Myllyharju & Kivirikko, 2001; Tzaphlidou, 2008). Additionally, a variant in the α-actinin-3 (*ACTN3*) gene has been frequently associated with sport and exercise related traits due to differences in the observed mechanical properties of muscle (North et al., 1999).

2.3.1.1 COL1A1 (rs1800012)

Collagens are fibrous proteins, which provide mechanical integrity as major constituents of the extracellular matrix and connective tissues (Kadler et al., 2007). Type I collagen is the main structural protein component of bone, ligament and tendon tissues, forming approximately 60-80% of tendon dry mass (Collins & Raleigh, 2009). The most abundant form of type 1 pro-collagen is composed of one COL1A2 and two COL1A1 sub-units, to form a heterotrimer, encoded for by the *COL1A2* and *COL1A1* genes, respectively (Ghosh, 2002; Myllyharju & Kivirikko, 2001). A noncoding G to T SNP

within the promoter region of the first intron of *COL1A1* (rs1800012) is associated with greater Sp1 transcription factor binding affinity and increased production of COL1A1 (Mann et al., 2001). The relative increase in COL1A1 polypeptide associated with the *COL1A1* (rs1800012) T allele results in greater formation of type 1 pro-collagen formed exclusively from COL1A1 polypeptides (Mann et al., 2001).

Type 1 collagen formed with a high proportion of three COL1A1 pro-collagen subunits has been suggested to be weaker than the COL1A1 / COL1A2 combination as the T allele has been repeatedly associated with increased risk of osteoporotic fracture in the elderly (Mann et al., 2001; Mann & Ralston, 2003). However, although the COL1A1 homotrimer has been associated with impaired mechanical strength, the mechanism remains unclear and the T allele has also been suggested to be protective against anterior cruciate ligament (ACL) ruptures in young physically active participants (Ficek et al., 2013; Khoschnau et al., 2008; Posthumus et al., 2009a). Nevertheless, these results are inconsistent as others have associated the T allele with increased risk of ACL rupture (Stępień-Słodkowska et al., 2013) and bone fractures in pre-pubertal children (Blades et al., 2010). Others have found no association between the *COL1A1* (rs1800012) T allele and fracture risk (Cosman et al., 2013; Korvala et al., 2010; Varley et al., 2018) nor tendon injuries (Erduran et al., 2014; Posthumus et al., 2009c) in physically active adults.

The influence of the *COL1A1* (rs1800012) SNP on injury risk remains unclear. Nevertheless, the physiological consequence and previous genetic association findings of this SNP indicate that it warrants inclusion in candidate gene association studies for injury risk. The genetic penetrance of this SNP appears to be influenced by age as differences are observed in the influence of the T allele between pre-pubertal, adult and post-menopausal individuals (Blades et al., 2010; Korvala et al., 2010; Mann & Ralston, 2003; Suuriniemi et al., 2006). Some have hypothesised that the effect of the T allele may result from differences in recovery (Baumert et al., 2016) and T allele carriers were weaker and reported greater muscle soreness following exercise induced muscle damage compared with G allele carriers (Baumert et al., 2018).

2.3.1.2 COL1A2 (rs412777)

A synonymous A to C SNP in exon 25 of *COL1A2* (rs412777) has also been associated with fracture risk and BMD in children (Blades et al., 2010; Suuriniemi et al., 2003). Interestingly, Blades et al. (2010) found that the risk of fracture was roughly halved for CC homozygotes whilst Suuriniemi et al. (2003) found that possession of one C allele resulted in a four-fold increase in fracture risk. Blades et al. (2010) discussed that the inconsistency of findings may result from variations in genetic background between the two studies or the potential that different SNPs were identified and mistaken using the PvuII restriction site. Nevertheless, Blades et al. (2010) also observed greater BMD in CC homozygotes which would support the hypothesis of a protective effect of the C allele for fracture risk and no difference in BMD was found between genotypes by Suuriniemi et al. (2003).

The *COL1A2* and *COL1A1* genes are located on chromosomes 7 and 17 respectively. Therefore, as expected, the *COL1A2* (rs412777) and *COL1A1* (rs1800012) SNPs are in linkage equilibrium and demonstrate random inheritance patterns across the population. Consequently,

these variants may present a cumulative influence on the genetic susceptibility to fracture injury in children and Blades et al. (2010) found that fracture risk for those who possessed both the *COL1A1* GG and *COL1A2* CC genotypes was more than halve that of other genotype combinations (OR = 0.39, 95% CI = 0.20–0.77, p = 0.008). Additionally, these individuals consistently displayed some of the highest bone quality measures although this did not reach statistical significance (Blades et al., 2010). Despite limited and contradictory associations of the *COL1A2* (rs412777) SNP with fracture risk the magnitude of reported effects has been large, and replication attempts to clarify the influence of this SNP in combination with the *COL1A1* (rs1800012) appear warranted, especially in youth athletes.

2.3.1.3 COL5A1 (rs12722)

The *COL5A1* gene encodes for the pro-α1 chain of type V collagen and has been more extensively studied than many others in relation to injury risk. Type V collagen co-assembles with type I collagen to regulate the diameter of heterotypic type I / V fibrils in tendon, ligament and bone and plays a significant role in the fibrilogenesis of developing connective tissue (Birk et al., 1990; Wenstrup et al., 2004). A noncoding C>T SNP within the 3' UTR of *COL5A1* (rs12722) causes transcription of more stable mRNA (Laguette et al., 2011). This suggests that more COL5A1 is produced from the *COL5A1* (rs12722) T allele than the C allele. The diameter of type I collagen fibres appears to decrease with an increased abundance of type V collagen (Birk et al., 1990). Therefore, the *COL5A1* (rs12722) C allele has been repeatedly associated with a reduced risk of tendon (Altinisik et al., 2015; Mokone et al., 2006; September et al., 2009) and ligament injury (O'Connell et al., 2015; Posthumus et al., 2009b). Altinisik et al. (2015) found the C allele was significantly higher in non-injured controls than patients suffering from tennis elbow and concluded that the CC genotype was protective against chronic degenerative tennis elbow similar to findings with Achilles tendinopathy by Mokone et al. (2006) and September et al. (2009). Additionally, the C allele appears to interact with other genes that regulate the extracellular matrix to modify the risk of Achilles tendon pathology (Brown et al., 2016). Although a GWAS investigating genetic associations with Achilles tendon injury, ACL tears and tendinopathy showed no association with the *COL5A1* gene (Kim et al., 2017b) the protective effect of the C allele on tendon and ligament injury was confirmed in a recent meta-analysis (Pabalan et al., 2018).

Despite the observed relationship between type V collagen abundance and type I collagen fibre diameter (Birk et al., 1990) a direct causal mechanism for the protective effect of the C allele has yet to be established in vivo. A study investigating this potential mechanism found that *COL5A1* (rs12722) C homozygotes had more extensible tendons than TT+CT individuals (Kubo et al., 2013). However, Foster et al. (2014) were unable to replicate the findings of Kubo et al. (2013) and found that the mechanical properties and dimensions of tendon were not influenced by the *COL5A1* (rs12722) SNP in vivo. Nevertheless, those with the C allele have been repeatedly associated with increased flexibility measures with differences between genotypes increasing with age (Brown et al., 2011b; Collins et al., 2009). Alternative findings have indicated that the CT genotype is detrimental to mobility, resulting in lower functional movement scores in adolescent team sport athletes when compared with TT homozygotes (Stastny et al., 2019). However, a higher proportion of *COL5A1*

(rs12722) C allele carriers has been observed in elite male rugby athletes compared to non-athlete controls (Heffernan et al., 2017). The authors hypothesised that this occurred due to innate injury resistance for C allele carriers conferring benefit to the high-risk environment of elite rugby. This protective effect was not observed in ballet dancers (Kim et al. 2014b) but the TT genotype of the *COL5A1* (rs12722) SNP was associated with increased muscle injury severity compared with the CT+CC combined group of professional male footballers (Massidda et al., 2015a). A similar tendency for greater injury severity in CT individuals was observed in another group of elite male footballers, which likely failed to reach statistical significance because the analysis was underpowered as none of the seventy-three participants possessed the TT genotype (Pruna et al., 2013). The T allele has also been associated with increased muscle soreness and extended recovery following exercise induced muscle damage when compared with C allele carriers (Baumert et al., 2016) and the SNP appears to impact the influence of muscle injuries in team sports which could be affected by loading exposure.

In addition to rugby player status, the *COL5A1* (rs12722) SNP has also been associated with differences in athletic performance (Brown et al., 2011a; Posthumus et al., 2011b). The CC genotype has been associated with a reduced risk of self-reported exercise-induced muscle cramping in endurance athletes (O'Connell et al., 2013). Alternatively, TT homozygotes showed significantly faster 42.2km run times than CC individuals in an ironman triathlon (Posthumus et al., 2011b) and 56km ultra-marathon (Brown et al., 2011a). Consequently, it has been suggested that the previously observed associations between CC homozygotes and reduced muscle cramping may result from a reduced endurance running capacity, thus reduced muscle exertion / fatigue inducing capability, rather than being directly protective (Pickering & Kiely, 2017a). The T allele was hypothesised to affect the elasticity of tendons and be advantageous to endurance performance via improvements in running economy (Posthumus et al., 2011b). However, this was not supported by subsequent work which found no significant differences in range of motion or running economy at two different running speeds (Bertuzzi et al., 2014). Therefore, although results vary and the mechanism remains unclear, the *COL5A1* (rs12722) SNP appears to play a complex role in both injury risk and athletic performance. Muscle, tendon and ligament injuries are common in football and understanding the potential influence of the *COL5A1* (rs12722) SNP on injury risk may be valuable to inform injury prevention strategies.

2.3.1.4 ACTN3 (rs1815739)

The *ACTN3* gene has also been frequently associated with sport and exercise related traits resulting from differences in the mechanical properties of muscle (North et al., 1999). A common C to T non-synonymous nonsense coding SNP in exon 16 of the *ACTN3* gene (p.R577X, rs1815739) results in a premature stop codon (X), and non-functional ACTN3, in place of the functional ACTN3 producing arginine codon (R) (North et al., 1999). The ACTN3 protein is almost exclusively found as a constituent of the Z-lines in fast twitch, type II, muscle fibre sarcomeres (Fridén & Lieber, 2001). As the Z-line is a common site of damage during unaccustomed eccentric exercise (Fridén & Lieber, 2001), R allele carriers, with functional ACTN3, are thought to possesses more robust sarcomeres (MacArthur & North, 2004). Furthermore, the R allele displays enhanced adaptive signalling and

force transmission than the X allele (Vincent et al., 2010). Consequently, the R allele has repeatedly been associated with improved strength and power performance (Alfred et al., 2011; Eynon et al., 2009; Kikuchi et al., 2013; Ma et al., 2013; MacArthur & North, 2007; Roth et al., 2008) and more recently protection against injury (Kim et al., 2014b; Massidda et al., 2017; Qi et al., 2016; Shang et al., 2015). Nevertheless, the X allele is not explicitly detrimental to athletic potential and has been associated with improved endurance performance and trainability (Seto et al., 2013; Silva et al., 2015) and has been observed in elite power-oriented sport athletes (Ginevičienė et al., 2016; Ruiz et al., 2013; Sessa et al., 2011).

The baseline strength of untrained R allele homozygotes appears greater than XX counterparts (Broos et al., 2015; Clarkson et al., 2005; Erskine et al., 2014). However, some have only observed this relationship in females but not males (Walsh et al., 2008) or not at all in trained individuals of both sexes (Ginevičienė et al., 2016; Hanson et al., 2010; Norman et al., 2009). These results suggest that the baseline differences in strength associated with the *ACTN3* (rs1815739) R allele can be overcome with training. Nevertheless, the R allele was associated with superior sprint and jump performance in elite strength and power athletes without significant differences in weightlifting performance (Ben-Zaken et al., 2019; Broos et al., 2015; Kim et al., 2014a). The authors suggest this indicates the *ACTN3* (rs1815739) SNP influences speed related exercise and not just strength. However, Norman et al. (2009) found no significant differences in peak power, mean power, torque-velocity or fatigability between the *ACTN3* (rs1815739) genotypes in moderately trained individuals during 30-s Wingate cycling tests. The RR individuals did display a significantly greater increase in peak torque compared with X allele homozygotes following repeated exercise bouts suggesting that the *ACTN3* genotype may influence the responsiveness to training (Norman et al., 2009).

Norman et al. (2009) al so found that *ACTN2* expression is affected by ACTN3 content, which may explain how the body could compensate for the ACTN3 deficiency associated with the X allele. A follow-up study found that sprint exercise resulted in increased glycogen depletion and hypertrophic signalling for R allele carriers (Norman et al., 2014). Norman et al. (2014) concluded that the observed differences in muscle mass and glycogen utilisation provided a mechanistic explanation for the influence of the *ACTN3* (rs1815739) genotype on human performance. These findings are supported by *Actn3* knockout mice models which show higher activity of key oxidative enzymes in fast twitch muscle fibres (MacArthur et al., 2007), which could enhance the oxidative capacity of type II muscle fibres during repeated and moderate-intensity muscle contractions (Del Coso et al., 2019). However, this potential performance advantage associated with the X allele is not reflected by an increased prevalence in endurance athletes (Grealy et al., 2013; Guilherme et al., 2018; Lucia et al., 2006; Papadimitriou et al., 2018; Saunders et al., 2007) and some have even shown a reduced frequency of the X allele (Ahmetov et al., 2010; Kikuchi et al., 2016; Li et al., 2017). Nevertheless, the prevalence of the X allele does not appear significantly underrepresented in elite power athletes (Ginevičienė et al., 2016; Ruiz et al., 2013; Sessa et al., 2011; Wang et al., 2013). Pickering and Kiely (2017) suggest that the X allele of the *ACTN3* (rs1815739) SNP may influence endurance trainability by improving the aerobic metabolism and physiological properties of type II fibres as shown in *Actn3* knockout mice (Seto et al., 2013). However, studies in untrained humans found that although XX individuals had greater baseline endurance performance, RR individuals

demonstrated more pronounced improvements in maximal oxygen uptake following training (Silva et al., 2015). Nevertheless, others have found no significant differences in aerobic capacity improvements over a 5-year period in cross country skiers (Mägi et al., 2016).

Norman et al. (2014) also found that RR homozygotes had greater activation of hypertrophic transcription factors than XX individuals following sprint exercise. This finding is supported by others who have shown that R allele carriers have larger type II fibre cross-sectional area (Broos et al., 2016), which also increased more following training (Gentil et al., 2011). Nevertheless, research in mice has indicated that *Actn3* deficiency results in higher calcineurin activity (Seto et al., 2013). Calcineurin plays a key role in the selective upregulation of genes specific to type I fibres, enhancing the responsiveness of X allele carriers to endurance training whilst supressing the response to strength and power training because of the inhibition of calcineurin on slow-to-fast fibre transformation (Chin et al., 1998; Garton et al., 2014; Seto et al., 2013). Indeed, Delmonico et al. (2007) found, contrary to others, that XX females between the ages of 50-85 had greater peak power performance than female R allele carriers, with no differences observed between the genotypes of males. However, following 10 weeks of training, females with the RR genotype showed greater improvement in peak power than XX females (Delmonico et al., 2007). Variations in recovery or resistance to damage following exercise are likely to alter the risk of musculoskeletal injury by mediating mechanical disruption of the affected fibres (Del Coso et al., 2019).

The R allele of the *ACTN3* (rs1815739) SNP has been associated with reduced exercise induced-muscle damage (Belli et al., 2017; Pimenta et al., 2012; Vincent et al., 2010) and / or enhanced recovery from such exercise (Baumert et al., 2016). Exercise induced muscle damage causes an inflammatory response and leakage of proteins such as creatine kinase into the blood. Vincent et al. (2010) explored the differences in gene expression and muscle damage following a single bout of eccentric exercise and found that RR individuals tended to have lower blood CK and muscle pain scores, but greater anabolic gene expression. These results have also been replicated in football players (Pimenta et al., 2012) and endurance athletes (Belli et al., 2017). Furthermore, others have shown significant differences in the functional recovery of physical performance with X allele carriers experiencing greater reductions in leg power and jump height than RR homozygotes following a marathon and half-ironman, respectively, along with higher muscle damage and pain (Del Coso et al., 2016, 2017a).

Exercise-induced muscle damage is a complex trait mediated by several genetic variants (Del Coso et al., 2017b) and others have found no difference in muscle damage between *ACTN3* (rs1815739) genotypes (Broos et al., 2019; Clarkson et al., 2005). Additionally, one study found that moderately active RR genotype males showed greater voluntary force decrements and slower recovery than XX counterparts following two muscle damaging exercise bouts (Venckunas et al., 2012). A recent study by Broos et al. (2019) examined the different responses of the RR and XX genotypes to an eccentric exercise bout in non-athletic young men. The number of participants was low (RR n = 4 and XX n = 4) and large interindividual variations were observed in both muscle damage markers, and strength reductions, independent of genotype. However, a significant increase in type II fibre stiffness was observed in RR individuals and the authors suggest that the *ACTN3* (rs1815739) SNP may not affect susceptibility to muscle damage acutely but could prevent subsequent damage in repeated bouts of eccentric exercise (Broos et al., 2019). Greater

understanding of an individual's genetic predisposition to strength, endurance and recovery could guide training prescription and recovery decisions to mediate injury risk. Although inconsistent the R allele appears to be associated with more rapid recovery and greater adaptive signalling following exercise induced muscle damage. Therefore, these individuals may be at less risk of injury and able to undertake more frequent training sessions than X allele carriers.

The risk of injury of XX individuals has been reported to be more than 2.5 times greater for male football players (Massidda et al., 2019) and 4.7 times greater in female ballet dancers (Kim et al., 2014b) than R allele homozygotes. However, evidence supporting an influence of the *ACTN3* (rs1815739) polymorphism on injury risk is inconsistent and another study of female athletes found the R allele increased the risk of muscle strain injury when compared to X allele carriers (Iwao-Koizumi et al., 2014). Iwao-Koizumi et al. (2014) suggested this result may occur due to R allele carriers expressing higher contraction forces during exercise which could cause a greater risk of muscle strain (Iwao-Koizumi et al., 2014). However, the X allele has also been repeatedly associated with increased risk of ankle ligament injury (Qi et al., 2016; Shang et al., 2015) which would suggest that the increased strength associated with R allele improves the ability of the muscle to maintain joint stability during exercise (Del Coso et al., 2019). Nevertheless, X allele carriers appear to have lower BMD (Yang et al., 2011) than R allele carriers indicating that the XX genotype may also increase bone injury risk during exercise (Del Coso et al., 2019). It is possible the observed relationship between the *ACTN3* R allele and BMD is related to the association with increased muscle mass and function causing greater bone loading, and therefore BMD maintenance rather than directly influencing bone metabolism (Pickering & Kiely, 2018). Nevertheless, some have hypothesised that the *ACTN3* (rs1815739) SNP may influence fracture risk but there is currently no genetic association evidence to support such an effect (Del Coso et al., 2019).

Although contradictory findings exist, a growing body of evidence suggests that the *ACTN3* (rs1815739) SNP influences numerous physical performance traits. The *ACTN3* (rs1815739) SNP may influence the risk of musculoskeletal injury directly via an alteration in the tissue-specific threshold tolerance of muscle, bone and ligament tissue. Alternatively, the *ACTN3* (rs1815739) SNP may indirectly influence tissue-specific injury risk via alterations in the recovery and adaptive response to exercise, which may be further mediated by physiological and / or environmental factors. Therefore, the *ACTN3* (rs1815739) SNP appears worthy of inclusion as a candidate for investigation in the present research with injury risk in elite male football development.

2.3.2 Hormone receptor genes

Genetic variants can directly influence the mechanical integrity of musculoskeletal tissue via alterations in its structural components. However, physiological tissue in healthy humans is alive and continuously degrading and regenerating (Krabbe et al., 1982; Magra & Maffulli, 2008; Quadrilatero et al., 2011; Rahim et al., 2017). Therefore, genes which affect the homeostasis and regulation of musculoskeletal tissue may also be considered as candidates to influence injury susceptibility. The structure, functionality and sensitivity of hormone receptors can differ due to interindividual genetic variations (Alonso et al., 2011; Danser et al., 1995; Rigat et al., 1990). Hormones and hormone receptors play a vital role in the regulation of the musculoskeletal system (DiGirolamo et al., 2012).

Subsequently, variants within the estrogen receptor 1 (*ESR1*), matrix metallopeptidase 3 (*MMP3*) and angiotensin I converting enzyme (*ACE*) genes have been examined as candidates of influence on injury risk and physical performance (Kumagai et al., 2018; Larruskain et al., 2018; Ma et al., 2013; Puthucheary, 2011; Tang et al., 2013).

2.3.2.1 ESR1 (rs2234693)

A noncoding T to C SNP (rs2234693) in the first intron of *ESR1* is associated with altered expression of the gene (Alonso et al., 2011) and injury risk (Kumagai et al., 2018; Tang et al., 2013). Estrogen is a key regulator of bone metabolism which is involved in bone remodelling, resorption and formation (Khosla et al., 2012). The actions of estrogen primarily occur through the intracellular activation of two principal receptors - ESR1 and ESR2 (Khosla et al., 2012). The C allele of the *ESR1* (rs2234693) SNP results in increased expression of *ESR1*, reduced BMD and greater risk of osteoporotic fracture in postmenopausal women (Alonso et al., 2011; Kurt et al., 2012; Sonoda et al., 2012). However, the increased risk of fracture associated with the C allele has not been replicated in young physically active participants (Cosman et al., 2013; Välimäki et al., 2005). Furthermore, others have even observed that the C allele to be associated with a reduced risk of muscle injury and hamstring stiffness (Kumagai et al., 2018). Currently there is limited replication of the *ESR1* (rs2234693) SNP with injury risk in young physically active participants. However, it may be that this SNP interacts with physiological processes of aging to mediate the risk of fracture and muscle injury.

2.3.2.2 MMP3 (rs679620)

The MMP3 protein stimulates the activity of other metalloproteinases (Toth et al., 2003). Metalloproteinases are involved in the control and integrity of the extra cellular matrix by catalytically degrading structural proteins including various types of collagen (Birkedal-Hansen et al., 1993; Somerville et al., 2003). A T>C missense coding sequence SNP in the *MMP3* (rs679620) gene results in a glutamic acid in place of a lysine codon. The functional consequences of this SNP are unclear, but the C allele has been independently associated with increased risk of Achilles tendinopathy (Gibbon et al., 2016) and appears to interact with the *COL5A1* (rs12722) T allele to further increase injury risk (Raleigh et al., 2009). The results of Raleigh et al. (2009) highlight how both gene-environment and gene-gene interactions can occur to significantly influence predisposition to injury. A clear mechanism to explain this interaction was unclear but the MMP3 molecule appears to interact with type V collagen as part of its role in regulating the extra cellular matrix (Birkedal-Hansen et al., 1993; Raleigh et al., 2009). Therefore, an interaction effect may occur if the *MMP3* (rs679620) C and *COL5A1* (rs12722) T alleles result in a direct change in the reactivity or regulation of MMP3 and type V collagen in tendinous tissue. Alternatively, an indirect effect may occur because of differences in tissue extensibility, associated with the *COL5A1* (rs12722) T allele (Kubo et al., 2013), or other unknown influences of the T allele, which interacts with MMP3 activity to increase tendinopathy risk (Raleigh et al., 2009). The C allele was also recently associated with a reduced risk of hamstring injury (Larruskain et al., 2018) and time lost from knee injury in football players (Hall et al., 2022), without any influence on range of motion (Posthumus et al., 2010b). Although there is limited evidence for the effect of the *MMP3* (rs679620) SNP with injury risk functional interactions

with other casual variants may exist. Therefore, understanding how this SNP may interact with other genetic variants, growth / maturation and loading exposure may allow targeted preventative strategies to be adopted to reduce individual injury risk.

2.3.2.3 ACE (rs1799752)

A 287 base pair insertion (I) / deletion (D) polymorphism within intron 16 of the *ACE* gene (rs1799752) is one of the most heavily researched genetic variants in association with differences in physical performance phenotypes (Ma et al., 2013; Puthucheary, 2011). The ACE protein is a key component of the endocrine renin-angiotensin system which regulates blood pressure, electrolyte and fluid homeostasis (Nehme et al., 2019). The ACE molecule supresses vasodilation by degrading vasodilators and forming potent vasoconstrictors like angiotensin II, which controls blood pressure and fluid-electrolyte balance (Nehme et al., 2019). The I allele of the *ACE* I/D (rs1799752) polymorphism results in decreased ACE activity in blood (Rigat et al., 1990) and cardiac tissue (Danser et al., 1995). This in turn is associated with increased exercise induced expression of genes involved in angiogenesis and aerobic metabolism (van Ginkel et al., 2016). Consequently, the II genotype has been repeatedly associated with improved endurance performance (Cieszczyk et al., 2009; Collins et al., 2004; Dionísio et al., 2017; Gayagay et al., 1998; Ma et al., 2013; Montgomery et al., 1998, 1999).

Gayagay et al. (1998) initially observed that the II genotype was overrepresented in elite Australian rowers. This finding was later replicated in elite male mountaineers and associated with a greater improvement in time to fatigue after training (Montgomery et al., 1998). A follow up study also found that male II genotype army recruits experienced greater increases in non-fat mass than D allele carriers following 10 weeks of military training (Montgomery et al., 1999). Others have subsequently found that II homozygotes have a greater proportion of slow twitch muscle fibres (Zhang et al., 2003), higher maximal oxygen uptake (Almeida et al., 2012; Hagberg et al., 1998), greater muscle capillarisation (Vaughan et al., 2016) and aortic distensibility (Tanrıverdi et al., 2005) than DD individuals. Additionally, the II genotype has been overrepresented in the fastest ironman triathlon finishers (Collins et al., 2004), elite rowers (Cieszczyk et al., 2009) and associated with greater YoYo intermittent endurance test performance in male football players (Dionísio et al., 2017).

Despite the association of the I allele with endurance performance, the D allele is not ubiquitously detrimental to athletic performance. Indeed, the DD genotype has been associated with greater strength measures (Colakoglu et al., 2005; Folland et al., 2000) and repeatedly appears to be overrepresented in sprint and power-oriented sports (Costa et al., 2009; Kikuchi et al., 2012; Morucci et al., 2014; Myerson et al., 1999; Nazarov et al., 2001; Papadimitriou et al., 2016; Tsianos et al., 2004; Woods et al., 2001). The relative frequency of the I allele appears to increase with competitive race distance and the D allele is more frequent in sports which are associated with shorter anaerobic activity (Di Cagno et al., 2013; Myerson et al., 1999; Nazarov et al., 2001; Tsianos et al., 2004). This observation of the D allele has been replicated by some with overrepresentation in sprint swimmers (≤ 400 m) (Costa et al., 2009; Woods et al., 2001) and superior performances in short duration aerobic running events (≤ 2000 m) (Cam et al., 2005, 2007) and sprint running (≤ 400 m) (Papadimitriou et al., 2016). It has been hypothesised that the cross-over, where the I allele

becomes more advantageous, occurs for aerobic sports lasting longer than 10-30 minutes as running economy becomes more important to performance success (Cerit et al., 2006). Prior to this point, Cerit et al. (2006) suggest that the increased anaerobic power and maximal oxygen uptake, which has been associated with the D allele (Kasikcioglu et al., 2004; Rankinen et al., 2000a), results in greater success in short to middle distance running performances.

The results of the *ACE* I/D (rs1799752) polymorphism and athletic performance are, however, inconsistent and several studies have found no significant association with elite endurance athletes (Ash et al., 2011; Papadimitriou et al., 2018; Rankinen et al., 2000b; Scott et al., 2005; Taylor et al., 1999) or sprint performance (Scott et al., 2010). Furthermore, two studies have even associated the D allele with marathon performance in elite Japanese (Tobina et al., 2010) and Israeli (Amir et al., 2007) athletes. As mentioned, the DD genotype has also been repeatedly associated with greater maximal oxygen uptake than II individuals (Kasikcioglu et al., 2004; Rankinen et al., 2000a), contrary to previous findings (Hagberg et al., 1998). This finding may be attributable to the increased hypertrophic response of the left ventricular previously observed for DD individuals which would also result in an increase in cardiac output (Di Mauro et al., 2010; Diet et al., 2001; Fatini et al., 2000; Hernández et al., 2003; Kasikcioglu et al., 2004; Montgomery et al., 1997). However, these results are also inconsistent as others have observed no differences in maximal oxygen uptake (Djarova et al., 2013; Orysiak et al., 2013; Roltsch et al., 2002) nor ventricular hypertrophy (Alves et al., 2009; Rizzo et al., 2003) between *ACE* I/D (rs1799752) genotypes.

Several studies have repeatedly found no significant difference in the genotype distribution of the *ACE* I/D (rs1799752) variant between football players and sedentary controls indicating that neither allele is clearly advantageous to performance success (Cięszczyk et al., 2016; Coelho et al., 2016; Fatini et al., 2000; Juffer et al., 2009; Massidda et al., 2012; Micheli et al., 2011). Juffer highlighted this finding et al. (2009) with the ID genotype more, and the II genotype less, prevalent in elite football players than runners but not different to sedentary controls. Micheli et al. (2011) suggested that the *ACE* I/D (rs1799752) genotypes could be used to differentiate individuals with the greatest genetic potential to succeed in football. However, this hypothesis is based solely on an increase in jump performance observed in ID individuals without differences in genotype distribution between elite youth football players and sedentary controls (Micheli et al., 2011). Only one study has found a significant difference in the genotype frequency of the *ACE* I/D (rs1799752) polymorphism in football players and untrained controls (Ginevičienė et al., 2014). Ginevičienė et al. (2014) found that the ID genotype was higher in all football players combined, whilst the DD genotype was lower in defenders and midfielders, than untrained controls. The differences in genotype distributions between positions found by Ginevičienė et al. (2014) may explain why others have not observed significant differences between football athletes overall and controls. Therefore, overall football performance success appears unaffected by the *ACE* I/D (rs1799752) polymorphism, but some genotype profiles may be more suitable to specific playing positions than others.

The local renin-angiotensin system affects cellular growth, proliferation, differentiation, extracellular matrix remodelling and inflammation in specific tissues (Dzau, 1993; Nehme et al., 2019; Ribeiro-Oliveira et al., 2008). The activity of ACE is linked with the exercise induced response to training (van Ginkel et al., 2016) and several authors have explored the influence of the *ACE* I/D (rs1799752) polymorphism on physical performance adaptation. Consequently, the D allele has been

associated with greater increases in strength and left ventricular hypertrophy in male military recruits and football players following training (Colakoglu et al., 2005; Fatini et al., 2000; Folland et al., 2000; Montgomery et al., 1997; Pereira et al., 2013b). However, again these findings appear inconsistent as other have not found any association of the *ACE* I/D polymorphism with strength gains following resistance training in trained (Erskine et al., 2014; Gentil et al., 2012) or untrained (Thomis et al., 2004; Williams et al., 2005) males. The DD genotype has been associated with significantly greater muscle volume (Charbonneau et al., 2008) and free fat mass (Lima et al., 2011) at baseline but the II genotype has been linked with a greater anabolic response (Lima et al., 2011; Montgomery et al., 1999). The I allele has been linked with endurance adaptation in military recruits with greater improvements in reps to fatigue (Montgomery et al., 1998) and muscle contraction efficiency (Williams et al., 2000). Nevertheless, others have found no difference in the endurance training response between *ACE* I/D (rs1799752) genotypes (Alves et al., 2018; Alves et al., 2009). An interesting paper by Pescatello et al. (2006) evaluated whether changes in muscle strength and size differed between *ACE* I/D (rs1799752) genotypes following unilateral resistance training in predominantly Caucasian, healthy young men and women. Baseline isometric maximal voluntary contraction, dynamic one-repetition maximum and muscle cross-sectional area were greater in the dominant untrained than non-dominant trained arm, but this did not differ between genotypes (Pescatello et al., 2006). However, the increase in maximal voluntary contraction of the trained arm in I allele carriers was greater (II/ID = 22%; DD = 17%), but not different for repetition max (51%) or cross-sectional area (19%). Furthermore, D allele carriers showed greater increases in repetition max (DD/ID = 11%; II = 7%) and cross-sectional area (DD = 1.5%; ID =1.7%) while that of II individuals reduced (II = -1%). Changes in maximal voluntary contraction were also greater in the untrained arm for I allele carriers (II/ID = 6.8%; DD = 2.0%) and the authors concluded that the *ACE* ID genotype influences the contralateral effects of unilateral resistance training but could not conclude with any certainty the mechanism of these effects (Pescatello et al., 2006).

As both the *ACE* I/D (rs1799752) and *ACTN3* (rs1815739) variants have shown numerous associations with athletic phenotypes, several authors have evaluated the potential interaction or cumulative influence of these variants together. Consequently, D and R allele carriers of the *ACE* and *ACTN3* variants have shown greater contractile forces than I and X carriers (Erskine et al., 2014) with significant interactions found for peak power, sprint and jump performances (Ahmetov et al., 2013; Dionísio et al., 2017; Eynon et al., 2009; Gómez-Gallego et al., 2009; Papadimitriou et al., 2016). Pereira et al. (2013a) found that the *ACTN3* (rs1815739) and *ACE* I/D (rs1799752) variants both independently, and in combination, influence the response to high-speed power training in 60- to 70-year-old Caucasian women. However, contradictory findings have shown that neither of these genetic variants influence muscle strength and contractile properties (Lima et al., 2011; McCauley et al., 2009; Orysiak et al., 2018; Rodríguez-Romo et al., 2010). Ahmetov et al. (2013) investigated the influence of the *ACE* (rs1799752) and *ACTN3* (rs1815739) variants with strength related traits in physically active middle school Caucasian males and females. No significant differences were observed in females, but male *ACE* (rs1799752) D homozygotes displayed significantly greater standing long jump and hand grip strength (Ahmetov et al., 2013). Furthermore, the grip strength differences became even more pronounced when the *ACE* (rs1799752) DD and *ACTN3* (rs1815739) RR genotypes were combined (Ahmetov et al., 2013).

The observed relationship between the *ACE* I/D (rs1799752) variant with physical performance and the response to exercise has led some to explore the influence on muscle damage (Yamin et al., 2007). Yamin et al. (2007) asked seventy young men and women to complete maximal eccentric contractions of the elbow flexor muscles. The authors found that I allele carriers had significantly greater increases in creatine kinase than DD individuals. The increase in muscle damage markers was stepwise with each addition possession of the I allele and the *ACE* I/D (rs1799752) genotype appeared the most powerful determinant of peak creatine kinase activity when included with age, sex and body mass index (Yamin et al., 2007). Comparable results have been observed following marathon performance with greater muscle damage markers observed in I allele carriers (Sierra et al., 2019). These findings suggest that the I allele represents an increased risk of exercise induced muscle damage and the D allele has been associated with a reduced risk of muscle injury in elite male footballers (Massidda et al., 2020) although no association with injury was observed in elite female ballet dancers (Kim et al. 2014b). Nevertheless, the risk of injury associated with the *ACE* I/D (rs1799752) variant appears specific to muscle injury.

Substantial evidence supports the influence of the *ACE* I/D (rs1799752) variant with athletic performance despite considerable contradictory findings. The *ACE* I/D (rs1799752) variant does not appear to prevent an individual achieving elite football performance but may influence the playing style and position. The *ACE* I/D (rs1799752) variant may interact with other genetic factors associated with athletic performance to mediate the response to training and physical performance traits which could affect individual injury risk when considering contributory influence of loading exposure. Specifically, although current evidence is emerging and remains limited the *ACE* I/D (rs1799752) variant appears to influence the risk of muscle injury via an increased susceptibility to muscle damage during exercise. Consequently, understanding how the *ACE* I/D (rs1799752) variant influences muscle injury in the present study could inform alterations in training exposure between genotypes to protect players from injury in elite male football.

2.3.3 Transcription factor genes

Transcription is the process of generating mRNA from DNA, which is later translated into protein. Regulation at the various stages of gene expression, from transcription through to translation, can influence physiological function by adjusting the amount and type of protein produced. The initiation of this process is regulated, following a training or nutritional stimulus, by transcription factors which upregulate or supress the expression of target genes to produce a specific adaptive response (Cunanan et al., 2018; Wackerhage, 2014). Therefore, interindividual variations in genes which encode for transcriptional regulators have also been implicated with differences in predisposition to injury (Ahmetov et al., 2006; Pickering et al., 2018; Strandberg et al., 2003). It is possible that these variants also interact with other physiological and genetic processes associated with growth, maturation and aging to further influence the individual susceptibility to injury risk. Understanding how the responsiveness of elite male footballers to training may vary at various stages of development could guide interventions to protect players from injury with bespoke programmes considering their individual needs at that time.

2.3.3.1 VDR (rs2228570)

The vitamin D receptor gene (*VDR*) encodes for the nuclear vitamin D3 receptor which stimulates the downstream expression of genes involved in mineral metabolism and immunity (Jurutka et al., 2000). A common C to T missense variant in the initiator codon of the *VDR* gene (rs2228570, previously rs10735810) results in reduced transcriptional potency of the VDR protein (Arai et al., 1997; Jurutka et al., 2000). The T allele has been associated with reduced BMD in healthy Japanese premenopausal (Arai et al., 1997) and Turkish postmenopausal (Kurt et al., 2012) women, in addition to, young male athletes (Nakamura et al., 2002a, 2002b; Strandberg et al., 2003). As BMD is considered an important determinant of bone strength (Fonseca et al., 2014; Hernandez & van der Meulen, 2017), the *VDR* (rs2228570) T allele SNP has also been linked with increased risk of stress fracture in young male (Chatzipapas et al., 2009; Korvala et al., 2010) and female (Varley et al., 2018) military recruits in basic training. This SNP is sometimes referred to as the *VDR* FokI restriction fragment length polymorphism (RFLP) and Tajima et al. (2000) examined the interaction between the *VDR* FokI (rs2228570) genotypes and bone metabolism following resistance exercise training in young Japanese males. Stimulation of bone formation and suppression of resorption occurred for all participants within one month, however, suppression of resorption was greater and longer lasting in C than T allele carriers (Tajima et al., 2000). This finding has been replicated by Rabon-Stith et al. (2005) who found a greater increase in BMD following resistance training in elderly male and female C allele carriers than T homozygotes. Others have also shown that male adolescent athletes with the CC genotype have significantly greater BMD than T allele carriers when exposed to weight bearing exercise (Nakamura et al., 2002a, 2002b; Strandberg et al., 2003). Additionally, T allele carriers appear to have lower serum vitamin D levels compared to CC homozygotes (Tuncel et al., 2019). Nevertheless, some have found directly contradictory results in adolescent football players with higher, and greater increases, in BMD associated with the T allele (Diogenes et al., 2010) and no difference in injury incidence or severity in elite male football players (Massidda et al., 2015b). Furthermore, CC homozygotes have been associated with an increased prevalence of lower back pain in athletes (Cauci et al., 2017) and risk of lumbar spine pathologies (Colombini et al., 2014), although this may be specific to Caucasian ethnicities (Pabalan et al., 2017).

Interestingly, Micheli et al. (2011) also found that the TT genotype was overrepresented in elite young male football players than sedentary controls. Micheli et al. (2011) suggest this result may occur due to previously observed differences in strength between the *VDR* (rs2228570) genotypes. However, the evidence the authors cite for this observation found significant differences in women only which became non-significant after adjustment for confounding (Windelinckx et al., 2007). Furthermore, Micheli et al. (2011) comment that any initial advantage associated with the T allele can be overcome with training as they observed no significant differences in physical performance between the *VDR* gene (rs2228570) genotypes. A finding which is consistent with others (Morucci et al., 2014). Our understanding of the importance of vitamin D and its regulation via VDR in response to different stimulus is increasing. Recent reviews have suggested that increased vitamin D levels may augment athletic performance via improvements in; recovery following training, skeletal muscle function, force and power production (Dahlquist et al., 2015). The *VDR* (rs2228570) SNP appears to have an influence on BMD and the risk of fracture injury. Differences in bone metabolism during adolescence may be more important as bone mass accrual is at its peak, but

bone fragility also appears to increase (Parfitt, 1994; Yilmaz et al., 2005). An impaired ability to regulate bone homeostasis during periods of intense exercise is important factor when considering the risk of bone fracture and overuse apophysitis injuries. Indeed, a transient weakness in bone strength during pubertal growth has been suggested as a potential cause of the increased apophysitis injury during adolescence (Wang et al., 2010). Fracture and apophysitis injuries are prevalent and potentially severe time-loss injuries in elite male youth football (Light et al., 2021; Read et al., 2018b). Therefore, identifying individuals with a genetic susceptibility for apophysitis injuries may help to inform targeted training modifications to mitigate this risk.

2.3.3.2 GDF5 (rs143383)

The pubertal growth spurt is associated with increased risk of apophysitis injury (Read et al., 2018b) and fracture risk (Parfitt, 1994). The growth and differentiation factor 5 gene (*GDF5*) encodes for a transcription factor that promotes the expression of genes responsible for joint tissue development, maintenance and repair (Mikic, 2004). Therefore, like VDR, the GDF5 protein plays an important role is the regulation of tendon, cartilage, ligament and bone (Mikic, 2004). A non-coding G to A intron SNP in the 5'UTR of *GDF5* (rs143383) results in reduced mRNA transcript production (Miyamoto et al., 2007; Southam et al., 2007) and is associated with increased risk of osteoarthritis (Yin & Wang, 2017), meniscal injury (Ge et al., 2014) and stress fracture (Zhao et al., 2016). AA homozygotes of the *GDF5* (rs143383) SNP have also been linked with an increased risk of all recorded injuries in football (McCabe & Collins, 2018), ACL rupture (Chen et al., 2015) and Achilles tendinopathy (Posthumus et al., 2010a). Nevertheless, *in vivo* analyses of the differences in human tendon properties were non-significant between *GDF5* (rs143383) genotypes (Kubo et al., 2013) and others have observed no association with ACL (Raleigh et al., 2013) or muscle injury in elite male footballers (Larruskain et al., 2018; Pruna et al., 2016). However, a recent study in adolescent team sport athletes found that both male and female GG homozygotes had worse reactive strength, and males greater flexibility, scores than A allele carriers (Stastny et al., 2019). Stastny et al. (2019) suggest that the *GDF5* (rs143383) and *COL5A1* (rs12722) SNPs interact to influence the functional physical performance capabilities of adolescents team sport athletes via disruptions to flexibility. The *GDF5* (rs143383) SNP may have particular relevance to the risk of apophysitis and fracture injuries associated during the adolescent growth spurt due to the regulatory role of the GDF5 protein in developing tissue. However, the reduced flexibility and functional capacity also associated with the *GDF5* (rs143383) SNP along with previous numerous other injury associations indicates that this variant may be informative for individual injury susceptibility in elite male football development.

2.4 Genetically individualised athletic development programmes

The individual responsiveness and improvements following a standardised training exposure can vary substantially (Hautala et al., 2006; Hubal et al., 2005). Some have found improvements ranging from 0% to +250% in strength and -2% to +59% changes in muscle cross sectional area following resistance training (Hubal et al., 2005), with similar observations reported in endurance training (Hautala et al., 2006). Consequently, coaches and sports scientists continue to acknowledge

the need to adapt to individual needs (Halson, 2014; Pickering & Kiely, 2019). Indeed, the use of individual workload intensity thresholds for ACWR monitoring have been presented (Abt & Lovell, 2009) and the physical performance and load tolerance demands vary between positions in football (Bradley et al., 2011; Bush et al., 2015; Tierney et al., 2016). Interindividual genetic differences will influence the individual tissue-specific threshold tolerance, adaptability and recovery from exercise (Baumert et al., 2016; Pickering & Kiely, 2019). Therefore, attempts have already been made to introduce genetically informed individual training programmes (Jones et al., 2016; Pickering et al., 2018).

Genetically individualised programmes are generated by calculating a total genotype score (TGS) for endurance or power performance based upon the findings of previous association studies (Ruiz et al., 2009, 2010; Williams & Folland, 2008). The TGS of an individual is then calculated based on their polygenic profile and transformed onto a 0-100 scale based on how many advantageous variants they possess (Williams & Folland, 2008). This calculation appears to originally be proposed in sport and exercise sciences to examine the probability of any individual possessing the optimal TGS in a population (Williams & Folland, 2008). Subsequently, others have examined the validity and application of TGS for injury susceptibility (Goodlin et al., 2015; Hall et al., 2022). Although unclear, the aim of Goodlin et al. (2015) appears to be to observe the participants' responses to their genetic information. However, statistical analysis, reporting and discussion of these results appears to be lacking. Arguably, the study by Goodlin et al. (2015) more closely aligns with a systematic review and meta-analysis of studies which include the variants in the 23andMe direct-to-consumer testing analysis. However, there is a lack of clarity around how search terms were defined, and searches completed, in addition to how screening, quality and risk of bias assessments were performed as recommended by expert guidelines for systematic reviews and meta-analyses (Moher et al., 2009). Consequently, the results of this study are somewhat questionable and need further verification.

A genuine attempt to explore the potential cumulative influence of candidate genes on musculoskeletal injury incidence found that elite male youth footballers who became injured during the observation period had a significantly higher TGS than uninjured players (Hall et al., 2022). This is a promising observation but understanding how tissue-specific TGSs may influence injury risk in football development could inform the design of training or nutritional interventions to support long term development. Hall et al. (2022) also examined how maturation influences the effect of genetic variants on injury incidence. The authors found that, along with other maturation specific genetic associations, all soft tissue and ligament injuries were more prevalent in Pre-PHV *COL5A1* (rs12722) CC individuals than Pre-PHV T allele carriers. This finding is in contrast to that observed in adults for whom the C allele appears protective from injury (O'Connell et al., 2015; Posthumus et al., 2009b), highlighting the potential importance of maturation in consideration of the genetic penetrance of variants with injury risk. Additionally, regardless of maturity, the *ACTN3* (rs1815739) and *MMP3* (rs679620) T alleles were associated with greater time loss following ankle injury and knee injury respectively (Hall et al., 2022). Nevertheless, others who have explored the relationship between TGSs for athletic performance have discussed the importance of contextualising the influence of genetic variants with consideration of their interactions with environmental factors to improve the estimation accuracy (Buxens et al., 2011; Grealy et al., 2015).

The *COL1A1* (rs1800012), *COL1A2* (rs412777), *COL5A1* (rs12722), *ACTN3* (rs1815739), *ESR1* (rs2234693), *MMP3* (rs679620), *ACE* I/D (rs1799752),*VDR* (rs2228570) and *GDF5* (rs143383) variants each appear to influence recovery and injury risk. Training and nutritional interventions have been shown to reduce injury incidence in elite male football (Lemes et al., 2021). Therefore, understanding how these variants may influence injury risk in elite male football development could support each individual to have a greater opportunity to achieve their potential through protection from injury with targeted interventions based on their genotype profile. A paradigm shift towards greater exploration of the real-world applicability of genetically informed training and decisions making practises to support individual athlete development has been highlighted (Pickering & Kiely, 2019). Nevertheless, it should not be forgotten that although current heritability estimates of athletic performance are significant, circa 50% (Bouchard et al., 1998; De Moor et al., 2007; Silventoinen et al., 2008), environmental and training factors are at least as important as genetics in injury risk. Therefore, understanding how environmental factors influence injury risk could improve the sensitivity of injury risk models and provide international opportunities to reduce injury risk with appropriate management. The genetic penetrance of included variants on injury risk appear to vary with age, sex and physiological stages of growth and maturation. Therefore, the aim of the research is to explore the potential influence of selected genetic, physiological and environmental factors on injury incidence in elite male football development to evaluate the potential applicability of this information to protect individuals from injury and support long term development.

2.5 Thesis aims and objectives

The overarching aim of this programme of research was to understand how genetic variants that influence injury risk could inform developmental programmes to reduce injury risk and support elite male football players to achieve their potential. Genetic association studies with injury risk continue to emerge with varying consistency, and systematic scrutiny is required to identify candidate variants with greater confidence for potential application in elite male youth football. Injury incidence is a complex emergent process and consideration of environmental and physiological factors, which contribute to injury incidence, may improve the accuracy of injury risk models. The influence of genetic variants on injury risk may vary depending on physiological stages of growth, maturation, and aging. Therefore, understanding how growth and maturation interact with genetic and environmental factors could provide greater opportunities for applied interventions to reduce injury incidence in elite male youth football development.

The aims of the thesis are:

1. To identify candidate genetic variants with potential utility of application to reduce injury risk in elite male youth football player development.

2. To explore how physiological development and candidate genetic variants independently influence the risk of injury in elite male football player development.

3. To develop potential applications of combining genetic, physiological, and environmental information to reduce injury risk in elite male youth football.

The objectives of the thesis are:

1. To review existent genetic association literature and identify candidate variants worthy of investigation with injury risk in young healthy physically active males.

2. To investigate how growth, maturation and aging influence the incidence of injury across elite male football development pathway age groups.

3. To assess if identified candidate genetic variants associated with tissue-specific injury risk are observed in elite male football development pathway age groups.

4. To develop tissue-specific injury risk models using the combination of selected genetic, growth, maturation, and loading information in elite male football players.

CHAPTER 3: Systematic Review and Meta-Analysis of Candidate Gene Association Studies with Fracture Risk in Physically Active Participants

This chapter details the systematic review and meta-analysis of candidate gene association studies with fracture risk in physically active participants. Meta-analyses of genetic association studies are important to examine previously identified variants as initial findings are frequently unable to be replicated (Ioannidis et al., 2001; Salanti et al., 2005). Initial systematic searching of online databases sought to examine all genetic association studies with all musculoskeletal injuries. However, this goal quickly appeared too broad to provide meaningful insight within the time period. Nevertheless, several reviews, systematic reviews, and meta-analyses, of genetic association studies with tendon, ligament and muscle injury were identified in the literature, which could inform the selection of candidate genes for investigation in the research project (Lv et al., 2017; Pabalan et al., 2018; Wang et al., 2017). However, although several studies had independently investigated the influence of genetic variants on fracture risk in young physically active participants, the qualitative assessment and quantitative pooling of results had not been completed. Fracture and apophysitis injury are frequent and potentially severe injuries affecting youth footballers (de Loës, 1995; Le Gall et al., 2006; Light et al., 2021; Materne et al., 2020; Wik et al., 2020a), which could be mediated with training and nutritional interventions (Faude et al., 2017; Lemes et al., 2021; Moran et al., 2013). Therefore, a systematic review and meta-analysis of candidate gene association studies with fracture risk in physically active participants was completed. The *COL1A1* (rs1800012), *COL1A2* (rs412777) and *VDR* (rs2228570) SNPs were identified as candidates for further investigation in subsequent research. Sex-specific analysis indicated a protective effect of the *COL1A1* (rs1800012) T allele in females despite previous associations with increased risk of osteoporotic fracture in the elderly. This suggests that the genetic penetrance of the T allele is influenced by sex / age and is not ubiquitously detrimental to bone strength as has been previously suggested.

3.1 Introduction

Fractures are major musculoskeletal injuries, accounting for 22% of all sport and recreation related injuries in the United States (Conn et al., 2003). Fracture rehabilitation requires substantial time away from competition / work for physically active populations, such as athletes or military personnel (Kaufman et al., 2000; Le Gall et al., 2006) and has a negative impact on performance (Hägglund et al., 2013). Fractures occur when exposure to extrinsic aetiological factors result in force transfer to bone, which exceeds the threshold tolerance of an individual (Meeuwisse et al., 2007) and may occur from acute impact forces or repeated loading with insufficient recovery (i.e., stress fractures) (Bennell et al., 1999). Physical activity provides an important stimulus for bone health and is recommended to protect against osteoporotic fracture (Kohrt et al., 2004). However, this stimulus also represent an exposure to potentially injurious, forceful impact or repeated loading of the musculoskeletal system (Bacon & Mauger, 2017; Launay, 2015; Meardon et al., 2015; Schuh-Renner et al., 2017) which cause non-osteoporotic fractures which are the focus of the present meta-analysis.

Genetic differences have been shown to influence the interindividual variability in fracture risk (Efstathiadou et al., 2001; Ji et al., 2010; Mann et al., 2001; Trajanoska et al., 2018) with heritable factors associated with between 20-54% of fracture liability depending on site and age (Andrew et al., 2004; Michaëlsson et al., 2005). Fracture risk is a complex trait, influenced by the cumulative effects of a currently unknown number of genetic variants, which interact to produce slight alterations in tissue composition, structure and regulation (Baumert et al., 2016; Herbert et al., 2018; Kozlovskaia et al., 2017). Several GWAS have identified SNPs which influence fracture risk in genes involved in skeletal structure and homeostasis via alterations in bone mineral density (Trajanoska et al., 2018). To confirm the findings of GWAS, or to identify novel genetic variants, contributing to variability in fracture risk, genetic association studies may select candidate genes, based on their mechanistic effect on fracture risk. SNPs can change the physiological functionality of a genetic product by altering the amino acid sequence or moderating expression directly. Others may not directly influence fracture risk but are frequently inherited, or in linkage disequilibrium, with unidentified variants that do. The major structural protein of bone is type 1 collagen (Mann et al., 2001), whilst vitamin D is also fundamental for bone homeostasis (DeLuca, 2005). SNPs in the collagen type 1 alpha 1 (*COL1A1*), vitamin D receptor (*VDR*) and LDL receptor related protein 5 (*LRP5*) genes have been associated with a three- to eight-fold increase of fracture risk among physically active participants in some studies (Blades et al., 2010; Chatzipapas et al., 2009; Korvala et al., 2010), yet others have shown no association with the same SNPs (Cosman et al., 2013; Varley et al., 2018). Several genetic variants within the *COL1A1*, *LRP5* and *VDR* genes, along with other candidate genes, have been inconsistently associated with fracture risk (Korvala et al., 2010; Varley et al., 2018). Researchers exploring genetic association with fracture risk often combine male and female participants, thus improving the statistical power of the analysis. It can be argued that autosomal (i.e., non-sex-specific) genes may be compared equivalently between the sexes. However, physically active females have a significantly greater incidence and absolute risk of fracture compared to males (Kaufman et al., 2000; Waterman et al., 2016; Wentz et al., 2011), which may influence the relative contribution of genetic differences to fracture risk. Therefore, combining physically active male and female participants in genetic association with fracture risk may contribute to the inconsistency observed across studies.

A meta-analysis of 370 studies found statistically significant heterogeneity in 14 out of 36 groups of genetic association studies on the same topic with stronger effects in the first study of a topic than subsequent replication attempts in 25 cases (Ioannidis et al., 2001). This may result from spurious findings which are not validated in subsequent research or because a gene effect may be stronger in some sub-populations than others (Ioannidis et al., 2001). Potential limitations such as linkage disequilibrium, population stratification and Hardy-Weinberg equilibrium (HWE) are inherent in genetic association studies, contributing to study heterogeneity (Ioannidis et al., 2003; Salanti et al., 2005). Additional variation resulting from issues relating to study design and quality, such as sample size calculations and reporting of participant characteristics are inconsistent in genetic association studies of fracture risk (Chatzipapas et al., 2009; Suuriniemi et al., 2006; Välimäki et al., 2005; Varley et al., 2016; Yanovich et al., 2012), yet omission of methodological details such as these can have a substantial influence on the study outcome (Ioannidis et al., 2003). Therefore, an independent analysis of study quality is necessary to understand the limitations of published genetic association studies in the extant literature.

Meta-analyses pool results from individual genetic association studies to evaluate overall effects with greater statistical power and identify heterogeneity between studies (Ioannidis et al., 2001; Salanti et al., 2005). It is unclear which genetic variants are consistently associated with fracture risk and whether the magnitude of the effect is dependent on factors such as gender or study quality. Therefore, the aim of this systematic review and meta-analysis was to evaluate the findings and quality of genetic association studies with non-osteoporotic fracture risk in physically active humans with sub-analysis of the influence of gender on overall findings.

3.2 Method

3.2.1 Search strategy

The current review is registered on the PROSPERO International prospective register of systematic reviews (Trial Registration: CRD42018115008) and was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement guidelines (Moher et al., 2009). A literature search to identify articles evaluating the association of genetic variants with fracture injury incidence was completed using a pre-determined search strategy in the PubMed, SPORTDiscus (EBSCO) and Science Direct databases from their inception on the 30th of October 2018. The exact search terms used were: Fracture OR Fractures AND Gene OR Allele OR Polymorphism OR SNP OR Variant OR Genetic. The title and abstract of search results were screened for relevant articles, which were selected for full text evaluation by two authors independently (ERM & MW) using predetermined eligibility criteria. The reference list of eligible articles was subsequently screened for further articles.

3.2.2 Inclusion and exclusion criteria

Genetic case-control association studies of fracture occurrence in physically active humans published in English in a scientific peer-reviewed journal were included in the analysis to identify previously investigated genetic variants. Participants were required to be healthy, and clearly reported as at least moderately physically active, as part of either their occupation (e.g., athletes and military personnel) or lifestyle, as defined by the ACSM's Guidelines for Exercise Testing and Prescription (Thompson et al., 2013). Any case studies or association studies with osteoporotic fracture, osteogenesis imperfecta, fracture recovery and genetic risk score evaluation studies were excluded.

3.2.3 Study selection and data extraction

Following the removal of duplicates, studies were screened independently by two reviewers (ERM and MW) with discrepancies concluded by consensus agreement. The following data were extracted from eligible articles: (1) study details (author, publication date, country of origin); (2) population characteristics (gender, age, ethnicity, physical activity); (3) genetic variant(s). Quality assessment and risk of bias assessments were conducted using the Q-Genie (Sohani et al., 2015)

and modified ROBINS-I (Qasim et al., 2019; Sterne et al., 2016) tools independently by two authors (ERM and YM). The Q-Genie tool categorizes studies as either poor, moderate, or good quality with the modified ROBINS-I determining risk of bias as low, moderate, serious, or critical. Study characteristics data are presented as means ± standard deviations.

3.2.4 Meta-analysis

Data analysis was performed by one author (ERM) and reviewed by another (MW). Data were extracted, where possible, in the form of genotype frequency distributions between fractures (cases) and non-fractures (controls) for males, females and combined if not reported separately. If only percentage distributions were reported, participant number for each group was calculated using overall participant number. If neither of the above was possible, authors were contacted directly for data.

A meta-analysis was performed to calculate overall fracture risk, with sub-analysis of males and females separately, as odds ratios (OR) for each SNP, with extracted data available from two or more studies using the following genetic association meta-analysis models of comparison: allele contrast, recessive, and homozygote contrast, as recommended by Lee (2015). The frequency distribution between fracture cases with the candidate risk allele, as theoretically identified by the studies, and non-injured controls was entered into a dichotomous Mantel-Haenszel meta-analysis for each model as shown in Figure 8 using RevMan 5.3 software (Cochrane Collaboration, Oxford, United Kingdom) to generate pooled ORs with 95% confidence intervals (CI).

Genetic models were analysed using either fixed ($I^2 < 20\%$) or random ($I^2 \geq 20\%$) effects models, depending on heterogeneity between studies, quantified with the I^2 statistic with sub-analysis of sex. To provide a qualitative indication of the magnitude of effect observed, the OR produced by meta-analysis were converted to the standard mean difference and described in line with those suggested for Cohen's d (Cohen, 1988; Sánchez-Meca et al., 2003). Funnel plots were generated using the outcome of all included SNPs for each genetic comparison models to allow visual interpretation of potential biases (Sterne et al., 2011). To evaluate the potential of bias between studies, Egger's Test (Egger et al., 1997) was conducted to indicate the presence of funnel plot asymmetry using SPSS (IBM SPSS Statistics for Windows, IBM Corp, Version 24.0. Armonk, New York) for each of the genetic comparison models.

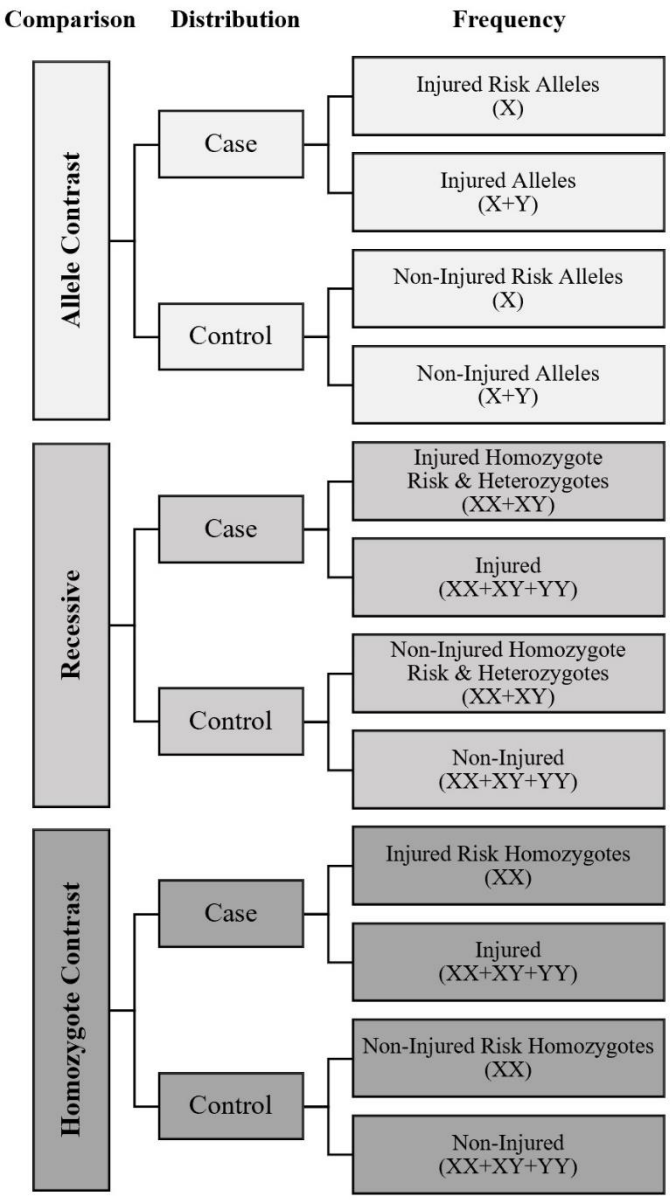

Figure 8. Meta-Analysis data input diagram. X, Risk Allele of genetic variant for fracture as defined by mechanistic rationale or candidate gene association study; Y, Non-Risk Allele of genetic variant. Symbols in parentheses indicate how the frequency counts were calculated to establish if differences in the risk allele distribution were present between cases and controls for each model.

3.3 Results

3.3.1 Study selection

Figure 9 outlines the results of the study selection process. Once duplicates were removed, reviews, case-studies and abstracts were excluded along with studies investigating clinical populations (including osteogenesis imperfecta or osteoporotic fractures patients) and fracture recovery. Reference list screening of the remaining articles provided one additional study, resulting in the full text review of 16 eligible studies. Five articles were excluded based on predetermined inclusion criteria, with qualitative assessments completed on the remaining 11 articles. Ten articles were included in the final meta-analysis due to lack of reported or available data in one study, determined after contacting the authors (Yanovich et al., 2012).

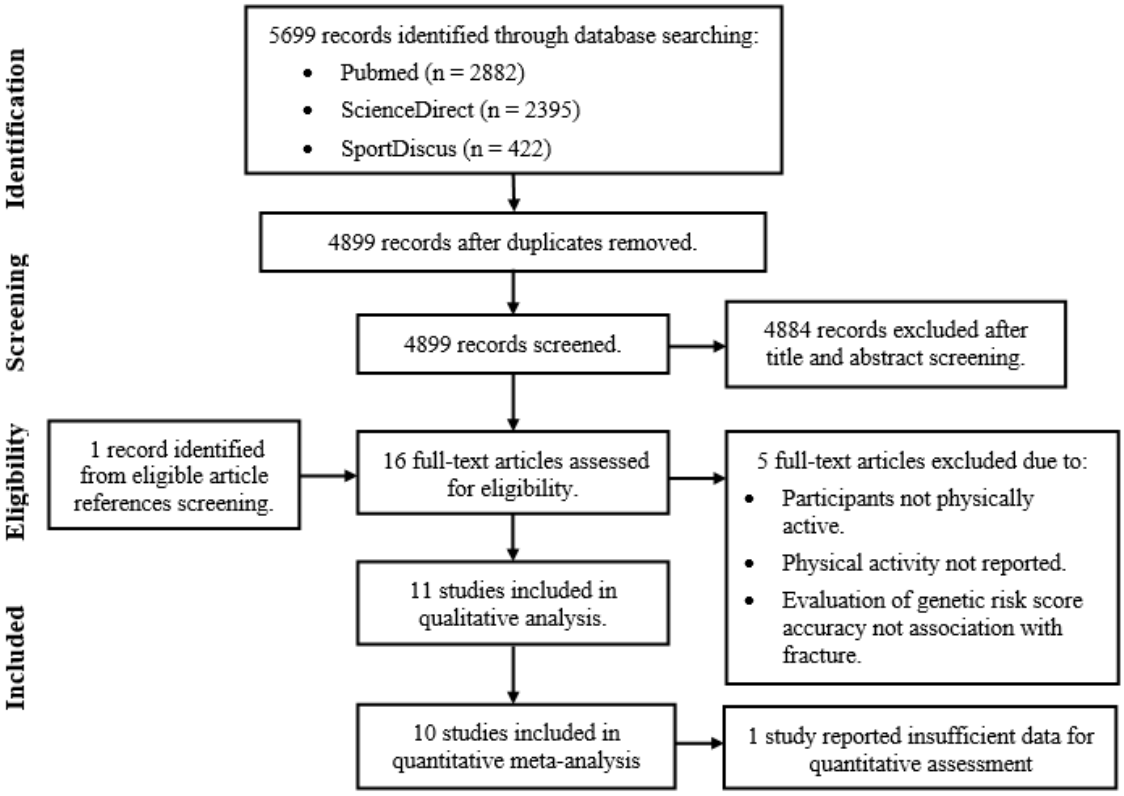

Figure 9. Genetic case-control association study of fracture risk in physically active participant systematic review and meta-analysis study selection process.

3.3.2 Study characteristics

The characteristics of each included study are summarised in Table 1. A total of 39 SNPs from 14 different genes were analysed at least once in the included studies. The mean sample of the studies was 499 ± 385 (males: 454 ± 400 & females: 117 ± 72). However, a convenience sample of the same 501 elite athletes from various sports (433 males and 68 females) was replicated in three studies evaluating different genetic variants with facture risk (Varley et al., 2015, 2016, 2018). Excluding these duplications, a total of 4462 (3676 males and 686 females) different physically active participants of various nationalities and ethnicities, aged 4 to 32 years are included. Of these, 961 were classified as fracture cases (779 males and 182 females) and 3501 considered non-fracture controls (2997 males and 504 females).

Two studies focused on acute fracture risk in children (Blades et al., 2010; Suuriniemi et al., 2003), the other nine evaluated stress fracture risk in professional adult military and / or athlete groups. Only one study investigated female participants alone (Suuriniemi et al., 2003), four investigated only males (Chatzipapas et al., 2009; Korvala et al., 2010; Välimäki et al., 2005; Zhao et al., 2016), five included both combined and separate analysis for male and female participants (Cosman et al., 2013; Varley et al., 2015, 2016, 2018; Yanovich et al., 2012) with one reporting only pooled results for males and females (Blades et al., 2010).

Only one paper achieved the highest classification of study quality (Blades et al., 2010), three were classified as moderate (Cosman et al., 2013; Korvala et al., 2010; Zhao et al., 2016), resulting in seven of the eligible studies defined as poor quality genetic association studies (Chatzipapas et

al., 2009; Suuriniemi et al., 2003; Välimäki et al., 2005; Varley et al., 2015, 2016, 2018; Yanovich et al., 2012). The overall risk of bias judgement varied from moderate to critical and was predominantly affected by bias due to confounding and participant selection. A summary of the assessment, including domain level judgments, are presented in Table 2. The funnel plots for the allele, recessive and homozygote comparison models, shown in Figures 10, 11 and 12 respectively, did not display a perfect funnel shape, but indicated no clear publication bias. The funnel plots generally display a cluster of large studies around the summary estimate and a lack of smaller studies spread beneath. The nine studies included in the Egger's test of this meta-analysis was just under the ten recommended as a rule of thumb for sufficient power by Sterne et al. (2011). However, the results of Egger's test indicated no sign of funnel plot asymmetry in any of the genetic meta-analysis comparison models, suggesting that there was no between-study bias within the included studies ($p > 0.24$).

3.3.3 Meta-analysis

Ten genetic variants from six different genes; *COL1A1* (rs1800012), *COL2A1* (rs412777), *CTR* (rs1801197), *ESR1* (rs2234693 & rs9340799), *LRP5* (rs3736228) and *VDR* (rs10735810, rs7975232, rs731236 & rs1544410) were replicated at least once in seven of the ten eligible studies, which constituted the quantitative meta-analysis. The summary statistics for each genetic comparison model meta-analysis are presented in Tables 3 to 6. Table 3 includes pooled analysis for all participants (male and females) of included studies, while Table 4 includes males and females from studies classified as good or moderate quality only. No statistically significant overall effect was found from the meta-analyses of any genetic model or SNP ($p > 0.06$). Tables 5 and 6 include summary statistics of male and female only sub-group analysis respectively. Sub-group analysis identified a significant reduction of fracture risk in female participants, with the T allele of the *COL1A1* rs1800012 SNP using the allele contrast model (OR = 0.48, 95% CI = 0.25 – 0.91, $p = 0.03$, $d = -0.18$), however this was not statistically significant in the recessive model (OR = 0.51, 95% CI = 0.24 – 1.06, $p = 0.07$, $d = -0.16$).

Significant overall heterogeneity was observed between studies in the *COL1A2* rs412777, *ESR1* rs9340799 and *VDR* rs10735810 meta-analyses, with significant sub-group heterogeneity found in the *COL1A1* rs1800012, *COL1A2* rs412777 and *ESR1* rs2234693 SNPs. Exclusion of poor-quality studies reduced the analysis to two genetic variants in two different genes (*COL1A1* rs1800012 & *VDR* rs1544410) from three studies, but this did not change the overall effect in these analyses.

Table 1. Summary of articles included from systematic review of genetic case-control association studies with fracture risk in physically active participants.

Study	Gene (SNPs)	Participant Characteristics					Result
		Sample	Case / Control & Sex	Physical Activity	Age (y)	Ethnicity	
Blades et al. (2010)	*COL1A1*: (rs1800012) *COL1A2*: (rs412777)	M & F English Children presented to A&E following impact trauma.	Fracture = 197 - (M = 124, F = 73) Control = 187 - (M= 106, F = 81) TOTAL = 384 - (M = 230, F = 154)	Recreational physical activity.	M = 11 ± 3 F = 11 ± 3 (4 - 16)	Caucasian	*COL1A2* 'PP' genotype halved fracture risk (p = 0.01, OR = 0.45, 95% CI = 0.24-0.82). *COL1A1* 's' allele trebled fracture risk in pre-pubertal children (p = 0.004, OR = 3.1, 95% CI = 1.43-6.63).
Chatzipapas et al. (2009)	*VDR*: (rs2228570, rs1544410, rs731236, rs7975232)	M only Military personnel	Stress Fracture = 32 Control = 32 TOTAL = 64	Military Duties	23 ± 3 (19 – 30)	Unknown	*VDR* rs2228570 'f' (p = 0.017, OR = 2.8, 95% CI = 1.2-6.3) and possibly rs1544410 'B' (p = 0.051, OR = 2.2, 95% CI = 1.0-4.4) alleles increase stress fracture risk.
Cosman et al. (2013)	*COL1A1*: (rs1800012) *ESR1*: (rs2234693, rs9340799) *VDR*: (rs1544410)	M & F US Military Recruits	Stress Fracture = 69 - (M = 43, F = 26) Control = 822 - (M = 712, F = 110) TOTAL = 891 - (M = 755, F = 136)	Basic Military Training	M = 19 ± 1 F = 18 ± 1 (18 – 20)	M: 86.5% Caucasian, 5% Asian, 8.5% Black. F: 79.4% Caucasian, 11% Asian, 9.6% Black.	No genetic association with stress fracture incidence. (p > 0.05).
Korvala et al. (2010)	*COL1A1*: (rs1800012, rs2696247, rs2586488, rs406226) *COL1A2*: (rs2301643, rs3216902, rs406226) *CTR*: (rs1801197) *IL-6*: (rs1800795) *LRP5*: (rs2277268, rs4988321, rs556442, rs3736228) *VDR*: (rs2228570, rs1544410, rs731236)	M only Finnish Military Conscripts	Stress Fracture = 72 Control = 120 TOTAL = 192	Basic Military Training	M = 20 ± 2 (18 – 27)	Unknown	Absence of CTR C allele and / or VDR C-A haplotype increased stress fracture risk (p = 0.007, OR = 3.22, 95% CI = 1.38-7.49). LRP5 haplotype A-G-G-C increased stress fracture risk (p = 0.031, OR = 2.72, 95% CI = 1.10 - 6.73) increasing when combined with the VDR C-A haplotype (p = 0.028, OR = 3.85, 95% CI = 1.16 – 12.84) but was mediated by body mass and BMI.
Suuriniemi et al. (2003)	*COL1A2*: (rs412777)	F Finnish children	Fracture = 37 Control = 221 TOTAL = 258	2.8 – 3.0 hrs / week	F = 11 ± 1 (10 – 12)	Unknown	COL1A2 P allele (either PP or Pp genotype) increased fracture risk compared to pp genotype (p = 0.007, OR = 4.1, 95% CI = 1.4 – 12.4).
Valimaki et al. (2005)	*ESR1*: (rs2234693, rs9340799)	M Finnish Military Conscripts	Stress Fracture = 15 Control = 164 TOTAL = 179	Basic Military Training	M = 19 ± 1 (18 – 20)	Unknown	No genetic association with stress fracture incidence. (p > 0.23).

(Continued)

Table 1. *Continued.*

Study	Gene (SNPs)	Participant Characteristics					Result
		Sample	Case / Control & Sex	Physical Activity	Age (y)	Ethnicity	
Varley et al. (2015)	*TNFSF11*: (rs1021188, rs9594738) *TNFRSF11A*: (rs3018362) *TNFRSF11B*: (rs4355801)	M & F Elite Athletes from USA and UK (SFEA Cohort)	Stress Fracture = 125 - (M = 98, F = 27) Control = 376 - (M = 335, F = 41) TOTAL = 501 - (M = 433, F = 68)	Professional Athletes of Various Sports	Stress Fracture = 27.2 ± 6.9 Control = 24.2 ± 5.5	Caucasian: Stress Fractures 83.2%, Controls 79.9% Other Unknown: Stress Fracture 16.8% Controls 20.1%	TNFSF11 rs1021188 AA (p = 0.024, OR = 2.9, 95% CI = 1.2 – 7.3) and TNFRSF11A rs3018362 GA+AA (p = 0.049, OR = 1.5, 95% CI = 1.0 – 2.4) individuals showed increased risk of stress fracture in comparison to GG individuals.
Varley et al. (2016)	*P2X7R*: (rs1653624, rs3751143, rs2230912, rs2230911, rs1718119, rs28360457, rs7958316, rs7958311, rs208294, rs28360447, rs17525809, rs35933842)	M & F Israeli Defence Force Soldiers and Elite Athletes from USA and UK (SFEA Cohort)	Military = 210 -(M = 198, F = 12), Stress Fracture = 43 -(M = 41, F = 2) Control = 167 -(M 157, F = 10) Elite Athletes = 501 - (M = 433, F = 68) Stress Fracture = 125 - (M = 98, F = 27) Control = 376 - (M = 335, F = 41) TOTAL = 711 -(M = 631, F = 80)	Military Training and Professional Athletes of Various Sports.	Military: Stress Fracture = 20.3 ± 1.6 Control = 18.9 ± 0.5 Athletes: Stress Fracture = 27.7 ± 7.5 Control = 24.4 ± 5.4	Elite Athletes: Stress fractures 83.2% Caucasian, 16.8% other. Controls 79.9%, Caucasian, 20.1% other. Military: Stress Fracture 36% non-Ashkenazi 64% Ashkenazi. Control 45% non-Ashkenazi and 55 % Ashkenazi.	P2X7R rs1718119 A allele (p = 0.01) and rs3751143 C allele (M only) (p = 0.04) associated with stress fracture occurrence in military participants. P2X7R rs3751143 C allele associated with stress fracture occurrence (p = 0.05) in elite athletes. After correcting for multiple comparisons using the false discovery rate test none of the findings remained significant (p > 0.05).
Varley et al. (2018)	*COL1A1*: (rs1800012) *CTR*: (rs1801197) *GC*: (rs4588, rs7041) *LRP5*: (rs3736228) *SOST*: (rs1877632) *VDR*: (rs2228570, rs7975232, rs731236, rs1544410) *WNT16*: (rs3801387)	M & F Elite Athletes from USA and UK (SFEA Cohort)	Stress Fracture = 125 - (M = 98, F = 27) Control = 376 - (M = 335, F = 41) TOTAL = 501 - (M = 433, F = 68)	Professional Athletes of Various Sports	Stress Fracture = 27.7 ± 7.5 Control = 24.4 ± 5.4	Caucasian: Stress Fractures 83.2%, Controls 79.9% Other Unknown: Stress Fracture 16.8% Controls 20.1%	SOST rs1877632 TT+TC v CC (p = 0.02), VDR rs2228570 (p = 0.01) & rs731236 (p = 0.01) C homozygotes (both M only) were associated with stress fracture occurrence. After correcting for multiple comparisons using the false discovery rate test none of the findings remained significant (p > 0.05).

(Continued)

Table 1. *Continued.*

Study	Gene (SNPs)	Participant Characteristics					Result
		Sample	Case / Control & Sex	Physical Activity	Age (y)	Ethnicity	
Yanovich et al. (2012)	*ANKH*: (rs4701616) *CALCR*: (rs12154667, rs1548456) *CBG*: (rs11629171, rs2281518) *COL1A2*: (rs420257, rs42517, rs42522, rs24531, rs413826) *IL6*: (rs1554606) *LRP4*: (rs2306033) *NR3C1*: (rs4244032, rs12656106) *ROR2*: (rs10992075) *VDR*: (rs4328262) Additional Not Reported.	M & F Israeli Defence Force Soldiers	Stress Fracture = 182 - (M = 165, F = 17) Control = 203 - (M = 162, F = 41) TOTAL = 385 - (M = 327, F = 58)	Military Training	Stress Fracture = 20.1 ± 1.7 (18 – 32) Control = 20.2 ± 1.3 (18 – 32)	Ashkenazi 49.5%, Non-Ashkenazi 38.1% and Unknown 12.4%.	NR3C1, ANKH, VDR, ROR2, CALCR, IL6, CBG, and COL1A2 associated with increased risk of stress fracture ($p < 0.05$). NR3C1, AR, VDR, CALCR, COL1A2, and LRP4 associated with reduced risk of stress fracture ($p < 0.05$). After correcting for multiple comparisons using the false discovery rate test none of the findings remained significant ($p > 0.05$).
Zhao et al. (2016)	*GDF5*: (rs143383)	M Chinese Military Recruits	Stress Fracture = 189 Control = 1209 TOTAL = 1398	Basic Military Training	Stress Fracture = 18.5 ± 1.4 Control = 18.5 ± 1.8	Unknown	GDF5 rs143383 T allele ($p < 0.001$, OR = 1.8, 95% CI = 1.4 – 2.3) and TT genotype ($p = 0.002$, OR = 1.8, 95% CI = 1.3 – 2.5) increased risk of stress fracture occurrence in comparison to C allele and TC+CC genotypes, respectively.

Note: M, Male; F, Female; SNP, Single Nucleotide Polymorphism; OR, odds ratio; CI, confidence interval.

Table 2. Risk of bias assessment judgements for genetic case-control association studies with fracture risk in physically active participants.

Author	Selection bias	Bias due to confounding	Bias in classification of exposure	Bias in assessment of outcome	Bias due to missing data	Bias in selection of reported results	Overall risk of bias
Blades et al. (2010)	Moderate	Serious	Moderate	Moderate	Low	Moderate	Serious
Chatzipapas et al. (2009)	Serious	Serious	Serious	Serious	Low	Moderate	Serious
Cosman et al. (2013)	Moderate	Moderate	Low	Low	Low	Moderate	Moderate
Korvala et al. (2010)	Serious	Serious	Serious	Serious	Moderate	Moderate	Serious
Suuriniemi et al. (2003)	Moderate	Moderate	Moderate	Moderate	Low	Moderate	Moderate
Valimaki et al. (2005)	Moderate	Moderate	Moderate	Moderate	Low	Moderate	Moderate
Varley et al. (2015)	Serious	Moderate	Low	Moderate	Low	Serious	Serious
Varley et al. (2016)	Serious	Serious	Serious	Moderate	Serious	Serious	Serious
Varley et al. (2018)	Serious	Serious	Serious	Moderate	Low	Serious	Serious
Yanovich et al. (2012)	Serious	Serious	Serious	Low	Serious	Critical	Critical
Zhao et al. (2016)	Moderate	Moderate	Low	Low	Low	Low	Moderate

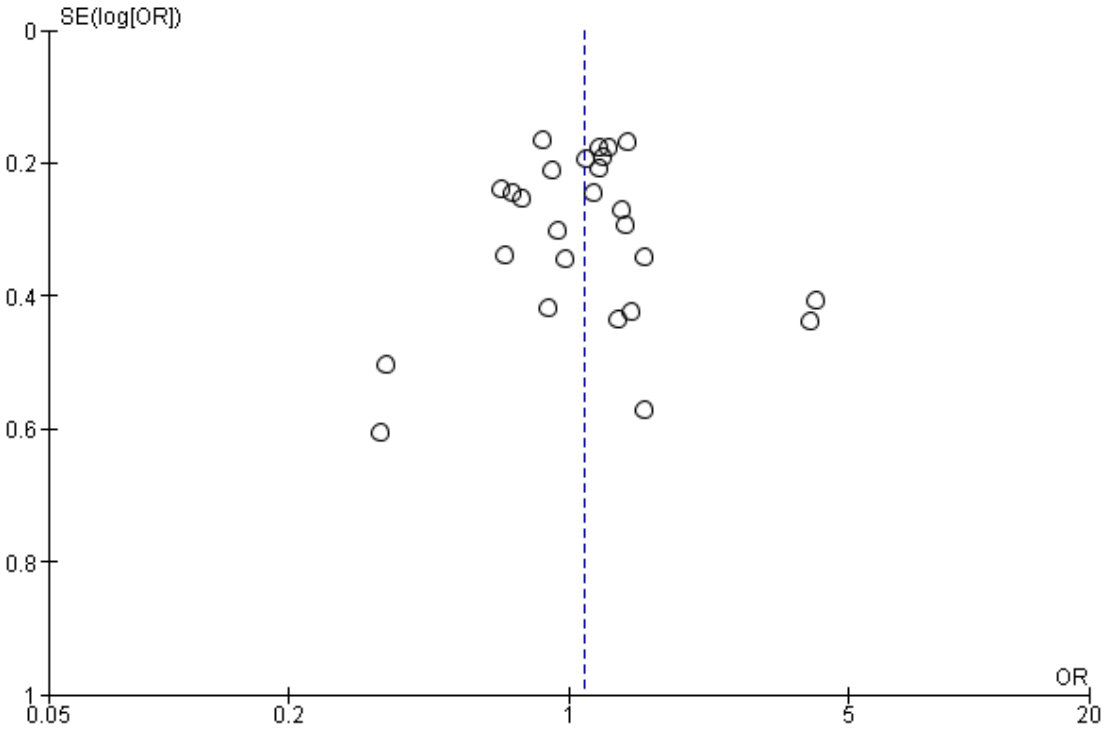

Figure 10. Funnel plot of single nucleotide polymorphisms replicated in studies investigating genetic association with fracture risk in physically active participants in the allele contrast model.

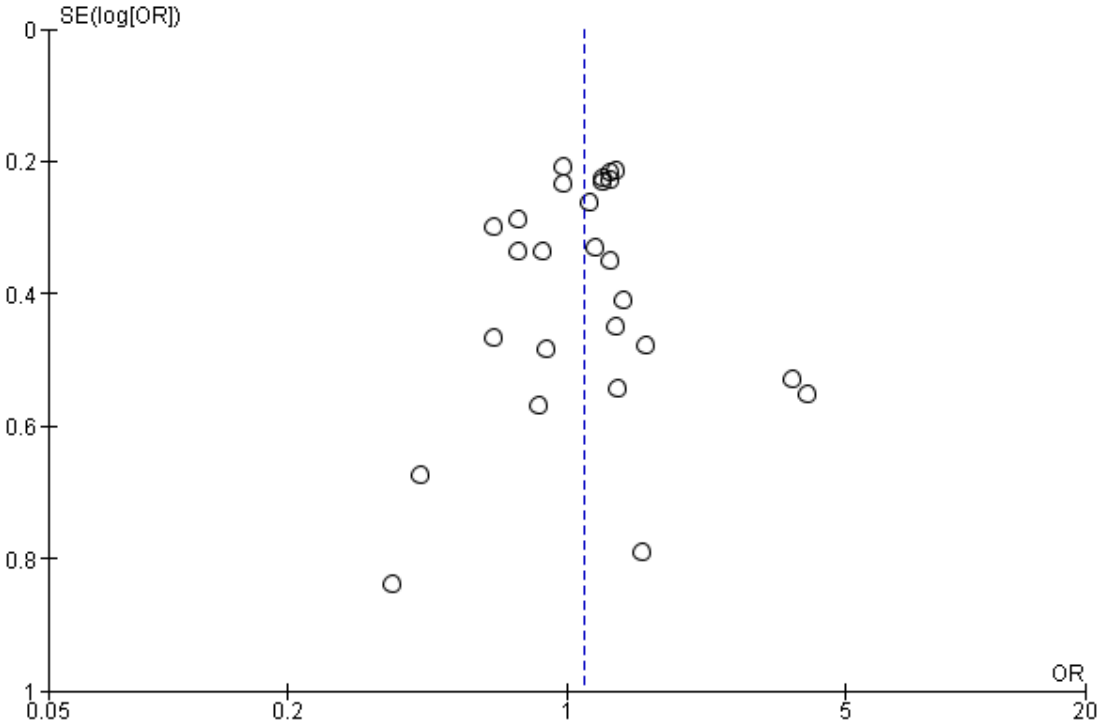

Figure 11. Funnel plot of single nucleotide polymorphisms replicated in studies investigating genetic association with fracture risk in physically active participants in the recessive model.

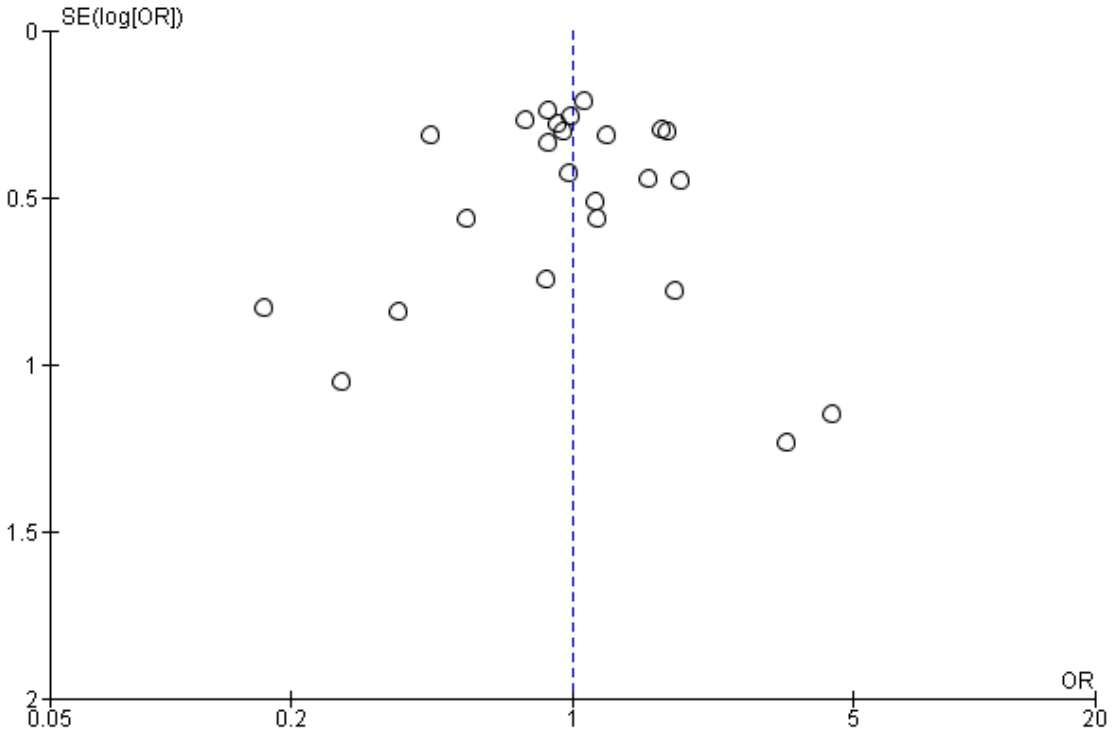

Figure 12. Funnel plot of single nucleotide polymorphisms replicated in studies investigating genetic association with fracture risk in physically active participants in the homozygote contrast model.

Table 3. Summary effects from the overall analyses of case-control association studies for genetic variants associated with fracture occurrence risk in physically active participants including all studies and sex sub-groups.

Genetic variant & comparison model	*n*	Test of heterogeneity				Test of overall association		
		Overall		*Sub-Group*		FE / RE	OR (95% CI)	*P*
		I^2	*P*	I^2	*P*			
COL1A1 (rs1800012)								
Allele contrast: T	4	17%	0.30	67%	0.05	FE	0.95 (0.76 - 1.77)	0.66
Recessive: TT+TG	4	1%	0.41	57%	0.10	FE	0.99 (0.77 - 1.27)	0.91
Homozygous: TT	4	0%	0.70	0%	0.63	FE	0.58 (0.25 - 1.32)	0.19
COL1A2 (rs412777)								
Allele contrast: C	2	92%	<0.001	92%	<0.001	RE	1.81 (0.39 - 8.52)	0.45
Recessive: CC+CA	2	83%	0.02	82%	0.02	RE	1.81 (0.46 - 7.17)	0.40
Homozygous: CC	2	85%	0.009	85%	0.009	RE	0.87 (0.22 - 3.52)	0.85
CTR (rs1801197)								
Allele contrast: T	3	0%	0.56	0%	0.88	FE	1.27 (0.82 - 1.97)	0.29
Recessive: TT+TC	3	0%	0.68	0%	0.96	FE	1.23 (0.67 - 2.27)	0.51
Homozygous: TT	3	25%	0.26	38%	0.20	RE	1.23 (0.81 - 1.87)	0.33
ESR1 (rs2234693)								
Allele contrast: C	2	0%	0.41	44%	0.18	FE	1.25 (0.80 - 1.95)	0.33
Recessive: CC+CT	2	0%	0.79	0%	0.49	FE	1.28 (0.69 - 2.35)	0.43
Homozygous: CC	2	48%	0.14	73%	0.05	RE	0.92 (0.43 - 1.98)	0.84
ESR1 (rs9340799)								
Allele contrast: G	2	68%	0.04	0%	0.81	RE	1.02 (0.35 - 3.00)	0.96
Recessive: GG+GA	2	33%	0.23	0%	0.80	RE	1.11 (0.40 - 3.09)	0.85
Homozygous: GG	2	4%	0.35	0%	0.49	FE	0.91 (0.57 - 1.45)	0.69
LRP5 (rs3736228)*								
Allele contrast: T	2	0%	0.92	0%	0.99	FE	1.14 (0.78 - 1.65)	0.50
Recessive: TT+TC	2	0%	0.94	0%	0.89	FE	1.18 (0.78 – 1.77)	0.43
VDR (rs2228570)								
Allele contrast: C	3	76%	0.006	0%	0.33	RE	1.60 (0.82 - 3.11)	0.17
Recessive: CC+CT	3	61%	0.05	0%	0.36	RE	1.49 (0.76 - 2.91)	0.25
Homozygous: CC	3	0%	0.71	0%	0.86	FE	1.49 (0.98 - 2.26)	0.06

(Continued).

Table 3. *Continued.*

Genetic variant & comparison model	*n*	Test of heterogeneity				Test of overall association		
		Overall		*Sub-Group*		FE / RE	OR (95% CI)	*P*
		I^2	*P*	I^2	*P*			
VDR (rs7975232)								
Allele contrast: C	2	0%	0.83	0%	0.66	FE	1.07 (0.76 - 1.51)	0.71
Recessive: CC+CA	2	0%	0.89	0%	0.91	FE	1.09 (0.68 - 1.73)	0.72
Homozygous: CC	2	11%	0.32	43%	0.19	FE	0.92 (0.60 - 1.41)	0.70
VDR (rs731236)								
Allele contrast: C	3	41%	0.17	28%	0.24	RE	1.05 (0.72 - 1.53)	0.80
Recessive: CC+CT	3	10%	0.34	0%	0.41	FE	1.01 (0.72 - 1.41)	0.96
Homozygous: CC	3	0%	0.52	10%	0.29	FE	0.98 (0.60 - 1.60)	0.95
VDR (rs1544410)								
Allele contrast: T	4	47%	0.09	0%	0.54	RE	0.97 (0.68 - 1.39)	0.87
Recessive: TT+TG	4	0%	0.50	0%	0.90	FE	1.08 (0.79 - 1.49)	0.62
Homozygous: TT	4	25%	0.25	0%	0.94	RE	0.85 (0.60 - 1.21)	0.37

Note: RE, random effects; FE, fixed effects; OR, odds ratio; CI, confidence interval. Sample size describes frequency of case and control counts for each model with risk variant frequency in parentheses. *LRP5 (rs3736228) TT homozygotes present in only one of the two included studies. Values in bold indicate significant heterogeneity and/or associations with fracture risk.

Table 4. Summary effects from the overall analyses of case-control association studies for genetic variants associated with fracture occurrence risk in physically active participants including only good and moderate quality studies with sex sub-groups.

Genetic variant & comparison model	*n*	Test of heterogeneity				Test of overall association		
		Overall		Sub-Group		FE / RE	OR (95% CI)	*P*
		I^2	*P*	I^2	*P*			
COL1A1 (rs1800012)								
Allele Contrast: T	3	36%	0.20	57%	0.10	RE	0.95 (0.66 - 1.36)	0.77
Recessive: TT+TG	3	33%	0.22	54%	0.11	RE	0.96 (0.65 - 1.42)	0.84
Homozygous: TT	3	0%	0.53	0%	0.47	FE	0.73 (0.28 - 1.91)	0.53
VDR (rs1544410)								
Allele Contrast: T	2	41%	0.18	0%	0.93	RE	0.96 (0.59 - 1.56)	0.87
Recessive: TT+TG	2	0%	0.45	0%	0.90	FE	1.00 (0.62 - 1.62)	1.00
Homozygous: TT	2	0%	0.39	0%	0.45	FE	0.93 (0.59 - 1.47)	0.75

Note: RE, random effects; FE, fixed effects; OR, odds ratio; CI, confidence interval. Sample size describes frequency of case and control counts for each model with risk variant frequency in parentheses. Values in bold indicate significant heterogeneity and / or associations with fracture risk.

Table 5. Summary effects of case-control association studies for genetic variants associated with fracture occurrence risk in physically active males only.

Genetic variant & risk comparison model	Sample size		Test of heterogeneity		Test of sub-group association		
	Participants	Studies	Within sub-group		FE / RE	OR (95% CI)	*P*
	Fracture / Control (Risk model frequency)		I^2	*P*			
COL1A1 (rs1800012)							
Allele Contrast: T	362 (17%) / 1960 (19%)	3	0%	0.99	FE	0.96 (0.70 - 1.31)	0.80
Recessive: TT+TG	208 (28%) /1138 (30%)	3	0%	0.93	FE	0.97 (0.68 - 1.37)	0.85
Homozygote Contrast: TT	208 (2%) / 1138 (3%)	3	0%	0.43	FE	0.87 (0.28 - 2.70)	0.81
CTR (rs1801197)							
Allele Contrast: T	271 (90%) / 706 (89%)	2	12%	0.29	FE	1.27 (0.80 - 2.08)	0.30
Recessive: TT+TC	163 (92%) / 442 (91%)	2	0%	0.38	FE	1.22 (0.62 - 2.38)	0.51
Homozygote Contrast: TT	163 (58%) / 442 (51%)	2	4%	0.31	FE	1.35 (0.94 - 1.95)	0.11
ESR1 (rs2234693)							
Allele Contrast: C	83 (81%) / 1294 (73%)	2	0%	0.99	FE	1.53 (0.80 - 2.91)	0.14
Recessive: CC+CT	54 (85%) / 849 (80%)	2	0%	0.96	FE	1.48 (0.69 - 3.20)	0.32
Homozygote Contrast: CC	54 (39%) / 849 (32%)	2	48%	0.14	FE	1.32 (0.75 - 2.32)	0.34
ESR1 (rs9340799)							
Allele Contrast: G	85 (91%) / 1351 (87%)	2	84%	**0.01**	RE	0.92 (0.13 - 6.62)	0.96
Recessive: GG+GA	54 (93%) / 849 (90%)	2	67%	0.08	RE	0.96 (0.15 - 6.36)	0.97
Homozygote Contrast: GG	54 (50%) / 849 (44%)	2	38%	0.21	RE	0.93 (0.43 - 2.01)	0.86
LRP5 (rs3736228)							
Allele Contrast: T	303 (12%) / 798 (12%)	2	0%	0.69	FE	1.14 (0.75 - 1.72)	0.54
Recessive: TT+TC	169 (21%) / 443 (21%)	2	0%	0.73	FE	1.16 (0.75 - 1.81)	0.50
Homozygote Contrast: TT	169 (0.6%) / 443 (1.1%)*	2*	N/A	N/A	N/A	N/A	N/A
VDR (rs2228570)							
Allele Contrast: C	304 (59%) / 710 (49%)	3	80%	**0.006**	RE	1.42 (0.64 - 3.15)	0.39
Recessive: CC+CT	194 (68%) / 454 (60%)	3	68%	**0.04**	RE	1.32 (0.59 - 2.95)	0.50
Homozygote Contrast: CC	194 (25%) / 454 (17%)	3	0%	0.51	FE	1.47 (0.95 - 2.28)	0.09
VDR (rs7975232)							
Allele Contrast: C	191 (71%) / 531 (71%)	2	0%	0.68	FE	1.03 (0.71 - 1.50)	0.86
Recessive: CC+CA	127 (78%) / 347 (76%)	2	0%	0.64	FE	1.08 (0.65 - 1.78)	0.78
Homozygote Contrast: CC	127 (28%) / 347 (31%)	2	0%	0.47	FE	0.82 (0.52 - 1.30)	0.40
VDR (rs731236)							
Allele Contrast: C	295 (47%) / 719 (48%)	3	43%	0.17	RE	0.96 (0.65 - 1.44)	0.85
Recessive: CC+CT	193 (60%) / 470 (60%)	3	25%	0.27	RE	0.96 (0.63 - 1.47)	0.85
Homozygote Contrast: CC	193 (12%) / 470 (13%)	3	0%	0.70	FE	0.90 (0.54 - 1.52)	0.70
VDR (rs1544410)							
Allele Contrast: T	313 (61%) / 1780 (70%)	4	67%	**0.03**	RE	0.89 (0.54 - 1.48)	0.65
Recessive: TT+TG	213 (71%) / 1162 (77%)	4	30%	0.23	RE	1.00 (0.64 - 1.56)	0.98
Homozygote Contrast: TT	213 (18%) / 1162 (30%)	4	47%	0.13	RE	0.84 (0.47 - 1.50)	0.54

Note: RE, random effects; FE, fixed effects; OR, odds ratio; CI, confidence interval. Sample size describes frequency of case and control counts for each model with risk variant frequency in parentheses. *LRP5 (rs3736228) TT homozygotes present in only one of the two included studies. Values in bold indicate significant heterogeneity and / or associations with fracture risk.

Table 6. Summary effects of case-control association studies for genetic variants associated with fracture occurrence risk in physically active females only.

Genetic variant & risk comparison model	Sample size		Test of heterogeneity		Test of sub-group association		
	Participants	Studies	Within sub-group		FE / RE	OR (95% CI)	*P*
	Fracture / Control (Risk model frequency)		I^2	*P*			
COL1A1 (rs1800012)							
Allele Contrast: T	92 (15%) / 242 (26%)	2	0%	0.64	FE	**0.48 (0.25 - 0.91)**	**0.03**
Recessive: TT+TG	52 (25%) /144 (38%)	2	0%	0.49	FE	0.51 (0.24 - 1.06)	0.07
Homozygote Contrast: TT	52 (2%) / 144 (6%)	2	0%	0.53	FE	0.41 (0.07 - 2.33)	0.31
VDR (rs1544410)							
Allele Contrast: T	76 (66%) / 222 (70%)	2	0%	0.69	FE	1.13 (0.63 - 2.05)	0.68
Recessive: TT+TG	51 (75%) / 144 (77%)	2	0%	0.90	FE	1.14 (0.52 - 2.48)	0.75
Homozygote Contrast: TT	51 (24%) / 144 (31%)	2	6%	0.30	FE	0.87 (0.40 - 1.89)	0.73

Note: RE, Random Effects; FE, Fixed Effects; OR, Odds Ratio; CI, Confidence Interval. Sample size describes frequency of case and control counts for each model with risk variant frequency in parentheses.

3.4 Discussion

The aim of this meta-analysis was to evaluate the findings of candidate gene association studies on non-osteoporotic fracture risk in physically active humans. Only ten SNPs from six different genes were independently replicated, despite the ten studies eligible for meta-analysis including 39 SNPs from 14 different genes. A sub-analysis indicated a sex-linked significant trivial reduction of fracture risk for physically active females with the T allele of the *COL1A1* rs1800012 SNP using the allele contrast model ($p = 0.03$, $d = -0.18$). However, no statistically significant overall effect was observed from the pooled results of any SNP ($p > 0.05$).

The discordance between the results of our pooled analysis and that reported in individual studies could be attributed to differences in methodological rigor, participant ethnicity and / or sex. The two studies that provided data for the *COL1A2* PvuII (rs412777) analysis presented conflicting results, with one reporting that the 'PP' genotype halved fracture risk (Blades et al., 2010), and the other suggesting that the P allele (either PP or Pp genotype) increased fracture risk (Suuriniemi et al., 2003). In the combined analysis performed herein, these contradictory results lead to null effects, which found no significant overall effect of the *COL1A2* PvuII (rs412777) SNP with fracture risk. The results of the studies may differ if two different proximal PvuII sites in the *COL1A2* gene have been assessed, only Blades et al. (2010) report the specific reference SNP number (rs412777); or if the intervention of the study on calcium and vitamin D supplementation, from which Suuriniemi et al. (2003) recruited their participants, influenced the effect of the P allele. It should also be considered that due to the age of participants in both studies, circa 11 years, which observed genetic associations with fracture cases could have been confounded by diseases which had yet to display symptoms or be diagnosed. Nevertheless, the ethnicity of participants was not reported by Suuriniemi et al. (2003) and may have differed from the Caucasian participants studied by Blades et al. (2010). Allele frequencies and baseline risk can vary across ethnic groups and failing to account for this may result in spurious associations with candidate genes (Pérez-Lezaun et al., 1997; Thomas & Witte, 2002). The investigated SNPs included in this meta-analysis may have no functional influence on fracture risk but exist in linkage disequilibrium with other SNPs that do. These patterns of linkage disequilibrium can differ across populations and associations in one but not another may be a result of these complex differences.

Five studies investigated the genetic association of stress fracture risk in Caucasian, or predominantly Caucasian, male adult military or athletic individuals. However, five additional studies did not report participant ethnicity; three of which provided the data for the *VDR* FokI (rs2228570) SNP analysis (Chatzipapas et al., 2009; Korvala et al., 2010; Varley et al., 2018). This suggests that the C allele had no overall effect on fracture risk using the random effects meta-analysis model. However, it has been argued that random effects models are not more conservative if the relative contribution of smaller low-quality studies on the overall effect are increased (Sterne et al., 2011). A fixed effects model was not considered appropriate for the *VDR* FokI (rs2228570) analysis, as heterogeneity was significantly high ($p = 0.006$, $I^2 = 76\%$) and the participants' ethnicity unknown. Nevertheless, all three studies reported accordance with Hardy-Weinberg equilibrium and a significant trivial increase of fracture risk is associated with the C allele using a fixed effects model

(OR = 1.37, 95% CI = 1.03 – 1.81, p = 0.03, d = 0.07) and ethnic variation across studies may have masked a valid genetic association.

The genetic architecture and interindividual variation of complex traits, such as fracture risk, are determined by numerous genetic variants with a range of effect sizes, which can be very small (Gibson, 2009). Additionally, heterogeneity between genetic association studies is often high, so several replication attempts are required to determine the physiological effect of genetic variants with confidence (Salanti et al., 2005). However, the SNPs included in this meta-analysis had only been examined in two to four studies, with many authors attempting to identify novel variants, instead of examining previous findings from GWAS or other candidate gene studies. The *LRP5* and *ESR1* genes, and the *LRP5* rs3736228 SNP, have been associated with fracture risk (Trajanoska et al., 2018) or bone mineral density (Kemp et al., 2017) in GWAS. However, these studies have focused on osteoporotic fracture and / or non-athletic individuals older than 18 and GWAS on fracture risk in young physically active healthy individuals appear absent from the literature. Many of the studies included in the current meta-analysis were of poorer quality and required further verification. Nevertheless, accurate replication would be challenging, as adequate reporting of participant characteristics was a common limitation of studies. One study failed to report if Hardy-Weinberg equilibrium was observed (Välimäki et al., 2005) and three reported disequilibrium for certain SNPs which were not included in the quantitative analysis (Varley et al., 2016, 2015; Yanovich et al., 2012). Sample size was another frequently observed limitation of included studies, with only two reporting a-priori power calculations (Blades et al., 2010; Zhao et al., 2016). As the effect of genetic variants may be small, a-priori power calculations are strongly recommended in genetic association studies (Salanti et al., 2005) and several authors suggest that their studies may have been underpowered (Cosman et al., 2013; Korvala et al., 2010; Välimäki et al., 2005; Varley et al., 2016; Yanovich et al., 2012; Zhao et al., 2016). Some authors also acknowledged the potential influence that differences in nutritional status could have on bone health and thus fracture risk (Blades et al., 2010; Cosman et al., 2013; Korvala et al., 2010; Varley et al., 2018). However, none of the studies included in this meta-analysis were able to control for dietary variation between groups.

Although no overall effect of the included SNPs was observed on fracture risk in this meta-analysis some genetic variants, such as the *COL1A2* PvuII (rs412777) and *VDR* FokI (rs2228570) SNPs, may still warrant further investigation. Indeed, genetic variants in *LRP5* and *ESR1* have been associated with osteoporotic fracture risk and bone mineral density in GWAS (Kemp et al., 2017; Trajanoska et al., 2018) and genuine physiological genetic effects could have been disguised by ethnicity dependent linkage disequilibrium with other influential variants or insufficiently powered analysis. Nevertheless, none of the included SNPs currently show a significant overall effect on fracture risk in physically active male and female combined analysis and could, indeed, have no physiological influence. However, further high-quality replication attempts would provide greater clarity of the influence of genetic risk factors for fracture risk in physically active participants. In the future, researchers should ensure a-priori power calculations are conducted and reported using clearly defined homogenous sample groups to inform the understanding of potential gene-environment and gene-gene interactions.

Only one of the studies included within this meta-analysis included female only participants (Suuriniemi et al., 2003). Six included both male and female participants, of these, five reported both

combined and sex specific analysis (Cosman et al., 2013; Varley et al., 2016, 2015, 2018; Yanovich et al., 2012) and one reported combined analysis only (Blades et al., 2010). Males and females are often combined in genetic association studies of injury risk, which will improve the sample size and statistical power of the analysis (Blades et al., 2010). This approach is rationalized, if the genetic variants are located on the autosomal regions of the genome, by stating that the region of interest is not linked to a specific sex. However, this explanation disregards the significant differences in the relative risk of bone injuries between the sexes (Arendt et al., 1999; Renstrom et al., 2008; Wentz et al., 2011). The division of sex in this meta-analysis identified more than five times the number of male than female participants. This resulted in larger standard errors in the female sub-groups with only the *COL1A1* Sp1(rs1800012) and *VDR* BsmI (rs1544410) SNPs replicated in females in more than one study. Stress fracture risk has been suggested to be three times greater in physically active female military personnel and 50% higher in female athletes than their male colleagues, due to biomechanical and physiological differences (Wentz et al., 2011). Indeed, the prevalence of fracture cases in the current meta-analysis was greater in females (27%) than males (21%). Significant sub-group heterogeneity was observed between sexes in the effect of the *COL1A1* Sp1 (rs1800012), *COL1A2* PvuII (rs412777) and *ESR1* PvuII (rs2234693) SNPs, highlighting the potential for sex-specific associations. Epidemiological data suggest that fracture incidence is greater in males between the age 18 and 49 than females in the general population (Curtis et al., 2016). However, the authors suggest this pattern reflects the increased prevalence of young males in high trauma events such as road traffic accidents (Curtis et al. 2016). Nevertheless, physically active females consistently demonstrate an increased risk of fracture when compared to their male counterparts (Kaufman et al., 2000; Waterman et al., 2016; Wentz et al., 2011). Therefore, the relative contribution of genetic susceptibility to fracture risk, and potential of preventative strategies, is likely to be greater in physically active females than males.

The *COL1A1* Sp1 (rs1800012) SNP, located in the 1st intron of the *COL1A1* gene, is one of the most extensively investigated genetic variants in the injury risk literature. Sub-group analysis within the current meta-analysis indicated a trivial reduction in fracture risk for the T allele of the *COL1A1* rs1800012 SNP in physically active females, but not males. However, the T allele was not associated with fracture risk when males and females were combined (Blades et al., 2010; Cosman et al., 2013; Varley et al., 2018), nor in males only (Korvala et al., 2010). The observed reduction in fracture risk associated with the T allele in females was not independently reported by the two studies included, which provided data for this meta-analysis (Cosman et al., 2013; Varley et al., 2018). Whilst it should be acknowledged that despite pooling data from these two studies this finding is only based on 196 females in total the T allele has been consistently associated with increased risk of osteoporotic fracture due to reduced bone mineral density in elderly postmenopausal females (Mann et al., 2001; Mann & Ralston, 2003). Nevertheless, the T allele has been repeatedly associated with reduced ligament injury risk in physically active mixed sex (Ficek et al., 2013; Khoschnau et al., 2008; Posthumus et al., 2009a) and male participants (Stępień-Słodkowska et al., 2013). These previous findings in addition to those of the current meta-analysis suggest the T allele could be associated with protection against some sport and exercise related injuries but further research is still required.

The T allele of the rs1800012 *COL1A1* SNP is associated with greater Sp1 binding affinity and COL1A1 production, which is similar between male and female carriers (Mann et al., 2001). This

results in an increased relative abundance of type 1 procollagen formed exclusively from COL1A1 polypeptides, which has been suggested to be weaker than the normal COL1A1 / COL1A2 combination (Mann et al., 2001). However, this is based on the increased osteoporotic fracture risk associated with the T allele in the elderly and is contradicted by the protective effect of the T allele observed in this meta-analysis and other studies of sport and exercise related ligament injuries (Ficek et al., 2013; Khoschnau et al., 2008; Posthumus et al., 2009a; Stępień-Słodkowska et al., 2013). Participants included in this meta-analysis and in the injury risk literature include physically active individuals, who are much younger than those studied in association with osteoporotic fracture. Mechanical loading of the musculoskeletal system is increased during sport and physical activity (Bacon & Mauger, 2017; Launay, 2015; Meardon et al., 2015; Schuh-Renner et al., 2017). Therefore, the T allele may result in mechanically stronger type 1 collagen, which is protective against ligamentous and bone injuries at younger ages in physically active individuals. The T allele may increase osteoporosis susceptibility in the elderly due to other pathogenic factors, such as excessive bone resorption. Alternatively, differences in the mechanical properties of bone and ligament may explain the observed variations in injury susceptibility and further investigation of the influence of the T allele on fracture risk in young physically active participants is needed.

Genetic variants do not necessarily result in dichotomous injured or non-injured states and genetic penetrance describes the probability that a carrier of a risk allele will express the disease / injury trait (D. N. Cooper et al., 2013). The genetic penetrance of the *COL1A1* rs1800012 SNP with fracture risk may be sex-specific and influenced by age. Therefore, it is possible that the T allele of the *COL1A1* rs1800012 SNP is concurrently associated with a reduced risk of bone fracture in young physically active females and an increased risk of osteoporotic fracture in elderly females. This finding is based on a total of 204 female participants (53 fractures and 151 controls), ten of which were TT homozygotes. This is lower than expected (~5%) considering the overall size of the sample as the minor T allele is present in 16-19% of Europeans and 9-13% of individuals globally. It may be that no association was observed in the recessive and homozygote contrast models or the individual studies (Cosman et al., 2013; Varley et al., 2018) as the number of T homozygotes, and overall participants, was low. Pooling the results of multiple genetic association studies becomes highly valuable to improve the statistical power of the analysis but the effect of the *COL1A1* rs1800012 SNP T allele on fracture risk should be replicated in a large group of physically active females in order to examine the finding of this meta-analysis.

3.5 Conclusion

The aim of the current meta-analysis was to synthesize the findings and quality of genetic case-control association studies on fracture risk in physically active participants. Sex-specific analysis indicated a protective effect of the *COL1A1* (rs1800012) T allele in females despite previous associations with increased risk of osteoporotic fracture in the elderly. This suggests that the genetic penetrance of the T allele is influenced by sex / age and is not ubiquitously detrimental to bone strength as has been previously suggested. The null effects observed in the overall analyses of SNPs included in this meta-analysis should not be considered finite due to potential limitations of the included studies. Paediatric participants, only present in the *COL1A1* (rs1800012) and *COL1A2* PvuII

(rs412777) combined sex analyses, are more likely to include individuals with undiagnosed asymptomatic diseases which could influence the genetic association results. Nevertheless, the overall findings *COL1A1* (rs1800012) combined sex analyses do not change if paediatric participants are removed. However, the *COL1A2* PvuII (rs412777) analyses is comprised exclusively of paediatric participants and should, therefore, be considered specific to this population and with the limitations discussed. Readers should also consider the potential influence that nutritional differences which could interact with the exposure of physical activity and genetic predisposition to mediate susceptibility to fracture occurrence in the included studies. Overall review of study designs indicated several recommendations for consideration in future research, such as the inclusion of *a-priori* power calculations, sex-specific analysis and greater clarity in the reporting of participant ethnicity. Consequently, further high-quality investigation of the *COL1A1* (rs1800012), *COL1A2* PvuII (rs412777) and *VDR* FokI (rs2228570) SNPs with fracture risk in a homogenous sample of physically active participants is warranted.

CHAPTER 4: Growth and Maturation as Risk Factors for Injury in Elite Male Youth Football

This chapter investigates the influence of growth and maturation on the incidence of all, non-contact, non-contact muscle and apophysitis injuries in elite male youth footballers when measured every 3-12 weeks. Growth and maturation have been implicated as risk factors for injury during adolescence (Kemper et al., 2015; Read et al., 2018b; Wik et al., 2020b). However, previous research has not examined the effect of changes in growth that occur in short measurement intervals which align with the requirement for current data to inform decision making in an applied setting. This chapter shows that increased age and maturation increase the risk of, all injuries, non-contact injuries and non-contact muscle injuries. However, growth rate alone was unable to differentiate individuals at increased risk of injury, even when dichotomised into low (< 0.6 $cm.m^{-1}$) and high (> 0.6 $cm.m^{-1}$) growth rate comparisons. Therefore, more complex and sensitive models are required to confidently identify players at risk of injury in an elite male youth football environment with more frequent measurements. Genetic variants previously associated with fracture risk in young physically active individuals could increase the sensitivity of models to identify individuals at risk of apophysitis injury during periods of rapid growth using measurements collected every 3-12 weeks.

4.1 Introduction

Among athletes, injuries occur when the repetitive and / or acute musculoskeletal tissue loading experienced exceeds the threshold tolerance of an individual at that time (Meeuwisse et al., 2007). Therefore, the high workloads of youth football players (Bowen et al., 2017, 2020) combined with tissue weakness during maturational phases of rapid growth (Wang et al., 2010; Wang & Seeman, 2009) inevitably place them at risk of musculoskeletal injury. The subsequent injury rehabilitation process disrupts access to developmental experiences for elite youth football players and may impede opportunities to realise their long-term performance potential (Jones et al., 2019). The majority of injuries in youth football occur via non-contact mechanisms, although the prevalence ranges between 34-72% of recorded injuries (Faude et al., 2013; Hall et al., 2020; Materne et al., 2020; Read et al., 2018b). Muscle strains and ligament sprains are the most frequently injured tissues, with the knee and ankle the most common locations of injury, each of which accounting for 20-30% of injuries in youth football (Le Gall et al., 2006; Price et al., 2004). The incidence of injuries in elite youth football peaks in the Under-13 / Under-14 years, which aligns with the adolescent growth spurt – also known as PHV - and pubertal / physical maturation (Johnson et al., 2020; Le Gall et al., 2006; Price et al., 2004; Read et al., 2018b). Youth players are particularly susceptible to severe time-loss tendinopathy and overuse apophysitis injuries, which are often associated with rapid growth (Le Gall et al., 2006; Light et al., 2021; Read et al., 2018b). These injuries are commonly termed “growth-related overuse injuries” accounting for 13-25% of all injuries between the ages of 11- and 14-years maturation (Johnson et al., 2020; Le Gall et al., 2006; Read et al., 2018b; Wik et al., 2020a), which can result in over 28 days of time-loss from training and matches per injury (Light et al., 2021; Materne et al., 2020). Furthermore, although these injuries are generally considered to resolve after six months, numerous studies have indicated potential long-term consequences on

physical performance and pain (Guldhammer et al., 2019; Holden & Rathleff, 2020; Kaya et al., 2013).

The overlap between the average age of PHV and peak injury incidence in youth football suggests that assessment of growth and maturation could provide identifiable risk factors for injury (Kemper et al., 2015; Wik et al., 2020b). Temporal disruption in motor-control (John et al., 2019), tissue weakness and / or tightness (Kerssemakers et al., 2009; Wang et al., 2010) of the musculoskeletal system are all candidate mediators of injury susceptibility during PHV due to rapid growth. Kemper et al. (2015) observed an injury incidence risk ratio (RR) of 1.63 ($p = 0.03$; 95% CI: 1.06-2.52) with a stature growth rate of 0.6 cm per month (7.2 cm per year) or more in elite male youth football players between the ages of 11 and 19 years compared to those with stature growth <0.6 cm per month. However, it is unlikely that a fixed physiological threshold such as this exists and discretisation of growth rates may not be appropriate, despite providing a practically useful risk threshold. For example, Wik et al. (2020b) examined the linear relationship between annual stature growth and various types of injury risk in a similarly aged (12-17 years) cohort of elite athletes using a generalised estimation equation. They reported the incidence risk ratio of bone injury increased by 1.47 (95% CI: 1.11-1.94) per standard deviation (SD) increase in annual stature growth rate (~0.7 cm per month or 8.9 cm per year) of their cohort, although this was not observed for all injuries, which contrasted with the findings of Kemper et al. (2015). However, annual measurements of stature are not sufficiently frequent to identify the considerable, and often saltatory, fluctuations in growth rate that occur during adolescence (Hermanussen, 1998; Lampl & Johnson, 1993). Furthermore, practitioners supporting the development of elite youth footballers might require more current information to support the purposeful and individually targeted training interventions that could be implemented to mitigate this injury risk. Nevertheless, a recent systematic review also concluded that insufficient evidence existed to indicate that biological maturity and growth in adolescence were associated with injury (Swain et al., 2018) and, therefore, further investigation is required.

Generalised estimation equations are similar to generalised linear mixed models (GLMM) but infer population-averaged rather than subject-specific effects (Hu et al., 1998). However, GLMM can account for subject-specific effects to determine the influence of growth and maturation variables on injury risk at both univariate and multivariate levels (Hu et al., 1998). Additionally, GLMM can accommodate for variation in observations and missing data between participants whilst controlling for the overrepresentation on any injury-prone individuals that may be otherwise overrepresented or excluded with different statistical methods. The exact timing, tempo, duration, and magnitude of PHV differs substantially between individuals (Abbassi, 1998; Sherar et al., 2005; Tanner et al., 1966) and despite several plausible theories, a direct causal link between rapid growth and injury is still unclear (Swain et al., 2018). Injury occurrence is multifaceted and the association of growth rate with bone injuries observed by Wik et al. (2020b) could be influenced by numerous hormonal, emotional and neurological changes, which occur alongside stature growth during pubertal maturity (Swain et al., 2018). Therefore, other factors should be considered in combination with stature growth rate, which might enhance the overall sensitivity of these measures to identify individuals at risk of injury. Indeed, associations between injury and many other growth and maturation-related measures, such as body mass index (BMI), leg length and percentage of predicted adult height (PAH), have been reported (Kemper et al., 2015; Wik et al., 2020b).

The risk of different types of injury appear to vary across time in youth football, along with pubertal growth and maturation (Hall et al., 2020; Light et al., 2021; Materne et al., 2020). Understanding which components of growth and maturation interact to influence the susceptibility of common injuries in elite youth football, such as non-contact injuries, non-contact muscle injuries and apophysitis injuries, could help to identify individuals for targeted training interventions. Therefore, the aim of the present study is to examine the relationship between growth and maturation measured periodically and the incidence of common injuries in elite male youth football players, using GLMM to investigate the influence of independent variables in isolated (univariate) and combined (multivariate) analyses.

4.2 Methods

4.2.1 Study design

A retrospective observational analysis was conducted on data collected, prospectively, as part of pre-existing growth, maturation, and injury monitoring processes. Study design and reporting were directed by published consensus statements on injury research in sport (STROBE-SIIS) (Bahr et al., 2020) and football (Fuller, 2006). Convenience sampling identified 165 eligible candidates for participation as male football players between 8 and 23 years of age, registered at an elite English Premier League category one academy with existing data for analysis. Of these, 80 participants, and parents of those under 18 years of age, provided written informed consent to participate in the study which, received institutional ethical approval (No. SMEC_2019-20_002). Only data collected as a registered player at the academy was included, which included data periodically collected between September 2014 and March 2020 from the Under-9 to Under-23 academy age groups.

4.2.2 Data collection

4.2.2.1 Anthropometrics

Stature, seated height, and body mass were measured every 5th, 6th, or 7th week, dependent on training schedule, including pre-season (July – May) of the standard football season, to better align with the needs of applied practitioners for current data. Stature and seated height were measured using a free-standing portable stadiometer (Seca 213 portable stadiometer; Seca, Birmingham, United Kingdom) to the nearest 0.1 cm by three different academy sports science staff, who followed the International Society for the Advancement of Kinanthropometry (ISAK) recommended procedures (Norton, 2018). The majority of measurements were collected by an ISAK Level one certified Anthropometrist and the primary author (ERM) of the study. Participants were asked to stand (stature) or sit (seated height) erect on the base of the stadiometer, without shoes and with their head in the Frankfort horizontal plane (Norton, 2018) for measurement. Measurements were repeated, and a third measure collected, if the difference between stature measures > 0.5 cm (> 1.0 cm for seated height). The two closest measures within this range were accepted and averaged to establish the recorded stature and seated height. Body mass was measured using portable digital scales (Seca 875 flat scale; Seca, Birmingham, United Kingdom) to the nearest 0.5

kg, with the final measurements stored in a secured central anthropometric database. Body mass index calculated as: body mass divided by squared stature ($kg.m^{-2}$) and leg length was calculated by subtracting seated height from stature. Participants who missed measurement within a scheduled window were measured as soon as possible over subsequent weeks.

4.2.2.2 Growth and maturation

Growth rate was calculated as the change in stature divided by the days from the previous measurement and reported in cm per month ($cm.m^{-1}$) and cm per year ($cm.y^{-1}$) to the nearest 0.1 cm. Growth rates were subsequently classified as low (< 0.3 $cm.m^{-1}$ = < 3.6 $cm.y^{-1}$), medium (0.3 - 0.6 $cm.m^{-1}$ = 3.6 - 7.2 $cm.y^{-1}$) or high growth (> 0.6 $cm.m^{-1}$ = > 7.2 $cm.y^{-1}$) to align with previously identified thresholds significantly associated with increased injury risk reported (Kemper et al., 2015; Wik et al., 2020b) and variation around normal growth through childhood, adolescence and into maturity (Abbassi, 1998). A dichotomous classification of growth as high growth (> 0.6 $cm.m^{-1}$ = >7.2 $cm.y^{-1}$) or not (<0.6 $cm.m^{-1}$ = <7.2 $cm.y^{-1}$) as a categorical variable was also evaluated for direct comparison with previous findings. A maturation estimate at each measurement was assessed based on the percentage attainment of PAH calculated with the Khamis-Roche method (Khamis & Roche, 1994) using self-reported biological parent statures and the measured stature and body mass of the participant at that age. Percentage attainment of PAH has been suggested as a non-invasive alternative estimate of maturation status (Malina et al., 2015) and was previously validated in youth athletes (Malina et al., 2007a, 2016). The maturation estimate was defined using previously established thresholds, which were aligned with the pubertal growth spurt (Malina et al., 2007a; Parr et al., 2020). Players were classified as Pre-, Circa- or Post-PHV when PAH attainment was <89%, 89-95% or >95% respectively (Johnson et al., 2020; Malina et al., 2007a).

4.2.2.3 Injuries and exposure

All players had access to academy medical staff, who recorded all complaints requiring attention, in accordance with consensus guidance on injury definition and data collection procedures in football (Fuller, 2006). The injury incidence, occurrence type (match or training) and date, onset (acute, overuse or mixed), mechanism (contact or non-contact), full training return date, diagnosis and location using the Orchard Sports Injury and Illness Classification System (OSIICS) version 10 (Rae & Orchard, 2007) were all recorded. This injury database was examined to identify time loss injuries within the study measurement period. Time loss injury was defined as tissue damage or disruption to normal physical function occurring from football or related training activities, resulting in at least one day of missed training or competition (Bahr et al., 2020; Fuller, 2006). Illnesses were excluded from the analysis. For the purposes of this study, we classified injury into four groups: all injuries, non-contact injuries, non-contact muscle injuries and apophysitis injuries (e.g., Osgood-Schlatter's and Severs disease). Only measures taken within 21-84 days (3-12 weeks) of the previous measurement were included as valid intervals for analysis with injury. Time loss injuries were assigned with the date of occurrence to the relevant valid anthropometric measurement interval. Only participants with six or more valid measurement intervals were included in the analysis and measurement intervals in which participants were already injured or ill were excluded from the

analysis. The type of occurrence, OSIICS code, onset and mechanism of each injury were verified by cross reference with injury rehabilitation / management notes and, where possible, video confirmation from match analysis archives. Injury severity was classified as minor (1-7 days), moderate (8-28 days) or severe (>28 days) based on the number of days between the date of time loss injury and return date to full participation in training based on published guidelines (Bahr et al., 2020). Incidence of injury was calculated as the number of injuries per 1000 hours of training and matches typically observed across the measurement intervals of included participants.

Table 7. Perceived meaningful change and average within-subject 2SD change in continuous independent variable.

Continuous independent variable	Perceived meaningful change	Average 2SD change in participant value
Age	0.5 years	1.8 years
Stature	5.0 cm	8.9 cm
Body mass	3.0 kg	9.2 kg
BMI	0.5 $kg.m^{-2}$	1.5 $kg.m^{-2}$
Monthly stature growth	0.6 $cm.m^{-1}$	0.5 $cm.m^{-1}$
Annual stature growth	2.5 $cm.y^{-1}$	6.4 $cm.y^{-1}$
Leg length	2.0 cm	4.3 cm
Seated height	2.0 cm	4.9 cm
Percentage of PAH	3.0 %	5.6 %

Note: SD; Within-subject standard deviation. BMI; Body Mass Index. PAH; Predicted Adult Height.

4.2.3 Data integrity and quality assurance

The intra-tester reliability of anthropometric measurements showed a technical error of measurement of 0.2 cm for stature and 0.3 cm for seated height, equating to a relative technical error of measurement of 0.1% and 0.4% respectively. The type of occurrence, OSIICS code, onset and mechanism of each injury were verified by cross-reference with injury rehabilitation / management notes and, where possible, video confirmation from match analysis archives. Any discrepancies in injury data, which could not be clarified with medical notes, were resolved and confirmed via consultation with medical staff who recorded initial assessment and/or injury management notes.

4.2.4 Statistical analysis

To evaluate the relationship between growth / maturation and injury incidence, we produced four sets of models, one for each injury classification: all injuries ($Injury_{ALL}$), non-contact injuries ($Injury_{NC}$), non-contact muscle injuries ($Injury_{MUSCLE}$) and apophysitis injuries ($Injury_{APOPH}$). GLMMs were used to analyse the data as they are robust against missing data, variations in measurement number between participants and permits entry of multiple measures / injury occurrences for each participant.

First, the relationship between all growth / maturation variables and injury incidence was evaluated in individual, univariate GLMMs. Injury occurrence dichotomously coded as either injured or non-injured was entered as the dependent variable. Age, stature, body mass, BMI, seated height, leg length, PAH attainment and growth rate ($cm.y^{-1}$ & $cm.m^{-1}$) were continuous independent

variables, with growth rate categories and maturation status the categorical independent variables. Participant ID was included as a random effect. Three values of the continuous independent variables were evaluated: a single unit change, a two within-subject SD (2SD) change, and the perceived meaningful change (Table 7). All continuous variables were evaluated as changes from the participant mean centred value. In total, thirty models were completed for each of $Injury_{ALL}$, $Injury_{NC}$, $Injury_{MUSCLE}$ and $Injury_{APOPH}$.

Following the univariate analyses, multivariate analyses of the statistically significant univariate predictors were conducted in each of the sets of injury classification models to provide further detail on the relationship between growth and maturation with the different types of injury ($Injury_{ALL}$, $Injury_{NC}$, $Injury_{MUSCLE}$ and $Injury_{APOPH}$). The relationship between independent variables were examined using Pearson's correlation (continuous-continuous variables) with PROC CORR and one-way ANOVA (continuous-categorical variables) with PROC ANOVA in SAS University Edition (SAS Institute Inc., Cary, NC, USA) to examine potential collinearity when included in a multivariate model. Injury was dichotomously coded as the dependent variable and participant ID was added as a random effect, as per the univariate models. The continuous fixed effects for the multivariate analyses were all calculated as a 2SD change from the participants' mean so that they were modelled on the same scale with categorical fixed effects entered as in the univariate analyses. They were entered using a forward forced stepwise entry method, with the most statistically significant variable entered first. Goodness of fit was compared with an intercept only model using the AIC fit statistic in a smaller-is-better form to establish the most suitable model for interpretation of results. Risk ratios (RR) were used to evaluate the influence of the independent variable on injury risk (e.g. between categorical variable groups, and resulting from a unit / 2SD continuous variable changes) with statistical significance set as $p<0.05$. Data were analysed with GLMM using PROC MIXED in SAS University Edition (SAS Institute Inc., Cary, NC, USA).

4.3 Results

A total of 1201 valid measurement intervals from eighty participants, across six academy football seasons, were included in the analysis. An overview of descriptive, injury, growth and maturation measures recorded within each age group is presented in Table 8.

4.3.1 Injuries and exposure

A total of 183-time loss injuries were observed for included participants throughout the study period, representing approximately 67,673 hours of elite academy football exposure (60,954 training hours & 6,719 match hours), equating to 2.7 injuries per 1000 h of training (1.4 injuries per 1000 h) and matches (8.9 injuries per 1000 h). Incidence rates varied across academy age group phases with 2.6 injuries per 1000 h of training (1.4 injuries per 1000 h) and matches (5.6 injuries per 1000 h) observed in the Under 9-12s, 2.5 injuries per 1000 h of training (1.3 injuries per 1000 h) and matches (7.6 injuries per 1000 h) in the Under 13-16s and 3.8 injuries per 1000 h of training (2.1 injuries per 1000 h) and matches (20.9 injuries per 1000 h) in the Under 18-23s. Twelve participants experienced

no time loss injury during the study period, twenty-one suffered a single injury and the remaining 47 suffered multiple injuries.

4.3.2 Injury location and pathology

The thigh (20.2%) was the most frequent region of injury followed by the ankle (18.6%) with muscle injury (26.8%) the most common pathology followed by tendon injury (20.2%), as shown in Table 9, along with classification of all injuries by location and pathology.

4.3.3 Injury onset, mechanism and severity

Of the 183 recorded time loss injuries, 101 were acute onset (55.2%), 60 were gradual (32.8%), nine displayed mixed, acute-on-chronic, onset (4.9%) with onset unspecified in the remaining 13 injuries (7.1%). Ninety-eight injuries occurred via non-contact mechanisms (53.6%), of which 35 were muscle injuries and 33 were apophysitis injuries (Sever's: 10; Osgood-Schlatter's: 15; Hip Apophysitis: 8). Seventy-one injuries resulted from a contact mechanism (38.8%) and 14 injuries did not have a recorded mechanism (7.7%). Of the contact injuries: 41 occurred following direct contact to the injured area from another player (22.4% all injuries; 57.7% of contact injuries); 15 resulted from ball kicks or saves (8.2% all injuries; 21.1% of contact injuries); 10 from direct contact resulting in an injurious fall (5.5% all injuries; 14.1% of contact injuries); and 5 were a result of indirect contact (2.7% all injuries; 7.0% of contact injuries). The average duration of recorded time loss injuries was 41 ± 59 days, with most categorised as moderate severity (43.2%), followed by severe (36.6%) and minor (20.2%) injuries, with apophysitis injuries the most prevalent severe injury, followed by muscle injuries as shown in Tables 10 and 11.

Table 8. Descriptive statistics, injury, anthropometric, growth and maturation data for elite male youth footballers.

Age group	Under-9	Under-10	Under-11	Under-12	Under-13	Under-14	Under-15	Under-16	Under-18	Under-23	All
Included participants (*n*)	8	16	28	35	51	42	34	28	21	6	80
Valid measurement intervals (*n*)	34	78	102	154	230	203	147	105	123	25	1201
All injuries (*n*)	3 (8.8%)	3 (3.8%)	13 (12.7%)	21 (13.6%)	34 (14.8%)	26 (12.8%)	24 (16.3%)	19 (18.1%)	31 (25.2%)	9 (36.0%)	183 (15.2%)
Non-contact injuries (*n*)	2 (5.9%)	2 (2.6%)	5 (4.9%)	13 (8.4%)	19 (8.3%)	13 (6.4%)	12 (8.2%)	11 (10.5%)	16 (13.0%)	5 (20.0%)	98 (8.2%)
Non-contact muscle injuries (*n*)	1 (2.9%)	0 (0.0%)	0 (0.0%)	3 (1.9%)	9 (3.9%)	4 (2.0%)	3 (2.0%)	3 (2.9%)	9 (7.3%)	3 (12.0%)	35 (2.9%)
Apophysitis injuries (*n*)	0 (0.0%)	2 (2.6%)	4 (3.9%)	9 (5.8%)	8 (3.5%)	5 (2.5%)	4 (2.7%)	1 (1.0%)	0 (0.0%)	0 (0.0%)	33 (2.7%)
Measurement interval (days)	52 ± 10	49 ± 10	53 ± 11	51 ± 12	49 ± 8	50 ± 7	50 ± 8	50 ± 10	50 ± 18	64 ± 18	51 ± 11
Age (years)	9.3 ± 0.2	10.1 ± 0.3	11.1 ± 0.3	12.1 ± 0.4	13.1 ± 0.4	14.1 ± 0.4	15.1 ± 0.3	15.9 ± 0.3	17.2 ± 0.6	19.0 ± 0.5	13.7 ± 2.2
Stature (cm)	136.2 ± 8.0	140.9 ± 6.9	147.8 ± 6.7	153.1 ± 8.5	159.6 ± 9.3	168.0 ± 8.2	174.4 ± 6.4	176.1 ± 8.1	179.9 ± 8.5	182.9 ± 8.8	163.1 ± 15.0
Body mass (kg)	30.7 ± 4.1	33.7 ± 5.1	37.8 ± 6.1	42.6 ± 7.6	47.8 ± 9.1	56.1 ± 9.0	64.2 ± 8.2	67.7 ± 9.1	72.2 ± 9.3	78.1 ± 7.9	53.2 ± 15.2
BMI ($kg.m^{-2}$)	16.5 ± 0.8	16.9 ± 1.6	17.2 ± 2.2	18.1 ± 1.9	18.6 ± 1.8	19.8 ± 1.8	21.1 ± 2.0	21.8 ± 2.1	22.2 ± 1.7	23.4 ± 1.7	19.5 ± 2.6
Seated height (cm)	69.5 ± 4.5	71.0 ± 3.1	73.9 ± 2.9	76.4 ± 4.2	79.5 ± 5.2	84.2 ± 5.2	88.2 ± 4.0	89.4 ± 4.3	91.6 ± 4.7	MD	81.4 ± 7.8
Leg length (cm)	66.7 ± 3.7	69.9 ± 4.5	73.9 ± 4.6	76.7 ± 5.2	80.1 ± 5.0	83.8 ± 4.2	86.2 ± 3.8	86.7 ± 4.9	88.1 ± 5.0	MD	80.6 ± 7.4
PAH attainment (%)	75.3 ± 1.2	78.1 ± 1.6	81.1 ± 1.6	84.4 ± 2.2	88.1 ± 2.8	92.2 ± 2.6	95.9 ± 1.6	97.8 ± 1.2	99.5 ± 0.8	100.3 ± 0.2	89.4 ± 7.1
Growth rate ($cm.m^{-1}$)	0.5 ± 0.3	0.4 ± 0.3	0.5 ± 0.3	0.5 ± 0.3	0.6 ± 0.3	0.6 ± 0.4	0.3 ± 0.3	0.2 ± 0.2	0.2 ± 0.2	0.1 ± 0.1	0.4 ± 0.3
Growth rate ($cm.y^{-1}$)	5.6 ± 3.5	4.9 ± 3.1	5.7 ± 3.6	5.6 ± 3.4	6.8 ± 4.2	6.8 ± 4.2	4.2 ± 3.4	2.4 ± 2.7	2.2 ± 2.6	1.0 ± 1.7	5.1 ± 4.0
Growth rate category	2.0 ± 0.8	1.8 ± 0.8	2.0 ± 0.8	2.0 ± 0.8	2.2 ± 0.8	2.2 ± 0.8	1.6 ± 0.8	1.3 ± 0.6	1.3 ± 0.6	1.1 ± 0.4	1.9 ± 0.8
Maturation status	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.1	1.4 ± 0.5	2.0 ± 0.5	2.8 ± 0.4	3.0 ± 0.1	3.0 ± 0.0	3.0 ± 0.0	1.9 ± 0.9

Note: Percentages of injury in relation to the total valid measurements. Data presented as means ± standard deviations of the measurements within the age group which may include multiple measures from the same participant. PAH: Predicted adult height. BMI: Body mass index. MD: Missing data. Growth rate category: 1 = Low (< 0.3 $cm.m^{-1}$ = < 3.6 $cm.y^{-1}$), 2 = medium (0.3 - 0.6 $cm.m^{-1}$ = 3.6 - 7.2 $cm.y^{-1}$) & 3 = high growth (> 0.6 $cm.m^{-1}$ = > 7.2 $cm.y^{-1}$); Maturation status: 1 = Pre-PHV, 2 = Circa-PHV & 3 = Post-PHV.

Table 9. Summary of injury region and pathology in an elite English Premier League academy of males in the Under-9 to Under-23 age groups.

Location	Bruising or haematoma	Muscle injury	Tendon injury	Joint sprains	Cartilage injury	Joint dislocations	Synovitis, impingement, or bursitis	Fracture	Stress fracture	Other stress or overuse	Nerve injury	Other not specified	Total
Ankle	4	0	2	22	0	0	4	2	0	0	0	0	34
Elbow	0	0	0	1	0	0	0	1	0	0	0	0	2
Foot	2	0	0	0	0	0	0	0	1	0	0	1	4
Head	1	0	0	1	0	0	0	0	0	0	8	0	10
Hip and groin	0	3	1	0	0	0	1	0	0	4	0	0	9
Knee	2	0	1	3	3	0	10	0	0	0	0	0	19
Lower leg	2	4	0	0	0	0	0	0	0	1	0	0	7
Lumbar spine	3	4	0	0	0	0	1	0	1	0	0	2	11
Neck	0	0	0	1	0	0	0	0	0	0	0	0	1
Pelvis and buttock	0	1	0	0	0	0	0	0	0	0	0	0	1
Shoulder	0	0	0	0	0	0	0	2	0	0	0	0	2
Thigh	2	35	0	0	0	0	0	0	0	0	0	0	37
Thoracic spine	0	2	0	0	0	0	0	0	0	0	0	0	2
Wrist and hand	1	0	0	1	0	1	0	8	0	0	0	0	11
Paediatric	0	0	33	0	0	0	0	0	0	0	0	0	33
Total	17	49	37	29	3	1	16	13	2	5	8	3	183

Note: Orchard Injury Classification 10.0 system (Rae & Orchard, 2007).

Table 10. Injury severity by region classification in an elite English Premier League academy of males in the under-9 to under-23 age groups.

Injured Region	Minor	Moderate	Severe	Total
Ankle	10	14	10	34
Elbow	1	0	1	2
Foot	2	0	2	4
Head	2	8	0	10
Hip and groin	2	5	2	9
Knee	3	9	7	19
Lower leg	3	2	2	7
Lumbar spine	1	7	3	11
Neck	0	0	1	1
Pelvis and buttock	1	0	0	1
Shoulder	0	0	2	2
Thigh	9	16	12	37
Thoracic spine	1	1	0	2
Wrist and hand	0	4	7	11
Paediatric	2	13	18	33
Total	37	79	67	183

Note: All paediatric injuries were apophysitis injuries as defined by the Orchard Injury Classification system (Rae & Orchard, 2007).

Table 11. Injury severity by pathology classification in an elite English Premier League academy of males in the Under-9 to Under-23 age groups.

Injury Pathology	Minor	Moderate	Severe	Total
Bruising or haematoma	8	8	1	17
Muscle injury	12	24	13	49
Tendon injury	4	14	19	37
Joint sprains	9	13	7	29
Cartilage injury	1	0	2	3
Joint dislocations	0	1	0	1
Synovitis, impingement, or bursitis	2	7	7	16
Fracture	0	1	12	13
Stress fracture	0	0	2	2
Other stress or overuse	0	2	3	5
Nerve injury	0	8	0	8
Other not specified	1	1	1	3
Total	37	79	67	183

Note: Apophysitis injuries included within tendon injury as defined by the Orchard Injury Classification system (Rae & Orchard, 2007).

4.3.4 Injury risk factors

The univariate effect of a single unit, 2SD and perceived meaningful change in anthropometric, growth and maturation variables on the incidence of $Injury_{ALL}$, $Injury_{NC}$, $Injury_{MUSCLE}$ and $Injury_{APOPH}$ are shown in Tables 12, 13 and 14, respectively. The raw values for a 2SD change in the continuous independent variables are provided in Table 7. All variables included in the multivariate analyses were significant and at least moderately correlated with every other ($p < 0.001$; $r > 0.51$). Growth rate was not associated with any significant change in injury risk for any injury classification analyses when considering a single unit (1 $cm.y^{-1}$ or 1 $cm.m^{-1}$) or participant mean centred meaningful unit change (2.5 $cm.y^{-1}$ or 0.6 $cm.m^{-1}$) ($p > 0.06$). Even when growth rate was discretised into categorical

segments, there was no significant difference between the risk of any injury classification analyses at high, medium or low growth rates ($p > 0.37$) or when dichotomised and growth rate >0.6 cm.m^{-1} compared to that <0.6 cm.m^{-1} ($Injury_{ALL}$: 1.08 [p = 0.64; 95% CI: 0.79-1.47]; $Injury_{NC}$: 0.83 [p = 0.38; 95% CI: 0.55-1.25]; $Injury_{MUSCLE}$: 0.86 [p = 0.68; 95% CI: 0.42-1.76]; $Injury_{APOPH}$: 0.67 [p = 0.26; 95% CI: 0.34-1.35]). However, injury incidence was significantly different between maturation status estimate categories for $Injury_{ALL}$, $Injury_{NC}$, $Injury_{MUSCLE}$ and $Injury_{APOPH}$ ($p < 0.01$) as shown in Table 15. The injury incidence was greatest when players were post-PHV for $Injury_{ALL}$, $Injury_{NC}$, $Injury_{MUSCLE}$ but when pre-PHV for $Injury_{APOPH}$, even when compared to circa-PHV 2.91 (p = 0.03; 95% CI: 1.09-7.76). This occurred despite the expected steady growth observed pre-PHV (5.9 ± 3.6 cm.y^{-1}), which then peaked circa-PHV (6.9 ± 4.2 cm.y^{-1}) and declined Post-PHV (2.7 ± 3.0 cm.y^{-1}).

A single unit increase in age (years) and stature (cm) was associated with a significantly greater injury incidence for $Injury_{ALL}$, $Injury_{NC}$ and $Injury_{MUSCLE}$ but a reduced injury incidence in $Injury_{APOPH}$ ($p < 0.05$). However, when examining the participant mean centred meaningful unit change, age and stature were only associated with a significant increase in injury incidence in $Injury_{ALL}$ and $Injury_{NC}$ ($p < 0.04$). The risk of injury incidence increased significantly in all models, except $Injury_{APOPH}$ ($p > 0.06$), for each single unit increase in body mass (kg), BMI (kg.m^{-2}), and percentage of PAH (%) ($p < 0.01$). However, these variables were only associated with a significant increase in risk for the $Injury_{ALL}$ analyses when compared with a meaningful increase from the participant mean centred value ($p < 0.03$). A 1.0 cm increase in seated height was associated with a significant increase in injury incidence for both $Injury_{ALL}$ and $Injury_{NC}$ ($p < 0.02$) but a decreased injury incidence for $Injury_{APOPH}$ ($p = 0.049$), with the $Injury_{NC}$ (which included both non-contact muscle & apophysitis injuries) not significant ($p = 0.07$). However, seated height was only associated with injury incidence in $Injury_{ALL}$ when considering a 2.0 cm perceived meaningful change from the participant mean centred value ($p = 0.02$). An increase in leg length was only associated with a significant increase in injury incidence in $Injury_{ALL}$ when considering the influence of a single unit change ($p = 0.04$).

Significant 2SD participant mean centred univariate variables were entered into a multivariate GLMM for the $Injury_{ALL}$ (maturation status, age, weight, percentage of PAH, stature, seated height, and BMI) and $Injury_{NC}$ (maturation status and age) analyses only because no significant 2SD participant mean centred univariate variables were present in the $Injury_{MUSCLE}$ or $Injury_{APOPH}$ analyses ($p > 0.07$). The multivariate model with the greatest fit compared to an intercept only model for $Injury_{ALL}$ (intercept only AIC = 1025.07; multivariate AIC = 869.41) incorporated maturation status, age, body mass, percentage of PAH, stature, and seated height. However, only seated height remained significant in this model ($p = 0.02$). The multivariate model for $Injury_{NC}$ included maturation status and age (AIC = 660.07) but was only slightly better than an intercept only model (AIC = 664.19) with neither variable significantly contributing to the overall model ($p > 0.06$).

Table 12. Univariate generalised linear mixed modelling analysis for a single unit change in anthropometric, growth and maturation injury risk factors in an elite male football academy.

Injury risk factor	All injuries		Non-contact injuries		Non-contact muscle injuries		Apophysitis injuries	
	RR (95% CI)	p	RR (95% CI)	p	RR (95% CI)	p	RR (95% CI)	p
Age (years)	1.16 (1.09-1.24)	*<0.0001*	1.17 (1.06-1.28)	*0.0012*	1.32 (1.13-1.55)	*0.0004*	0.82 (0.68-1.00)	*0.0495*
Stature (cm)	1.02 (1.01-1.03)	*0.0001*	1.02 (1.00-1.03)	*0.0108*	1.04 (1.01-1.06)	*0.0053*	0.97 (0.94-1.00)	*0.0337*
Body mass (kg)	1.02 (1.01-1.03)	*0.0001*	1.02 (1.01-1.03)	*0.0033*	1.04 (1.01-1.06)	*0.0015*	0.97 (0.94-1.00)	0.0621
BMI (kg.m^{-2})	1.10 (1.04-1.16)	*0.0011*	1.13 (1.04-1.22)	*0.0025*	1.22 (1.07-1.39)	*0.0022*	0.93 (0.78-1.12)	0.4692
Seated height (cm)	1.03 (1.01-1.05)	*0.0010*	1.03 (1.00-1.06)	0.0678	1.07 (1.01-1.12)	*0.0206*	0.95 (0.90-1.00)	*0.0486*
Leg length (cm)	1.02 (1.00-1.05)	*0.0437*	1.01 (0.98-1.05)	0.4626	1.04 (0.98-1.11)	0.2050	0.96 (0.91-1.01)	0.0954
PAH attainment (%)	1.04 (1.02-1.07)	*0.0003*	1.04 (1.01-1.08)	*0.0103*	1.10 (1.03-1.16)	*0.0055*	0.96 (0.90-1.01)	0.1201
Growth rate (cm.m^{-1})	0.98 (0.64-1.50)	0.9345	1.31 (0.75-2.30)	0.3476	1.18 (0.44-3.12)	0.7417	2.47 (0.95-6.46)	0.0642
Growth rate (cm.y^{-1})	1.00 (0.96-1.03)	0.9345	1.02 (0.98-1.07)	0.3476	1.01 (0.93-1.10)	0.7417	1.08 (1.00-1.17)	0.0642

Note: RR; Risk ratio. CI; Confidence interval. p-values in italics indicate significant association with p<0.05.

Table 13. Univariate generalised linear mixed modelling for participant mean centred meaningful unit change in anthropometric, growth and maturation injury risk factors in an elite male football academy.

Injury risk factor	All injuries		Non-Contact injuries		Non-Contact muscle injuries		Apophysitis injuries	
	RR (95% CI)	p	RR (95% CI)	p	RR (95% CI)	p	RR (95% CI)	p
Age (years)	1.13 (1.05-1.21)	*0.0011*	1.10 (0.99-1.21)	*0.0423*	1.18 (0.98-1.41)	0.0737	0.91 (0.77-1.07)	0.2440
Stature (cm)	1.15 (1.02-1.30)	*0.0219*	1.09 (0.88-1.17)	0.2237	1.23 (0.91-1.67)	0.1785	0.81 (0.63-1.06)	0.1269
Body mass (kg)	1.10 (1.02-1.18)	*0.0174*	1.06 (0.96-1.18)	0.1865	1.15 (0.95-1.38)	0.1531	0.87 (0.73-1.04)	0.1304
BMI (kg.m^{-2})	1.10 (1.01-1.19)	*0.0302*	1.07 (0.95-1.20)	0.2053	1.12 (0.91-1.38)	0.2672	0.89 (0.72-1.10)	0.2788
Seated height (cm)	1.12 (1.02-1.23)	*0.0230*	1.06 (0.93-1.20)	0.3306	1.08 (0.85-1.36)	0.5275	0.86 (0.71-1.06)	0.1512
Leg length (cm)	1.06 (0.96-1.18)	0.2357	1.01 (0.88-1.17)	0.7343	1.09 (0.84-1.42)	0.5265	0.87 (0.71-1.06)	0.1633
PAH attainment (%)	1.17 (1.02-1.34)	*0.0208*	1.10 (0.92-1.33)	0.2184	1.23 (0.88-1.74)	0.2276	0.82 (0.62-1.08)	0.1594
Growth rate (cm.m^{-1})	1.15 (0.87-1.51)	0.3385	1.33 (0.91-1.96)	0.1382	1.31 (0.68-2.52)	0.4204	1.54 (0.79-2.99)	0.2025
Growth rate (cm.y^{-1})	1.05 (0.95-1.15)	0.3385	1.11 (0.97-1.26)	0.1382	1.10 (0.87-1.38)	0.4204	1.16 (0.92-1.46)	0.2025

Note: RR; Risk ratio. CI; Confidence interval. p-values in italics indicate significant association with p<0.05.

Table 14. Univariate generalised linear mixed modelling for within-subject 2SD change in anthropometric, growth and maturation injury risk factors in an elite male football academy.

Injury risk factor	All injuries		Non-Contact injuries		Non-Contact muscle injuries		Apophysitis injuries	
	RR (95% CI)	p	RR (95% CI)	p	RR (95% CI)	p	RR (95% CI)	p
Age (years)	1.58 (1.20-2.09)	*0.0011*	1.50 (1.01-2.20)	*0.0423*	1.86 (0.94-3.66)	0.0737	0.69 (0.36-1.29)	0.2440
Stature (cm)	1.37 (1.05-1.80)	*0.0219*	1.26 (0.87-1.84)	0.2237	1.58 (0.81-3.10)	0.1785	0.63 (0.35-1.14)	0.1269
Body mass (kg)	1.38 (1.06-1.80)	*0.0174*	1.28 (0.89-1.85)	0.1865	1.60 (0.84-3.04)	0.1531	0.62 (0.33-1.15)	0.1304
BMI ($kg.m^{-2}$)	1.33 (1.03-1.73)	*0.0302*	1.27 (0.88-1.83)	0.2053	1.43 (0.76-2.72)	0.2672	0.70 (0.36-1.34)	0.2788
Seated height (cm)	1.40 (1.05-1.86)	*0.0230*	1.21 (0.82-1.80)	0.3306	1.25 (0.63-2.49)	0.5275	0.64 (0.35-1.18)	0.1512
Leg length (cm)	1.19 (0.89-1.60)	0.2357	1.07 (0.72-1.61)	0.7343	1.27 (0.60-2.70)	0.5265	0.66 (0.37-1.18)	0.1633
PAH attainment (%)	1.40 (1.05-1.87)	*0.0208*	1.29 (0.86-1.92)	0.2184	1.56 (0.76-3.23)	0.2276	0.65 (0.35-1.19)	0.1594
Growth rate ($cm.m^{-1}$)	1.13 (0.88-1.47)	0.3385	1.31 (0.92-1.87)	0.1382	1.28 (0.70-2.36)	0.4204	1.50 (0.80-2.79)	0.2025
Growth rate ($cm.y^{-1}$)	1.13 (0.88-1.47)	0.3385	1.31 (0.92-1.87)	0.1382	1.28 (0.70-2.36)	0.4204	1.50 (0.80-2.79)	0.2025

Note: SD; Within-subject standard deviation. RR; Risk ratio. CI; Confidence interval. p-values in italics indicate significant association with p<0.05.

Table 15. Univariate generalised linear mixed modelling analysis of maturation status estimate with injury risk in an elite male football academy.

Maturation Status	All Injuries			Non-Contact Injuries			Non-Contact Muscle Injuries			Apophysitis Injuries		
	Injury Risk Estimate (95% CI)	RR (95% CI)	p	Injury Risk Estimate (95% CI)	RR (95% CI)	p	Injury Risk Estimate (95% CI)	RR (95% CI)	p	Injury Risk Estimate (95% CI)	RR (95% CI)	p
Overall	0.14 (0.12-0.17)			0.08 (0.06-0.10)			0.03 (0.02-0.05)			0.01 (0.01-0.03)		
Pre-PHV	0.12 (0.09-0.15)	0.60 (0.43-0.83)	0.0019	0.07 (0.05-0.10)	0.59 (0.38-0.98)	0.0231	0.02 (0.01-0.03)	0.28 (0.12-0.64)	0.0028	0.03 (0.02-0.06)	5.39 (1.55-18.67)	0.0080
Circa-PHV	0.13 (0.09-0.18)	0.66 (0.46-0.95)	0.0256	0.06 (0.03-0.09)	0.48 (0.28-0.84)	0.0108	0.03 (0.01-0.06)	0.46 (0.20-1.09)	0.0795	0.01 (0.00-0.03)	1.85 (0.43-7.86)	0.4049
Post-PHV	0.20 (0.16-0.25)	1.00		0.11 (0.08-0.16)	1.00		0.06 (0.03-0.11)	1.00		0.01 (0.00-0.02)	1.00	

Note: RR; Risk ratio. CI; Confidence interval. PHV; Peak Height Velocity. Pre-PHV = <89%, Circa-PHV = 89-95% and Post-PHV = >95% attainment of predicted adult height. p-values in italics indicate significant difference from post-PHV or influence of maturation status with p<0.05.

4.4 Discussion

The main finding of the current study is that growth rate was not associated with the incidence of all injuries, non-contact injuries, non-contact muscle injuries or apophysitis injuries, when using anthropometric measures collected every 5-7 weeks in elite male youth football players. This finding remained consistent for both continuous and categorical analysis of high, medium, or low growth rate, and is contrary to previous observations of increased risk of all injuries (Kemper et al., 2015) and bone or growth plate injuries (Wik et al., 2020b) with growth above 0.6 $cm.m^{-1}$. From a mechanistic perspective, it is unclear why growth rate would increase the risk of all injuries ubiquitously (Swain et al., 2018), and, unfortunately, Kemper et al. (2015) did not discuss the potential reasons for their findings. Others have proposed adolescent motor awkwardness and, in turn, reduced ability to avoid potentially injurious contact situations following rapid growth as a potential risk factor to explain an increase in all football injuries (Van Der Sluis et al., 2014). However, a previous systematic review of research on adolescent awkwardness concluded that, although experience by some, it is unclear how this could affect injury risk (Quatman-Yates et al., 2012). Alternatively, if the risk of all other injuries remained equal, the observed risk of all injuries could increase during periods of rapid growth if the incidence of a particular type of injury, such as apophysitis injuries, increased dramatically. This was not observed in the present study, as the risk of all injuries increased as a function of chronological age and maturation, whilst the risk of apophysitis injury decreased in line with previous reports (Le Gall et al., 2007; Materne et al., 2020). Therefore, the results of the present study do not support the suggestion that growth rate significantly increases the risk of all injuries in youth football, which is consistent with reports among adolescent track and field athletes (Wik et al., 2020b).

Apophysitis injuries are thought to result from microfractures at the tendon-bone attachment site, considered the weakest bone structure in the young athlete (Arnold et al., 2017; Holden & Rathleff, 2020). These injuries are more common during maturational phases of rapid growth and might be related to a transient reduction in cortical bone mineral density (Wang et al., 2010; Wang & Seeman, 2009). However, in the present study, a single unit (1.0 $cm.m^{-1}$) increase in growth rate did not produce a significant increase in apophysitis injury incidence ($p > 0.06$), despite other authors previously observing a significant association with rapid growth and bone and growth plate injuries >0.7 $cm.m^{-1}$ (Wik et al., 2020b). The average growth rate measured across the entire cohort in the present study was 0.4 ± 0.1 $cm.m^{-1}$. This peaked in the Under-13 (0.6 ± 0.3 $cm.m^{-1}$) and Under-14 (0.6 ± 0.4 $cm.m^{-1}$) age groups, as expected in boys (Abbassi, 1998). These values align with the high growth rate categorisation threshold (>0.6 $cm.m^{-1}$) and are sufficient to evaluate the influence of rapid growth with injury. It is therefore surprising that a similar influence of growth rate on injury was not observed in the present study. Unlike Wik et al. (2020b), where annual measures of growth were reported, the present study collected measurements every 5-7 weeks, which was intended to match the needs of applied practitioners when planning training for youth players. It is possible that no significant association was established in the present study due to this increased measurement frequency. Although an increased measurement frequency allows for more accurate monitoring of an individual's changes in growth rates, it also results in shorter growth intervals, which may result in more extreme values being reported than over longer intervals. For example, one participant at

94.7% of his PAH, on the border of the circa- and post-PHV maturation category threshold, experienced high stature growth (0.7 and 1.1 cm.m^{-1}) in two consecutive measurement intervals (42 and 49 days from the previous measure), as shown in Figure 13. However, an annual stature measure over the same period would have been classified as medium growth (0.4 cm.m^{-1} 355 days from previous measure). Furthermore, measurements collected when players were classified as circa-PHV reached up to 1.6 cm.m^{-1} in a 42-day interval. Indeed, several measures of growth rate over 1.1 cm.m^{-1} were observed (Figure 13), which is much higher than expected from reference values using annual stature growth (Abbassi, 1998). These extreme growth rates, experienced across short periods of time, might not always lead to injuries, creating greater uncertainty in the estimates provided and contributing to non-significant findings, as shown by the high degree of uncertainty in the effect of growth rate on apophysitis injury (e.g., 95% CI range was 1.4 times greater than the mean OR for a perceived meaningful 0.6 cm.m^{-1} growth rate).

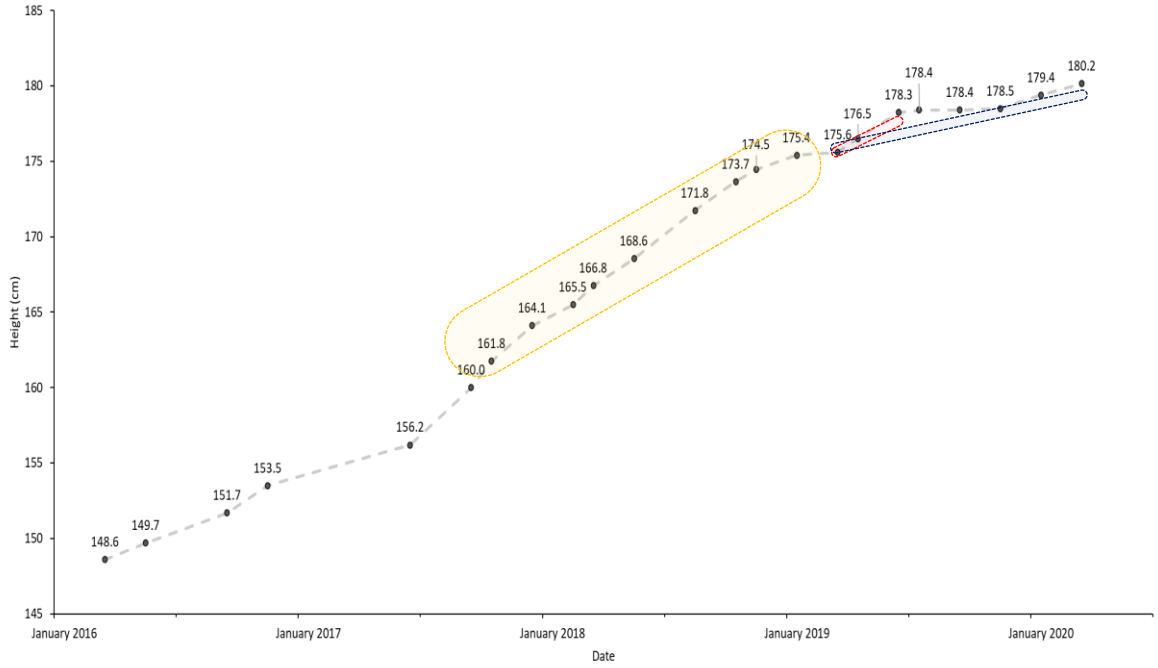

Figure 13. Example of participant stature growth throughout the observation period. Note: Yellow highlighted area indicates peak height velocity phase determined as current height measure between 89-95% of predicted adult height. Red area indicates two short measurement intervals classified as high growth (42 days at 0.7 cm.m^{-1} and 49 days at 1.1 cm.m^{-1}). Blue area highlights the corresponding annual measurement period, which would be classified as medium growth (355 days at 0.4 cm.m^{-1}).

Moderate (p < 0.001; r > 0.51), relationships were observed between all measured variables in the present study, which likely resulted in an inability for the model to effectively differentiate the contribution of the independent variables to injury incidence on a multivariate level. However, numerous physiological, psychological, social, and environmental changes occur alongside rapid growth, which could confound the association with injury risk (Swain et al., 2018). It is possible that annual measurements of growth rate, as used by Wik et al. (2020b), identify spurious significant relationships with apophysitis type injury, which may be attributable to other factors. Therefore, the results of the present study indicate that regular growth rate measurements are insensitive to identify individuals at risk of injury at that time in any practically meaningful timeframe for elite youth football. Nevertheless, the significant associations of previous research with rapid growth rate and injury

(Kemper et al., 2015; Wik et al., 2020b), alongside those showing a temporal reduction in bone strength (Wang et al., 2010; Wang & Seeman, 2009) and injury susceptibility during PHV (Johnson et al., 2020), suggest some relationship may still exist. It may be that a potential latent effect or extended period of risk follows rapid growth, which would not have been observed in the present study, with injury incidence directly aligned with growth rate within the same period. This could explain why others observed a significant association with injury using annual growth measures (Wik et al., 2020b) but not in the present study when using 5-7 weekly measurements, if injury risk remains elevated following, rather than occurring alongside, rapid growth. Future research may look to evaluate if prolonged or latent injury risk affects follow periods of rapid growth to provide greater clarity on the potential influence of rapid growth on injury risk.

Peak apophysitis injury incidence in the present study more closely aligns with the average age of take-off (i.e., when growth rate begins accelerating) (Abbassi, 1998). This occurred in the Under-12 age group when the average growth rate would be classified as medium growth (0.5 ± 0.3 $cm.m^{-1}$) and like that of younger age groups, as shown in Table 8. Although others have shown similar findings (Hall et al., 2020), apophysitis injury incidence usually peaks in older players around the Under-13 or 14 age groups (Materne et al., 2020; Read et al., 2018b). It may be argued that the earlier peak incidence in apophysitis injuries observed could also reflect an influence of maturation offset (early, on-time of late maturing), which was not assessed or controlled in the current study. However, this finding was observed along with the greatest risk of apophysitis injury associated with the pre-PHV phase ($p < 0.03$) and an average attainment of PAH of 84.4 ± 2.2% suggesting that this was not the case. Others have also observed later maturing (further from adult maturation than peers) footballers to have a significantly higher risk of apophysitis injury in the Under-14 age group (Le Gall et al., 2007) and greater maturation consistently coincides with a reduced incidence in apophysitis injury after 14 years of age (Le Gall et al., 2006; Price et al., 2004; Wik et al., 2020a). Wik et al. (2020b) adjusted their results to account for chronological age but acknowledged that the majority (68.5%) of their participants were early maturing, with baseline percentage of PAH and age, 92.5 ± 5.6% and 13.3 ± 0.9 years, respectively, and low average absolute annual stature growth of 3.4 cm. In contrast, the equivalent level of maturation in the current study was not observed until the Under-14 age group (92.2 ± 2.6% and 14.1 ± 0.4 years) with the average growth rate across similar ages (Under-12 to Under-16) also higher (5.2 $cm.y^{-1}$). Rapid growth in limb length precedes rapid torso growth (Malina et al., 2004a), as observed in the current study. Therefore, the association between leg length and injury risk observed by Wik et al. (2020b) likely represents the injuries occurring in later maturing individuals. This may also explain why Wik et al. (2020b) observed that stature growth >0.7 $cm.m^{-1}$ was associated with increased bone and growth plate injury incidence, whilst the average absolute change in stature across participants was low. Adolescent PHV is typically observed between 85-96% of PAH (Parr et al., 2020), which broadly aligns with Tanner genitalia stage 4 in males (Malina et al., 2005b; Marshall & Tanner, 1970). This led to the use of height and height growth velocity as a viable, non-invasive method of maturation estimation (Malina et al., 2005b, 2007a, 2012). However, 22% of the boys in the original study actually experienced PHV during genital stage 5 (Marshall & Tanner, 1970) and others have shown that around 30% of boys had still not experienced PHV even by this stage (Granados et al., 2015). Therefore, although useful for practical consideration, somatic estimations of maturation may be limited to confidently identify PHV for any given individual (Teunissen et al., 2020). Taken together, these findings suggest that,

although linked, consideration of growth and maturation as independent components of a complex and dynamic injury susceptibility process would be beneficial, and caution should be used when advocating for individual training programme modifications for based on somatic maturation estimates alone.

The increased risk of injury observed for players during PHV (Johnson et al., 2020) might still be evident but growth rate alone appears not to be the cause, which may be the result of other unknown mechanisms. The long-term implications of apophysitis injury appear to extend beyond the cessation of rapid growth (Guldhammer et al., 2019; Kaya et al., 2013), thus indicating the potential influence of other contributory factors. Therefore, individual inherent and / or temporal strength of bone tissue may differentiate the influence of growth rate on apophysitis injury incidence. Physical activity represents a potentially injurious exposure to the susceptible skeleton but appropriate mechanical loading prior to the end of puberty will significantly augment bone mass accrual (Khan et al., 2000; Specker et al., 2015). Wik et al. (2020b) observed only a moderate correlation between rapid growth and skeletal maturation ($r = 0.45$), and the dynamic interaction between rapid growth, skeletal maturation and physical activity is probably the most important consideration when managing the risk of apophysitis injury at the individual level. Modification or offload in training for youth players experiencing rapid growth to mediate injury risk in the short-term has been suggested (DiFiori et al., 2014; Faigenbaum et al., 2009). However, considering the importance of pubertal bone mass accrual and the potential benefits resulting from physical activity, a reduction in load during rapid growth may not be appropriate for long-term physical resilience for individuals not experiencing, and unlikely to suffer, an injury. Indeed, roughly half of life-time bone mass accrual occurs during puberty (Parfitt, 1994). This highlights the importance of current information to support the decision-making processes of applied youth practitioners, who may need to dynamically adjust the progressive overload of elite footballers to match the threshold tolerance, and support the long-term athletic development, of every individual.

The participants of Kemper et al. (2015) and Wik et al. (2020b) ranged from 11 - 19 and 11 - 16 years of age, respectively, with that of the current study between 8 - 20 years. This broader age range could have permitted the present study to include more injuries occurring further outside the expected age of PHV (Abbassi, 1998; Malina et al., 2007a; Parr et al., 2020), which would subsequently be expected to occur alongside slower or more steady growth rates. The timing, duration, and magnitude of PHV can vary between individuals (Abbassi, 1998; Malina et al., 2007a; Parr et al., 2020) and considering this broader age range reduces the risk of missing the effect of growth rate in individuals who grow rapidly outside of the normal timing. Increased age, and subsequent maturation (in relation to greater percentage attainment of PAH), was associated with increased injury risk and has consistently been shown to significantly increase the risk of youth football injury (Hall et al., 2020; Light et al., 2021; Materne et al., 2020; Read et al., 2018b; Wik et al., 2020a). As players age, physical performance improves through maturation as players gain significant increases in size, mass, strength, power, and endurance (Hansen et al., 1999; Philippaerts et al., 2006). This results in significantly increased expression and exposure to potentially injurious forces on the musculoskeletal tissues, both by the individual on themselves and from opposition players (Caine et al., 2008; Rumpf & Cronin, 2012; Wik et al., 2020a).

Maturational changes, such as PHV and peak weight velocity (Abbassi, 1998; Malina et al., 2007a; Parr et al., 2020), also broadly coincide with academy age group regulatory phase changes. Transition into a new Premier League Academy age group phase represents increased training and match demands on players as they progress to new regulatory constraints (e.g., 9 *vs.* 9 to 11 *vs.* 11 formats, larger pitch sizes and competition time 80 min to 90 min), which may also increase injury risk (Read et al., 2018b). Non-contact and non-contact muscle injuries are generally much more likely to occur because of increased training load exposure above the individual threshold tolerance at that time (Bowen et al., 2017; McCall et al., 2018) and the increase in all injuries with age and maturation occurred alongside both these injury analyses in the present study. Muscle injuries were much more prevalent in the Under-18 and Under-23 age groups than younger ages but a rapid increase in injury incidence is also seen in the Under-13 age group as growth rate begins to accelerate and some enter PHV. Others have also seen similar spikes in the Under-13 and Under-18 age groups (Hall et al., 2020; Light et al., 2021; Materne et al., 2020; Read et al., 2018b) which are the first in the Premier League Academy Youth Development (YDP) & Professional Development (PDP) phases. The injuries per 1000 h data indicate that the transition from Foundation Phase (Under 9-12s) to the YDP (Under 13-16s) had little influence on training injury incidence in the present study, although this increased in transition to the PDP (Under 18-23s). However, the injury incidence in matches increased substantially with each phase transition aligning with previous observations across age groups (Read et al., 2018b; Rumpf & Cronin, 2012; Wik et al., 2020a). Therefore, spikes in non-contact muscle injury incidence observed in the present study likely reflect the increased exposure that a new age phase and / or maturational development represents. Attention focused on managing the progressive overload and conditioning to game exposure based on growth and maturation between phases may provide beneficial reductions in injury incidence.

The findings of the present study should be interpreted contextually and considered alongside some limitations. No distinction was made herein between apophysitis injuries, which could have influenced the overall findings, although, this did not appear to occur from the spread of Osgood-Schlatter's Disease, Sever's Disease or Hip Apophysitis injuries. Individual measures rarely displayed smooth growth velocity curves and saltatory increases in stature are well documented (Hermanussen, 1998; Lampl & Johnson, 1993). Therefore, alternative growth velocity modelling strategies may provide greater clarity into the individual effect of growth on injury than point to point measurement assessments. Nevertheless, the findings of the present study suggest greater complexity in the injury susceptibility of elite adolescent male footballers than can be identified by growth rate alone. Future research should consider the potential of a latent effect of growth rate on injury incidence and / or the sensitivity of periodic measures compared with annual measurements to identify injury risk and support the decision-making processes of applied practitioners. Assessing the interaction of numerous injury risk factors across broader age ranges with skeletal maturation may provide further insight into the effect that growth timing has on injury incidence and to better identify individuals for whom rapid growth may increase injury risk as large variation is observed.

4.5 Conclusion

Although previous research has identified a significant association between growth rate and injury, this was not observed in the present study using measurements collected every 5-7 weeks. It is mechanistically unlikely that growth rate is related to all injuries but a causal link with apophysitis injury remains unclear due to the potential contribution of the numerous physiological changes and temporal disturbances in bone strength associated with maturation. Ultimately, growth and maturation alone are insufficient to confidently identify players at risk of injury in an elite youth football environment and more complex and sensitive models are required. To further understand factors affecting injury incidence in elite male football, chapter five explores genetic associations with apophysitis, muscle, bone, ligament, and tendon injuries across development age groups. The findings of the current study suggest that increased age and maturation increase the risk of all injuries, non-contact injuries and non-contact muscle injuries. Jumps in injury incidence appear to align with football academy phase transition age groups and attention should be placed upon managing the progression of exposure around these times. Chapter six draws upon the findings of chapters four and five to examine the combined influence of loading exposure, genetic variants and growth or maturation on injury incidence to identify potential applications in elite male football development.

CHAPTER 5: Candidate Gene Association Study with Injury in Elite Male Youth Footballers

This chapter assesses the influence of candidate genes, previously associated with common injuries in elite male youth footballers. Candidate genes were selected based on a bottom-up logical framework based on the genetic consequence of the variant resulting in a plausible physiological mechanism for injury risk. The *COL1A2* (rs412777) and *GDF5* (rs143383) SNPs appeared to influence fracture risk. The *MMP3* (rs679620) and *VDR* (rs2228570) SNPs were associated with the incidence of all non-contact injuries, whilst the *COL5A1* (rs12722) and *COL1A1* (rs1800012) SNPs were related to non-contact muscle and apophysitis injuries, respectively. Including the combined effect of these, and other, genetic variants with growth, maturation and/or loading exposure may enable genetically individualised tissue specific injury risk models to be developed. This information could then be used to develop bespoke interventions to mitigate injury risk and support long-term development based on individual genetic differences.

5.1 Introduction

Progressive overload in training exposure is designed by football coaches to stimulate an adaptive response, develop physical capabilities, and support competitive performance success (Cunanan et al., 2018; Helgerud et al., 2001; Kraemer & Ratamess, 2004). Musculoskeletal injuries occur when forces, distributed across the body, from the training and competition exposure result in tissue-specific load that exceeds the individual threshold tolerance at that time (Kalkhoven et al., 2021; Nielsen et al., 2018). Time loss injury and the requirement for rehabilitation before return to play is detrimental to senior team performance success (Hägglund et al., 2013) and restricts access to important developmental opportunities in youth footballers (Larruskain et al., 2021). Therefore, substantial research, debate and discussion has focused on the appropriate amount of loading exposure to improve physical performance without significantly increasing injury risk in football (Bowen et al., 2020; Dalen-Lorentsen et al., 2021; Fanchini et al., 2018; Malone et al., 2019). Nevertheless, musculoskeletal injury occurrence can be fundamentally considered as resulting from the dynamic interaction between both loading exposure and individual tissue threshold tolerance (Kalkhoven et al., 2021).

Heritability estimates provide an indication of the contribution of genetic differences between individuals on the observed variability in complex traits, such as injury risk (Visscher et al., 2008). Heritability estimates for ACL rupture, tennis elbow and wrist fracture injuries are around 69% (Magnusson et al., 2020), 40% (Hakim et al., 2003) and 54% (Andrew et al., 2004), respectively. However, identifying specific genetic variants which mediate this effect is challenging (Gibson, 2009) as the interaction of numerous DNA sequence variants with other genetic and environmental factors act to predispose, rather than determine, injury occurrence (Zondervan & Cardon, 2007). These variations in the sequence of the genetic code can interact with other genetic and environmental factors to have far-reaching consequences on the form and / or function of the resultant proteins and higher-level physiology throughout the body, which in turn can affect the tissue-specific load

tolerance and injury susceptibility of individuals with different genotypes (Lim et al., 2021; Rahim et al., 2019). Therefore, greater understanding of the genetic variations which mediate tissue-specific injury susceptibility in football may be informative for interindividual injury risk and prevention strategies.

Muscle, bone, ligament, and tendon tissue injuries are particularly prevalent in elite football (Le Gall et al., 2006; Price et al., 2004; Read et al., 2018b) and genetic variations associated with these tissues may be informative to inherent individual injury susceptibility. Bone fracture and apophysitis are particularly common, and often severe, injuries in children and adolescents footballers (Light et al., 2021; Read et al., 2018b; Tinkle & Wenstrup, 2005). The susceptibility and incidence of these injuries varies by age and maturation (Hall et al., 2020; Materne et al., 2021; Read et al., 2018b), with bone fracture and apophysitis injury incidence being greater prior to, and during, the peak adolescent growth spurt, known as PHV, and muscle, ligament and tendon injury increasing post-PHV and into adulthood (Emery et al., 2005; Light et al., 2021; Materne et al., 2020; Rumpf & Cronin, 2012). Therefore, the genetic penetrance, or probability that individuals with a given genotype will become injured (Lander & Schork, 1994), for specific injuries may evolve over time with exposure to both environmental and biological processes, such as training, maturation or ageing (Baumert et al., 2019). A G to T SNP (rs1800012) of the *COL1A1* gene appears a typical example of this and has been repeatedly associated with increased fracture risk in children and the elderly but reduced risk of both ACL rupture and fracture risk in young physically active populations, as shown in Table 16.

Apophysitis injuries are thought to result from repetitive loading which cause micro-avulsion fractures at the tendon-bone attachment site as rapidly growing long bones experience a transient increase in bone fragility due to delayed mineral acquisition (Faulkner et al., 2006; Wang & Seeman, 2009). Consequently, apophysitis injuries frequently occur alongside rapid growth and are common in children and adolescent footballers (Read et al., 2018b; Wik et al., 2020b). Nevertheless, large variation is seen in the influence of growth rate on injury incidence, as seen in chapter three, which may be partly attributable to interindividual genetic differences associated with fracture risk and bone properties. The *COL1A1* and *COL1A2* genes encode for procollagen sub-units which coalesce to form fundamental structures of bone, ligament and tendon (Birk et al., 1990). The T allele of the *COL1A1* (rs1800012) SNP has been shown to have a greater influence on fracture risk during adolescent growth (Blades et al., 2010) and is linked with inferior bone properties during early puberty (Suuriniemi et al., 2006). The T allele of the *COL1A1* (rs1800012) SNP has also been linked with impaired recovery following exercise-induced muscle damage (Baumert et al., 2018). A synonymous A to C SNP in the *COL1A2* (rs412777) gene, which shows random inheritance patterns across the population (linkage equilibrium) with the *COL1A1* (rs1800012) SNP, has also been associated with fracture risk and reduced bone mineral density (BMD) in children (Blades et al., 2010; Suuriniemi et al., 2003). Indeed, children who possessed both the *COL1A1* (rs1800012) GG and *COL1A2* (rs412777) CC genotypes consistently displayed some of the highest bone quality measures and suffered fewer fracture injuries than other genotype combinations (Blades et al., 2010). However, the only other study to investigate the influence of the *COL1A2* (rs412777) SNP with fracture risk in physically active children reported strong but directly contradictory results with the C allele associated with a four-fold increase in fracture risk (Suuriniemi et al., 2003). The *GDF5* (rs143383), *VDR* (rs2228570) and *ESR1* (rs2234693) SNPs influence gene transcription (Alonso et al., 2011; Jurutka

et al., 2000; Miyamoto et al., 2007) and have also been associated with fracture risk and bone fragility, as shown in Table 16. Therefore, the *COL1A1* (rs1800012), *COL1A2* (rs412777), *GDF5* (rs143383), *VDR* (rs2228570) and *ESR1* (rs2234693) SNPs appear to be potential candidates for investigation with fracture and apophysitis injury in elite male football players.

The *COL5A1* (rs12722)*, MMP3* (rs679620), *ACE* (rs1799752) and *ACTN3* (rs1815739) gene variants have been associated with several different musculoskeletal injuries in football, as shown in Table 16. Type V collagen co-assembles with type I collagen and plays a significant role in the fibrilogenesis of developing connective tissue (Birk et al., 1990; Wenstrup et al., 2004). An increased abundance of type V collagen appears to reduce the diameter of type I collagen fibres (Birk et al., 1990) in tendon, ligament and bone and a noncoding C to T SNP in the *COL5A1* (rs12722) gene results in more stable mRNA transcript (Laguette et al., 2011) indicating greater COL5A1 production. Consequently, the *COL5A1* (rs12722) C allele has been repeatedly associated with a reduced risk of tendon (Altinisik et al., 2015; Mokone et al., 2006) and ligament injury (O'Connell et al., 2015; Posthumus et al., 2009a). The *COL5A1* (rs12722) T allele also appears to interact with a T to C missense *MMP3* (rs679620) SNP to result in an even greater risk of tendinopathy (Raleigh et al., 2009). MMP3 stimulates the activity of other metalloproteinases (Toth et al., 2003), which regulate the extra cellular matrix by catalytically degrading structural proteins and collagens (Birkedal-Hansen et al., 1993; Somerville et al., 2003) and the *MMP3* (rs679620) C allele has also been independently associated with a reduced risk of hamstring injury in football players (Larruskain et al., 2018) and an increased risk of Achilles tendinopathy in high level athletes (Briški et al., 2021).

The *ACTN3* (rs1815739) SNP and *ACE* (rs1799752) insertion / deletion variants are amongst the most prolifically researched in association with physical performance related traits in football (McAuley et al., 2020) and have also been linked with injury risk as shown in Table 16. The ACTN3 protein is almost exclusively found as a constituent component of fast twitch muscle fibre sarcomeres, which is a common site of damage during unaccustomed eccentric exercise (Fridén & Lieber, 2001). The *ACTN3* (rs1815739) C to T SNP is a nonsense variant resulting in a premature stop codon (X allele) and production of non-functional ACTN3 (North et al., 1999). Those who possess the functional ACTN3 arginine coding allele (R) are thought to benefit from more robust sarcomeres (MacArthur & North, 2004), greater force transmission capabilities and enhanced adaptive signalling (Vincent et al., 2010). Indeed, R allele homozygotes (RR genotype individuals) appear to suffer significantly less muscle damage than X allele carriers (RX or XX genotype individuals) (Belli et al., 2017; Del Coso et al., 2017a) and express reduced bone remodelling markers (Levinger et al., 2017) following exercise, which may mediate the risk of muscle and fracture injury. Consequently, the R allele is associated with improved strength and power performance (Ma et al., 2013; Tharabenjasin et al., 2019) and reduced injury risk, although results are inconsistent, as shown in Table 16. ACE is important in blood pressure, electrolyte and fluid homeostasis (Nehme et al., 2019). A 287 base pairs insertion (I) / deletion (D) polymorphism in the *ACE* (rs1799752) gene affects ACE activity (Danser et al., 1995; Rigat et al., 1990) and the I allele is associated with increased exercise induced expression of genes involved in angiogenesis and aerobic metabolism (van Ginkel et al., 2016). However, the D allele is associated with greater increases in strength following training (Colakoglu et al., 2005; Folland et al., 2000) and appears protective against muscle injury in football (Massidda et al., 2020) as II homozygotes appear to suffer greater muscle damage following exercise

(Sierra et al., 2019). Genetic variants may influence injury susceptibility by directly affecting the structural integrity or homeostatic regulation of muscle, tendon, or ligament. However, injury susceptibility may also be affected by genetic variants which mediate recovery from exercise if these tissues are weakened for subsequent exposure to future exercise (Baumert et al., 2016). Therefore, the *ACTN3* (rs1815739), *ACE* I/D (rs1799752), *COL1A1* (rs1800012), *GDF5* (rs143383), *ESR1* (rs2234693), *MMP3* (rs679620), *COL5A1* (rs12722) and *VDR* (rs2228570) variants may influence the risk of muscle, tendon, and ligament injuries. However, previous genetic association findings remain inconsistent as shown in Table 16.

The individual tissue-specific threshold tolerance can be improved with training and / or supported with nutritional guidance, which differ by targeted tissue (Anderson et al., 2017; Lemes et al., 2021; Pasiakos et al., 2014). Therefore, identifying genetic variants that mediate tissue-specific injury susceptibility in elite football could improve the current understanding of injury prevention by exploring factors inherent to the individual which support targeted training interventions to mitigate injury risk. Genetic variants in several candidate genes affecting the form, function and injury susceptibility of muscle, bone, ligament, and tendon tissue have been previously identified in genes which code for structural constituents (*COL1A1* rs1800012, *COL1A2* rs412777, *COL5A1* rs12722 and *ACTN3* rs1815739) and / or regulatory components (*ACE* rs1799752, *ESR1* rs2234693, *MMP3* rs679620, *VDR* rs2228570 and *GDF5* rs143383) of the musculoskeletal system. The risk of different injuries varies with growth and maturation (Kemper et al., 2015; Light et al., 2021; Read et al., 2018b) so understanding how candidate genes, previously associated with injury susceptibility, influence the incidence of muscle, bone, ligament, and tendon injuries across the development pathway of elite footballers may provide valuable insight into inherent injury predisposition to support the long-term development of every individual. Therefore, the aims of the present study were twofold: firstly, to investigate the genetic association between previously identified candidate variants with muscle, bone, ligament, and tendon injuries across elite male football development age groups and, secondly; to identify if candidate genes previously associated with fracture risk are also associated with apophysitis injuries in elite male adolescent footballers.

Table 16. Overview of previous research on candidate genes for association with bone, ligament, tendon and muscle injury in elite male footballers.

Gene	Variant & Alleles	Location	Genetic Consequence	Physiological Effect	Study (Reference)	Participants	Genetic Variant / Minor Allele Injury Association Findings
COL1A1	rs1800012 G→T	17 Promoter region of first intron	↑ Sp1 Transcription factor binding	↑ COL1A1 Production and ↓ tissue strength	Mann & Ralston (2003) Blades et al. (2010) Ficek et al. (2013) Khoschnau et al. (2008) Posthumus et al. (2009a) Chapter three Cosman et al. (2013) Korvala et al. (2010) Varley et al. (2018) Erduran et al. (2014) Posthumus et al. (2009c) Gibbon et al. (2020) Larruskain et al. (2018) Pruna et al. (2013) Pruna et al. (2015) Bell et al. (2012)	M&F Elderly M&F Children M Adult Footballers M&F Patients M&F Patients M&F Physically Active M&F Military Recruits M Military Recruits M&F Athletes M&F Patients M&F Patients M&F Patients M Adult Footballers M Adult Footballers M Adult Footballers M&F Physically Active	↓ BMD & ↑ Fracture Risk ↑ Fracture Risk ↓ ACL Rupture Risk ↓ ACL Rupture & Shoulder Dislocation Risk ↓ ACL Rupture Risk ↓ Fracture Risk in F only No Association with Fracture Risk No Association with Fracture Risk No Association with Fracture Risk No Association with Tennis Elbow Risk No Association with Tendinopathy Risk ↓ ACL Rupture Risk No Association with Hamstring Muscle Injury Risk No Association with Soft Tissue Injury Risk No Association with Soft Tissue Injury Risk ↑ Joint Laxity
COL1A2	rs412777 A→C	7 Exon 25	Synonymous Proline coding	Unknown	Blades et al. (2010) Suuriniemi et al. (2003)	M&F Children F Children	↑ BMD & ↓ Fracture Risk ↑ Fracture Risk
GDF5	rs143383 G→A	20 5'UTR	Non-coding	↓ mRNA transcript production	Zhao et al. (2016) McCabe & Collins (2018) Chen et al. (2015) Ge et al. (2014) Posthumus et al. (2010a) Raleigh et al. (2013) Larruskain et al. (2018) Pruna et al. (2016)	M Military Recruits M Footballer M&F Patients M Military Recruits M&F Patients M&F Patients M Adult Footballers M Adult Footballers	↑ Stress Fracture Risk ↑ Ankle & Knee Injury Risk ↑ ACL Rupture Risk ↑ Meniscal Injury Risk ↑ Tendinopathy Risk No Association with ACL Rupture Risk No Association with Muscle Injury Risk No Association with Muscle Injury Risk
VDR	rs2228570 C→T	12 Initiator codon variant	Missense	↓ VDR transcriptional potency	Nakamura et al. (2002a) Nakamura et al. (2002b) Strandberg et al. (2003) Chatzipapas et al. (2009) Korvala et al. (2010) Varley et al. (2018) Diogenes et al. (2010) Massidda et al. (2015b)	M Athletes M Athletes M Adolescents M Military Recruits M Military Recruits M&F Athletes M Adolescent Footballers M Adult Footballers	↓ BMD ↓ BMD ↓ BMD ↑ Fracture Risk & ↓ Bone Strength No Independent Association with Fracture Risk ↑ Fracture Risk ↑ BMD No Association with Muscle Injury Risk
ESR1	rs2234693 T→C	6 Intron variant	Unknown	↑ gene expression	Cosman et al. (2013) Välimäki et al. (2005) Kumagai et al. (2018)	M&F Military Recruits M Military Recruits M&F Mixed Sport Athletes	No Association with Fracture Risk No Association with Fracture Risk ↓ Muscle Injury Risk

(Continued…)

Table 16. Continued.

Gene	Variant & Alleles	Location	Genetic Consequence	Physiological Effect	Study (Reference)	Participants	Genetic Variant / Minor Allele Injury Association Findings
MMP3	rs679620 T→C	11 Exon 2	Missense	Unknown	Larruskain et al. (2018) Pruna et al. (2016) Briški et al. (2021) Nie et al. (2019) Gibbon et al. (2016) Raleigh et al. (2009) Posthumus et al. (2012) El Khoury et al. (2016)	M Adult Footballers M Adult Footballers M&F Mixed Sport Athletes M&F Patients M&F Patients M&F Patients M&F Patients M&F Patients	↓ Hamstring Muscle Injury Risk No Association with Muscle Injury Risk ↓ Tendon Injury Risk ↑ Tendon Injury Risk No Association with Tendon Injury Risk ↑ Tendon Injury Risk ↓ ACL Rupture Risk with *COL5A1* rs12722 C Allele ↑ Tendon Rupture Risk
COL5A1	rs12722 C→T	9 3'UTR	Non-coding	↑ mRNA transcript stability	Larruskain et al. (2018) Pruna et al. (2013) Pruna et al. (2015) Pruna et al. (2016) Massidda et al. (2015a) Pabalan et al. (2018) Altinisik et al. (2015) Mokone et al. (2006) September et al. (2009) O'Connell et al. (2013) O'Connell et al. (2015) Posthumus et al. (2009b) Kim et al. (2014b) Bell et al. (2012)	M Adult Footballers M Adult Footballers M Adult Footballers M Adult Footballers M Adult Footballers M&F Caucasians Patients M&F Patients M&F Patients M&F Patients M&F Endurance Athletes M&F Patients F Patients F Ballet Dancers M&F Physically Active	No Association with Hamstring Muscle Injury Risk No Association with Soft Tissue Injury Risk No Association with Soft Tissue Injury Risk ↑ Muscle Injury Severity Risk ↑ Muscle Injury Severity Risk ↑ Tendon & Ligament Injury Risk ↑ Tendon Injury Risk ↑ Tendon Injury Risk ↑ Tendon Injury Risk ↓ Muscle Cramping Risk ↑ ACL Rupture Risk ↑ ACL Rupture Risk No Association with Injury Risk ↑ Joint Laxity
ACE	rs1799752 I→D	17 Intron variant	Indel variant	Altered hormone function	Larruskain et al. (2018) Massidda et al. (2020) Iwao-Koizumi et al. (2014) Kim et al. (2014b)	M Adult Footballers M Adult Footballers F Mixed Sport Athletes F Ballet Dancers	No Association with Hamstring Muscle Injury Risk ↓ Muscle Injury Risk ↓ Non-Contact Muscle Injury Risk (Heterozygotes) No Association with Injury Risk
ACTN3	rs1815739 C→T	11 Exon 15	Nonsense	Non-Functional ACTN3	Larruskain et al. (2018) Massidda et al. (2019) Clos et al. (2019) Iwao-Koizumi et al. (2014) Kim et al. (2014b) Shang et al. (2015) Qi et al. (2016) Kim et al. (2017b) Miyamoto et al. (2018) Moreno et al. (2020) Gutiérrez-Hellín et al. (2021) Zouhal et al. (2021)	M Adult Footballers M Adult Footballers M Adult Footballers F Mixed Sport Athletes F Ballet Dancers M Military Recruits M&F Patients M&F Patients M University Students M&F Adult Marathon Runners M&F Endurance Athletes M&F Athletes	No Association with Hamstring Muscle Injury Risk ↑ Muscle Injury Risk ↑ Non-Contact Muscle Injury Incidence ↓ Muscle & Ligament Injury Risk ↑ Ankle Joint Injury Risk ↑ Ankle Joint Injury Risk ↑ Ankle Joint Injury Risk No Association with Ankle Joint Injury Risk ↓ Hamstring Stiffness but No Injury Association ↑ Muscle Injury Risk No Overall Association with Injury Risk ↑ Non-Contact Injury Risk

Note: UTR; Untranslated region.

5.2 Methods

5.2.1 Study design

The study design and reporting were developed in concordance with previously published guidelines for genetic association studies (Little et al., 2009; Romero et al., 2002) and consensus statements for injury research in sport (STROBE-SIIS) (Bahr et al., 2020) and football (Fuller, 2006). A retrospective observational case-control genetic association experimental design was adopted to determine whether the distribution of risk alleles in candidate genes were higher in injured than non-injured players. The study received institutional ethical approval and was registered at ClinicalTrials.gov (NCT04952662).

5.2.2 Participants

Convenience sampling identified 185 eligible candidates who were invited to participate as elite male football players between 10 and 35 years of age at the time of recruitment. Of these, 113 players and the parents of those under 18 years of age, provided written informed consent to participate in the study. One participant was excluded as >50% of his genotypes could not be clearly determined using the saliva sample provided, which could not be resampled. A summary of included participant characteristics at the time of recruitment and self-reported participant ethnicity can be found in Tables 19 & 24, respectively. Players were considered to be elite, as they were registered players at an English Premier League Category One Football Academy or as professional players competing in the top two tiers of English football (English Football League Championship or English Premier League Divisions). Only data collected as a registered player at the football club between July 2013 and June 2021 was included from the under-9 to under-23 academy and senior male teams. Cases and controls were considered to experience similar training and environmental exposure as players within the same football teams through the academy and senior groups.

5.2.3 Injury surveillance

Injury data were retrospectively collected from injury monitoring databases prospectively recorded as part of normal working practice at an elite football club between July 2013 and June 2021. Medical staff, who were blinded to the genotype data, recorded all complaints requiring attention, in accordance with consensus guidance on injury definition and data collection procedures in football (Fuller, 2006). Time loss injury was defined as tissue damage or disruption to normal physical function occurring from football or related training activities (e.g., gym-based strength or field-based conditioning sessions), resulting in at least one day of missed training or competition (Bahr et al., 2020; Fuller, 2006). The time-loss injury incidence occurrence type (match or training) and date, onset (acute, overuse or mixed), mechanism (contact or non-contact), full-training return date, diagnosis and location using the Orchard Sports Injury and Illness Classification System (OSIICS) version 10 (Rae & Orchard, 2007) were all recorded. Injury severity was classified as minor (1-7 days), moderate (8-28 days) or severe (>28 days) based on the number of days between the

date of time loss injury and return date to full participation in training as recommended (Bahr et al., 2020). Incidence of injury was calculated as the number of injuries per 1000 hours of training and matches typically observed across the injury surveillance period for included participants. Illnesses were excluded from the analysis.

5.2.4 Injury phenotype classification

Time loss injuries were classified into seven groups for analysis: all injuries, non-contact injuries, non-contact muscle injuries, fracture injuries, apophysitis injuries, ligament injuries and tendon injuries. These injury classifications were selected to provide clear and accurately definable phenotypes, from which genetic associations may then guide applied decision making as recommended by Zondervan and Cardon (2007). Club medical staff performed initial injury diagnoses, with assessments verified by second opinion and / or scans were considered practically appropriate in the presence of reasonable doubt. All injuries include all time-loss injuries recorded as defined above. Non-contact injuries include all time-loss injuries occurring via non-contact mechanisms and non-contact muscle injuries those specific to muscle tissue. Non-contact mechanisms of injury were considered to have occurred without any direct or indirect contact from another player but included injuries resulting from the injured player striking a ball. Fracture injuries were always confirmed by X-ray scanning and included both contact and non-contact mechanisms. Ligament injuries included both contact and non-contact time loss injuries. Tendon injuries were time-loss injuries determined to be predominantly affecting tendinous tissue. However, tendinopathy injuries frequently display complex aetiology, which can result from referral of other symptoms, such as adductor tendinopathy from pubic bone oedema, for example. Apophysitis injuries were gradual onset and non-contact time-loss injuries. The apophysitis injury analysis included only participants who had injury surveillance data between the ages of 9 and 15 years of age, as apophysitis injury risk dramatically declines after this age as shown in chapter four and by others (de Loës, 1995; Hall et al., 2020; Read et al., 2018b).

5.2.5 Genetic variant selection and ethnicity dependent control

The selection of genetic variants as candidates for association with common injuries in elite male football for the present study followed a bottom-up mechanistic approach, using previously published recommendations (Lander & Schork, 1994; Romero et al., 2002). Initially, potential variants of interest were identified from systematic review and meta-analysis when appropriate, as shown in chapter three, of existing literature on genetic associations with musculoskeletal injury risk. These variants of interest were then scrutinised to establish if sufficient evidence of a functional consequence on genes with biological relevance to injury risk could be observed (Lander & Schork, 1994; Romero et al., 2002). This process identified nine genetic variants - *ACTN3* (rs1815739), *ACE* I/D (rs1799752), *COL1A1* (rs1800012), *COL1A2* (rs412777), *GDF5* (rs143383), *ESR1* (rs2234693), *MMP3* (rs679620), *COL5A1* (rs12722) and *VDR* (rs2228570) - as candidates for investigation with musculoskeletal tissue injuries in the present study as outlined in Figure 14. Table 18 outlines which genetic variants were considered to be candidates for investigation with the different tissue injury

classifications based upon the selection process and previous findings outlined in Table 16. The genetic variants included in the present study are all in linkage equilibrium and thus display random inheritance patterns with each other. Participant self-identified ethnicity was categorised into the 1000 Genomes Project five continental super population groups: African (AFR), ad mixed American (AMR), East Asian (EAS), European (EUR) and South Asian (SAS) (The 1000 Genomes Project Consortium, 2015) to account for the potential influence that unknown genetic variants associated with, and the frequency of which may vary by, ethnicity may have on the results. Self-reported ethnicity has been shown to be sufficient to account for ethnicity dependent genetic variance between groups (Tang et al., 2005).

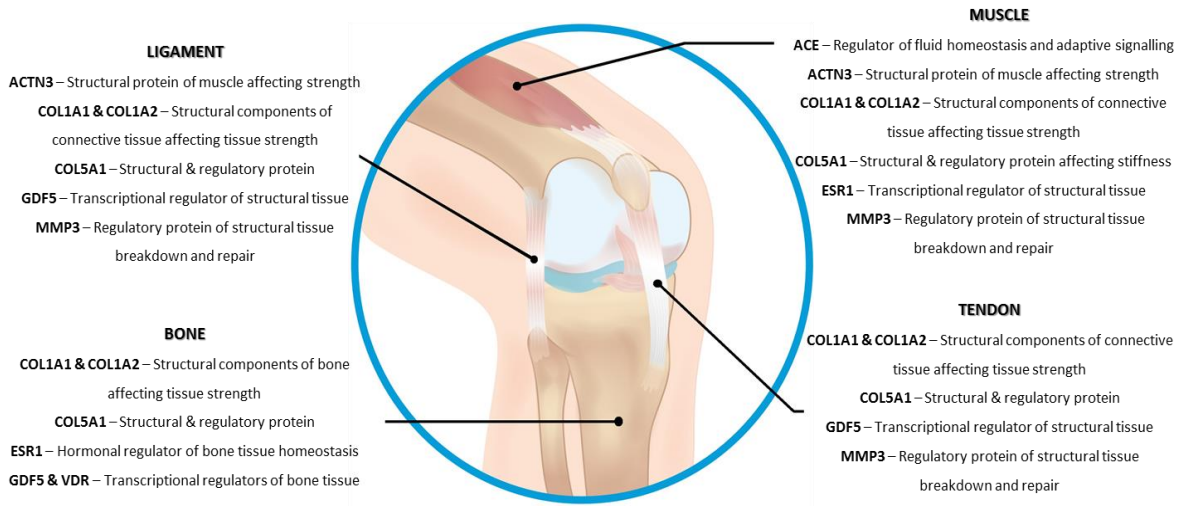

Figure 14. Influence of genetic candidate gene variants on common tissues of injury in football. Note: Source: Injurymap.

5.2.6 Genetic testing

Participants were asked to provide a 2 mL saliva sample into a collection vial (SalivaGene Collection Module II; Stratec Molecular GmbH) using the manufacturers guidelines under the supervision of the main investigator (ERM). This method was selected as a non-invasive method to provide ample genetic material for genotyping and more appropriate for use with children (Romero et al., 2002). A stabiliser solution provided by the manufacturer (SalivaGene Collection Module II; Stratec Molecular GmbH) was then mixed with the saliva sample, and the container subsequently sealed and labelled with an anonymous identification code, known only to the main investigator of the study. Sealed and labelled samples were then each placed into an individual grip seal plastic envelope, which was also labelled with the participant identification code and then transported to and stored within a Human Tissue Authority certified laboratory following certification guidelines at low temperatures. Stored saliva samples were transferred into 1.6 mL screw-top tubes, which were labelled and shipped to LGC genomics (LGC Limited, United Kingdom), who were blinded to the injury surveillance data, for DNA extraction and genotyping services.

Nine genetic variants (*COL1A1* rs1800012, *COL1A2* rs412777, *COL5A1* rs12722, *ACTN3* rs1815739, *ACE* rs1799752, *ESR1* rs2234693, *MMP3* rs679620, *VDR* rs2228570 and *GDF5* rs143383) were genotyped using KASP™ assays designed by LGC Genomics (LGC Limited, United

Kingdom). Allele frequencies along with primer sequences, which were validated by LGC Genomics prior to the analysis of experimental DNA samples, are shown in Table 17. KASP is a homogeneous, fluorescence-based genotyping technology based on allele-specific oligo extension and fluorescence resonance energy transfer for signal generation (Semagn et al., 2014). Variant specific primer and master mix were added to 10 ng of DNA according to the manufacturer's instructions. PCR amplification was performed (Veriti 384 thermal cycler, Applied Biosystems) and fluorescent signals read out using the 7900HT Fast Real-Time PCR System (Applied Biosystems) and converted to genotype information with the SDS program (version 2.3 Applied Biosystems). The call rate of all samples tested was greater than 94.6% and a negative control was included on each plate to check for non-specific amplification.

5.2.7 Data integrity and quality assurance

The genotype call rate of all samples tested was greater than 94.6% and a negative control was included on each plate to check for non-specific amplification. Genotypes were not automatically assigned to samples which failed to amplify consistently with the rest of the cluster and genotype classifications were confirmed by visual inspection of cluster plots by two of the main investigators (YM and ERM) using SNPViewer2 (KBiosciences UK Ltd., Hoddesdon, Herts, UK) as recommended (Semagn et al., 2014). The type of occurrence, OSIICS code, onset and mechanism of each injury were verified by cross-reference with injury rehabilitation / management notes and, where possible, video confirmation from match analysis archives. Any discrepancies in injury data, which could not be clarified with medical notes, were resolved and confirmed via consultation with medical staff who recorded initial assessment and/or injury management notes.

5.2.8 Statistical analysis

To examine the association between candidate genetic variants with injury incidence and injury frequency, the injury surveillance data were classified into dichotomous (injured / non-injured) and continuous (injuries per year) dependent variables for each tissue injury classification ($Injury_{ALL}$, $Injury_{NC}$, $Injury_{MUS}$, $Injury_{FRA}$, $Injury_{APO}$, $Injury_{LIG}$ and $Injury_{TEN}$). The association of genetic variants with tissue injury classifications, which were determined to demonstrate a plausible rationale for investigation - see table 18 as outlined above, were then evaluated with generalised linear models using PROC GLIMMIX in SAS OnDemand for Academics (SAS Institute Inc., Cary, NC, USA). Generalised linear models are robust against missing data and variations in observation duration between participants. Injury classification analyses with a total frequency greater than 0.3 per year were entered into the model as continuous dependent count variables, with genetic model entered as the independent variable using a Poisson distribution. The natural log of years under injury surveillance was also entered as an offset to account for the different observational periods between participants. All other injury classification variables were examined as binomially distributed (injured / non-injured) dependent variables, with the genetic comparison model entered as the independent variable, which did not account for different participant observation periods.

All injury classification and genetic variant analyses were compared the across four genetic association models previously outlined (Attia, 2003; Horita & Kaneko, 2015). Assuming the Y allele is a risk variant, these models were as follows: 1) the codominant model (XX *vs.* XY *vs*. YY) with three levels, whereby each genotype exerts a different (potentially additive) risk of injury; 2) the dominant model (XX *vs.* XY + YY), which assumes that possession of a single risk allele is sufficient to cause increased risk; 3) the recessive model (XX + XY *vs.* YY), which suggests possession of both risk alleles is required to cause an increased risk; and 4) the overdominance model (XX + YY *vs.* XY), which suggests that one of each allele causes the increased risk. The Hardy–Weinberg Equilibrium (HWE) of overall genotype frequencies (cases and controls) and those within each ethnicity group were evaluated using $\chi 2$ tests. Risk ratios (RR) and 99% confidence intervals (CIs) between genetic association comparison model groups were used to evaluate differences in injury risk. It has been argued that adjustment for multiple testing is not required for hypothesis-driven genetic association analyses, which are guided by previous observations (Gibbon et al., 2020). Therefore, statistical significance was set at $p < 0.05$. However, as the current study includes several independent genetic associations tests, a second significance level of $p \leq 0.01$ was chosen to indicate significant genetic associations with greater confidence than $p < 0.05$. Those tests significant at the $p<0.01$ level were considered to be significant with high confidence, whereas those at $p<0.05$ were deemed worthy of further investigation, in the future, given the limitations of our study. The *a priori* sample size calculations using G*Power software (Faul et al., 2007) suggested that 128 participants would be needed to achieve 80% statistical power using a two-tailed test of difference with medium effect size (Cohen's D = 0.5), significance set at 0.05, and minor allele frequency equal to 0.45.

Table 17. Candidate genetic variant 1000 Genomes Project allele frequencies and primer sequences used for KASP™ genotyping by LGC genomics Ltd.

Candidate Gene (Variant)	Allele X	1000G All Populations X Allele Frequency	Allele X Primer	Allele Y	1000G All Populations Y Allele Frequency	Allele Y Primer
ACE (rs1799752)	128bp Insertion	Unknown	CATTTCTCTAGACCTGCTGCCTT	Deletion	Unknown	CATTTCTCTAGACCTGCTGCCTA
ACTN3 (rs1815739)	C	60%	ACACTGCCCGAGGCTGACC	T	40%	CAACACTGCCCGAGGCTGACT
COL1A1 (rs1800012)	C	91%	CAGCCCTCATCCCGCCCC	A	9%	CAGCCCTCATCCCGCCCA
COL1A2 (rs412777)	A	67%	GGCTCTCCTCTTGCTCCAGCT	C	33%	GCTCTCCTCTTGCTCCAGCG
COL5A1 (rs12722)	C	65%	CACGCTCTGTCCACACCCAC	T	35%	CCACGCTCTGTCCACACCCAT
ESR1 (rs2234693)	T	55%	ATCTGAGTTCCAAATGTCCCAGCT	C	45%	CTGAGTTCCAAATGTCCCAGCC
GDF5 (rs143383)	G	55%	CGTTCTTGAAAGGAGAAAGCCG	A	45%	ACTCGTTCTTGAAAGGAGAAAGCCA
MMP3 (rs679620)	T	35%	GAAATATCTAGAAAACTACTACGACCTCA	C	65%	AATATCTAGAAAACTACTACGACCTCG
VDR (rs2228570)	G	67%	GAACACTTGAAGCTTGATATCTAGTTTC	A	33%	GAACACTTGAAGCTTGATATCTAGTTTG

Note: 1000G = 1000 genomes project phase 3.

Table 18. Candidate gene and injury association relationships to be investigated in elite male football players.

Candidate Gene (Variant)	All Injuries	Non-Contact Injuries	Fracture Injuries	Apophysitis Injuries	Non-Contact Muscle Injuries	Ligament Injuries	Tendon Injuries
ACE (rs1799752)	MAYBE	MAYBE			MAYBE		
ACTN3 (rs1815739)	MAYBE	MAYBE	MAYBE	MAYBE	LIKELY	LIKELY	
COL1A1 (rs1800012)	MAYBE	MAYBE	LIKELY	MAYBE		LIKELY	
COL1A2 (rs412777)	MAYBE	MAYBE	MAYBE	MAYBE			
COL5A1 (rs12722)	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	LIKELY	LIKELY
ESR1 (rs2234693)	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE		
GDF5 (rs143383)	MAYBE	MAYBE	LIKELY	MAYBE	MAYBE	LIKELY	LIKELY
MMP3 (rs679620)	MAYBE	MAYBE			MAYBE	MAYBE	LIKELY
VDR (rs2228570)	MAYBE	MAYBE	LIKELY	MAYBE			

Note: Maybe indicates gene-injury relationships which show limited previous association evidence and / or have a plausible mechanistic pathway; Likely indicates gene-injury relationships with considerable previous evidence of association in physically active participants.

5.3 Results

A descriptive summary of included participants and their injuries recorded during the injury surveillance period based on player age group at recruitment into the study is shown in table 19. The duration of each participant's injury surveillance period by age is shown in Figure 15, with 85 participants possessing injury surveillance data within the apophysitis risk period. The total injury surveillance period represents 227,273 elite football training and match exposure hours across 564 player seasons over an eight-year period with an estimated overall injury incidence of 2.9 injuries per 1000 hours. Of the 113 saliva samples collected, all but yielded sufficient DNA for genotyping analyses, the subsequent success rate of which ranged from 97.3 to 100% for included genetic variants. The genotype frequencies of included variants were all in HWE for all participants ($p > 0.18$) and both European ($p > 0.05$) and African ($p > 0.18$) ethnicity groups except for the *ESR1* rs2234693 variant ($p = 0.003$) for all participants as shown in Table 24 along with genotype frequencies based on participant self-identified ethnicity. A similar observation was made when examining the HWE of injured cases and non-injured controls across all included genetic variants and planned analyses, with the *ESR1* (rs2234693) variant the only variant not in HWE ($p < 0.05$) in the $Injury_{ALL}$ $Injury_{NC}$, $Injury_{FRA}$ and $Injury_{APO}$ analyses, as shown in tables 20 & 21.

All genetic association analyses can be seen in tables 20 to 23 below. Significant genetic associations were observed between the *COL1A2* (rs412777) variant and fracture risk with the codominant and recessive genetic models ($p \leq 0.01$). Pairwise comparisons indicated that those possessing the CC genotype had a significantly greater risk of fracture injury than both CA (RR = 1.88 [99% CI = 0.84-4.20], p = 0.04) and AA (RR = 2.67 [99% CI: 1.10-6.45], p = 0.004) individuals with the CA+AA genotypes combined being more than half as likely to sustain a fracture injury than C allele homozygotes (RR = 0.45 [99% CI: 0.22-0.94], p = 0.005). Non-contact injury frequency per season was significantly associated with the *MMP3* (rs679620) and *VDR* (rs2228570) SNPs using the recessive and overdominant models, both ($p \leq 0.01$). Possession of the CT or TT genotypes of the *MMP3* (rs679620) SNP was associated with a significantly increased frequency of non-contact injury per season compared with CC homozygotes (RR = 1.40 [99% CI: 1.03-1.92], p = 0.006). Homozygotes (AA or GG) of the *VDR* (rs2228570) SNP appeared to have a significantly reduced frequency of non-contact injuries when compared with AG heterozygotes (RR = 0.76 [99% CI 0.58-1.00], p = 0.01). Conversely, homozygotes (TT and CC) of the *COL5A1* (rs12722) SNP showed a significantly greater frequency of non-contact muscle injuries per season than heterozygotes (RR = 1.49 [99% CI 0.98-2.26], p = 0.01). These findings were supported by the following associations with other models at $p < 0.05$: *COL5A1* (rs12722) and non-contact muscle injury using the dominant and codominant models; *MMP3* (rs679620) and non-contact injury for the codominant and overdominant models; and *VDR* (rs2228570) with non-contact injuries with codominant and dominant models.

Genetic associations observed with $p < 0.05$ only, and which should be considered with less certainty include: *GDF5* (rs143383) and fracture with recessive, dominant and codominant models; *COL1A1* (rs1800012) and apophysitis injuries using the dominant model; *COL5A1* (rs12722) and tendon injury risk using a recessive model; and *VDR* (rs2228570) with all injuries using overdominant and dominant models. These results, which should be treated with less certainty, indicate that G

allele carriers (GG+GA) of the *GDF5* (rs143383) SNP are at reduced risk of fracture injury compared with A allele homozygotes (RR = 0.55 [99% CI 0.26-1.13], p = 0.03). Additionally, the risk of apophysitis injury appears to be increased for those with at least one A allele (AA+AC) of the *COL1A1* (rs1800012) SNP compared with C allele homozygotes (RR = 1.80 [99% CI 0.90-3.61], p = 0.03). Furthermore, the risk of tendon injury may be substantially reduced for those with the *COL5A1* (rs12722) TC or CC genotypes (RR = 0.41 [99% CI 0.15-1.10], p = 0.02). However, T allele carriers appear to be at reduced risk of non-contact muscle injury compared to C allele homozygotes (RR = 0.71 [99% CI 0.48-1.06], p = 0.03), with a significant difference observed between the protective effect of the *COL5A1* (rs12722) TC genotype compared to C homozygotes (RR = 0.62 [99% CI 0.39-0.99], p = 0.01). Finally, the *VDR* (rs2228570) homozygotes appeared to suffer significantly less time-loss injuries than AG heterozygotes (RR = 0.82 [99% CI 0.66-1.01], p = 0.02).

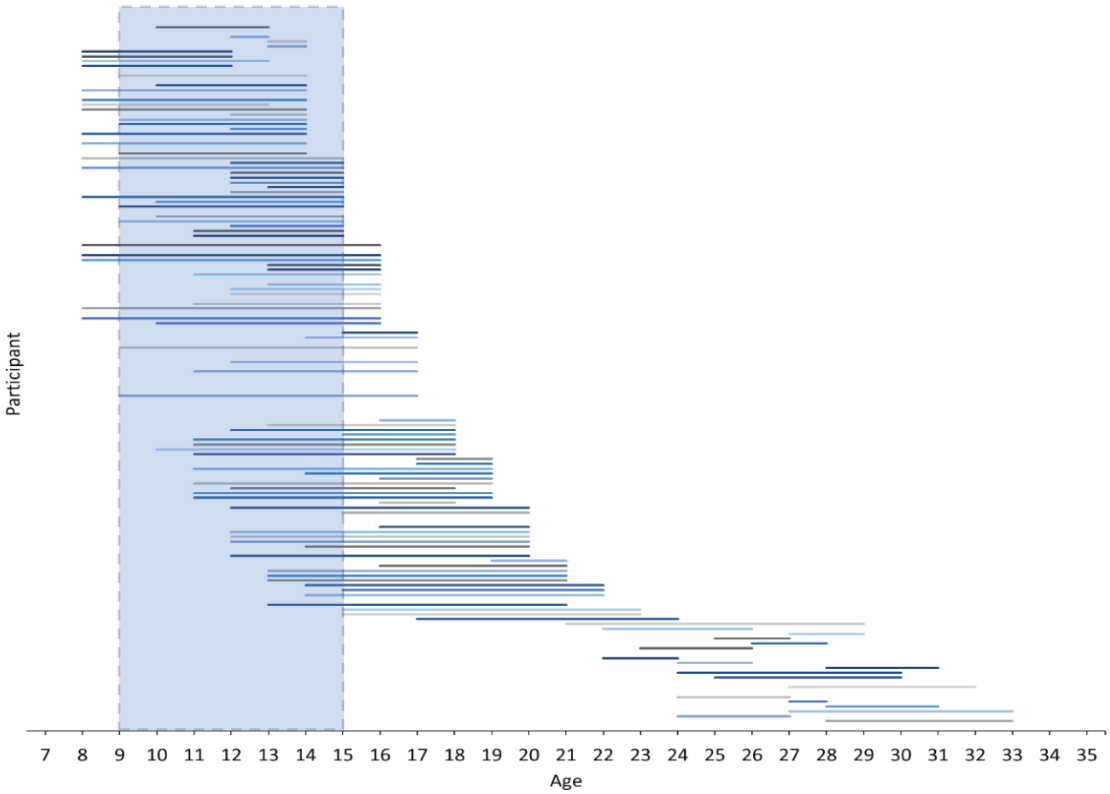

Figure 15. Injury surveillance age range per participant and the apophysitis injury risk period highlighted in blue.

Table 19. Descriptive statistics of included participants and injury summary by age group for elite male youth footballers.

Age group	Under-11	Under-12	Under-13	Under-14	Under-15	Under-16	Under-18	Under-23	First Team	Total
Included participants (*n*)	3	3	15	16	13	6	18	18	20	112
All injuries (*n*)	8	9	47	70	59	31	115	195	117	651
Injuries per year (*n*)	0.7 ± 0.6	1.1 ± 0.4	0.7 ± 0.4	1.0 ± 0.6	0.8 ± 0.4	1.0 ± 0.5	1.4 ± 0.8	1.7 ± 0.8	1.7 ± 1.2	1.2 ± 0.8
Non-contact injuries (*n*)	6	3	22	38	39	17	63	116	88	392
Non-contact injuries per year (*n*)	0.5 ± 0.4	0.4 ± 0.3	0.3 ± 0.3	0.6 ± 0.5	0.5 ± 0.4	0.5 ± 0.4	0.8 ± 0.5	1.0 ± 0.6	1.4 ± 1.2	0.8 ± 0.7
Non-contact muscle injuries (*n*)	4	2	5	19	13	3	22	62	56	186
Non-contact muscle injuries per year (*n*)	0.3 ± 0.3	0.3 ± 0.4	0.1 ± 0.1	0.3 ± 0.3	0.2 ± 0.2	0.1 ± 0.2	0.2 ± 0.3	0.5 ± 0.4	0.9 ± 0.9	0.4 ± 0.5
Tendon injuries (*n*)	0	0	1	0	0	2	11	11	9	34
Tendon injuries per year (*n*)	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.1	0.2 ± 0.3	0.1 ± 0.1	0.1 ± 0.3	0.1 ± 0.2
Joint injuries (*n*)	2	2	4	9	8	6	21	39	18	109
Joint injuries per year (*n*)	0.2 ± 0.3	0.3 ± 0.4	0.1 ± 0.1	0.1 ± 0.3	0.1 ± 0.2	0.2 ± 0.2	0.3 ± 0.3	0.3 ± 0.3	0.3 ± 0.5	0.2 ± 0.3
Fracture injuries (*n*)	1	2	9	11	7	3	6	10	3	52
Fracture injuries per year (*n*)	0.1 ± 0.1	0.2 ± 0.2	0.1 ± 0.2	0.2 ± 0.2	0.1 ± 0.2	0.1 ± 0.2	0.1 ± 0.1	0.1 ± 0.1	0.0 ± 0.1	0.1 ± 0.2
Apophysitis injuries (*n*)	1	1	9	10	18	7	9	4	0	59
Apophysitis injuries per year (*n*)	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.2	0.1 ± 0.2	0.2 ± 0.3	0.2 ± 0.2	0.1 ± 0.2	0.0 ± 0.1	0.0 ± 0.0	0.1 ± 0.2
Age at start of observation (years)	8 ± 0	10 ± 2	10 ± 2	11 ± 2	10 ± 2	12 ± 3	13 ± 2	14 ± 2	24 ± 4	14 ± 6
Age at end of observation (years)	12 ± 0	13 ± 0	14 ± 0	15 ± 0	17 ± 0	17 ± 0	18 ± 1	21 ± 1	28 ± 3	19 ± 5
Observation years (years)	4.0 ± 0.6	3.1 ± 1.7	4.5 ± 1.8	4.4 ± 1.9	5.6 ± 2.1	5.2 ± 2.6	5.3 ± 2.4	6.6 ± 1.7	3.6 ± 2.0	4.9 ± 2.2

Note: Data presented as means ± standard deviations of the measurements within the age group of participants at recruitment which may include injuries sustained in other age groups included in the observation period.

Table 20. Genetic association candidate gene analysis with all injuries & non-contact injuries observed in elite male footballers.

Candidate Gene Variant	Genetic Comparison Model	All Injuries			Non-Contact Injuries		
		Injured	Non-Injured	Incidence per year (99% CI)	Injured	Non-Injured	Incidence per year (99% CI)
ACE	Codominant						
(rs1799752)	II	0 (0)	0 (0)	N/A	0 (0)	0 (0)	N/A
	ID	26 (27)	25 (1)	1.30 (1.07-1.60)	26 (25)	21 (3)	0.76 (0.59-0.99)
	DD	74 (79)	75 (3)	1.14 (1.01-1.29)	74 (71)	79 (11)	0.70 (0.60-0.82)
	Dominant						
	II+ID	26 (27)	25 (1)	1.30 (1.07-1.60)	26 (25)	21 (3)	0.76 (0.59-0.99)
	DD	74 (79)	75 (3)	1.14 (1.01-1.29)	74 (71)	79 (11)	0.70 (0.60-0.82)
	Recessive						
	II	0 (0)	0 (0)	N/A	0 (0)	0 (0)	N/A
	ID+DD	100 (106)	100 (4)	N/A	100 (96)	100 (14)	N/A
	Overdominant						
	ID	26 (27)	25 (1)	1.30 (1.07-1.60)	26 (25)	21 (3)	0.76 (0.59-0.99)
	II+DD	74 (79)	75 (3)	1.14 (1.01-1.29)	74 (71)	79 (11)	0.70 (0.60-0.82)
ACTN3	Codominant						
(rs1815739)	TT (XX)	9 (10)	0 (0)	1.23 (0.86-1.77)	9 (9)	8 (1)	0.64 (0.39-1.06)
	TC (XR)	37 (40)	25 (1)	1.17 (0.99-1.39)	37 (36)	39 (5)	0.70 (0.57-0.87)
	CC (RR)	53 (57)	75 (3)	1.18 (1.03-1.36)	54 (53)	54 (7)	0.73 (0.61-0.88)
	Dominant						
	XX+XR	47 (50)	25 (1)	1.18 (1.02-1.38)	46 (45)	46 (6)	0.69 (0.57-0.84)
	RR	53 (57)	75 (3)	1.18 (1.03-1.36)	54 (53)	54 (7)	0.73 (0.61-0.88)
	Recessive						
	XX	9 (10)	0 (0)	1.23 (0.86-1.77)	9 (9)	8 (1)	0.64 (0.39-1.06)
	XR+RR	91 (97)	100 (4)	1.18 (1.06-1.31)	91 (89)	92 (12)	0.72 (0.63-0.83)
	Overdominant						
	XR	37 (40)	25 (1)	1.17 (0.99-1.39)	37 (36)	39 (5)	0.70 (0.57-0.87)
	XX+RR	63 (67)	75 (3)	1.19 (1.04-1.35)	63 (62)	61 (8)	0.72 (0.61-0.85)
COL1A1	Codominant						
(rs1800012)	TT	4 (4)	0 (0)	1.45 (0.87-2.40)	4 (4)	0 (0)	0.86 (0.45-1.65)
	TG	21 (22)	25 (1)	1.30 (1.05-1.60)	22 (21)	14 (2)	0.79 (0.60-1.04)
	GG	76 (81)	75 (3)	1.13 (1.00-1.28)	74 (72)	86 (12)	0.67 (0.58-0.79)
	Dominant						
	TT+TG	24 (26)	25 (1)	1.32 (1.08-1.60)	26 (25)	14 (2)	0.80 (0.62-1.03)
	GG	76 (81)	75 (3)	1.13 (1.00-1.28)	74 (72)	86 (12)	0.67 (0.58-0.79)
	Recessive						
	TT	4 (4)	0 (0)	1.45 (0.87-2.40)	4 (4)	0 (0)	0.86 (0.45-1.65)
	TG+GG	96 (103)	100 (4)	1.17 (1.05-1.30)	96 (93)	100 (14)	0.70 (0.61-0.80)
	Overdominant						
	TG	21 (22)	25 (1)	1.30 (1.05-1.60)	22 (21)	14 (2)	0.79 (0.60-1.04)
	TT+GG	79 (85)	75 (3)	1.14 (1.01-1.29)	78 (76)	86 (12)	0.68 (0.59-0.80)

Continued…

Table 20. Continued.

Candidate Gene Variant	Genetic Comparison Model	All Injuries			Non-Contact Injuries		
		Injured	Non-Injured	Incidence per year (99% CI)	Injured	Non-Injured	Incidence per year (99% CI)
COL1A2 (rs412777)	Codominant						
	CC	8 (8)	33 (1)	1.25 (0.90-1.76)	7 (7)	15 (2)	0.74 (0.48-1.15)
	CA	45 (48)	0 (0)	1.15 (0.97-1.36)	46 (44)	31 (4)	0.67 (0.54-0.84)
	AA	47 (50)	67 (2)	1.22 (1.06-1.41)	47 (45)	54 (7)	0.75 (0.62-0.90)
	Dominant						
	CC+CA	53 (56)	33 (1)	1.17 (1.01-1.36)	53 (51)	46 (6)	0.69 (0.56-0.84)
	AA	47 (50)	67 (2)	1.22 (1.06-1.41)	47 (45)	54 (7)	0.75 (0.62-0.90)
	Recessive						
	CC	8 (8)	33 (1)	1.25 (0.90-1.76)	7 (7)	15 (2)	0.74 (0.48-1.15)
	CA+AA	92 (98)	67 (2)	1.19 (1.07-1.33)	93 (89)	85 (11)	0.72 (0.62-0.82)
	Overdominant						
	CA	45 (48)	0 (0)	1.15 (0.97-1.36)	46 (44)	31 (4)	0.67 (0.54-0.84)
	CC+AA	55 (58)	100 (3)	1.23 (1.08-1.40)	54 (52)	69 (9)	0.75 (0.63-0.88)
COL5A1 (rs12722)	Codominant						
	TT	24 (26)	0 (0)	1.18 (0.97-1.44)	23 (22)	29 (4)	0.69 (0.54-0.90)
	TC	43 (46)	25 (1)	1.13 (0.96-1.33)	44 (43)	29 (4)	0.67 (0.54-0.84)
	CC	33 (35)	75 (3)	1.25 (1.05-1.49)	33 (32)	43 (6)	0.77 (0.62-0.97)
	Dominant						
	TT+TC	67 (72)	25 (1)	1.15 (1.01-1.31)	67 (65)	57 (8)	0.68 (0.58-0.80)
	CC	33 (35)	75 (3)	1.25 (1.05-1.49)	33 (32)	43 (6)	0.77 (0.62-0.97)
	Recessive						
	TT	24 (26)	0 (0)	1.18 (0.97-1.44)	23 (22)	29 (4)	0.69 (0.54-0.90)
	TC+CC	76 (81)	100 (4)	1.18 (1.05-1.33)	77 (75)	71 (10)	0.72 (0.62-0.84)
	Overdominant						
	TC	43 (46)	25 (1)	1.13 (0.96-1.33)	44 (43)	29 (4)	0.67 (0.54-0.84)
	TT+CC	57 (61)	75 (3)	1.22 (1.07-1.39)	56 (54)	71 (10)	0.74 (0.62-0.87)
ESR1 (rs2234693)	Codominant						
	CC	22 (23)*	0 (0)	1.24 (0.99-1.56)	20 (19)*	29 (4)	0.76 (0.57-1.02)
	CT	65 (68)*	100 (4)	1.19 (1.05-1.35)	65 (62)*	71 (10)	0.72 (0.62-0.85)
	TT	13 (14)*	0 (0)	1.02 (0.73-1.41)	15 (14)*	0 (0)	0.53 (0.34-0.83)
	Dominant						
	CC+CT	87 (91)	100 (4)	1.20 (1.08-1.34)	85 (81)	100 (14)	0.73 (0.64-0.84)
	TT	13 (14)	0 (0)	1.02 (0.73-1.41)	15 (14)	0 (0)	0.53 (0.34-0.83)
	Recessive						
	CC	22 (23)	0 (0)	1.24 (0.99-1.56)	20 (19)	29 (4)	0.76 (0.57-1.02)
	CT+TT	78 (82)	100 (4)	1.16 (1.03-1.31)	80 (76)	71 (10)	0.69 (0.60-0.81)
	Overdominant						
	CT	65 (68)	100 (4)	1.19 (1.05-1.35)	65 (62)	71 (10)	0.72 (0.62-0.85)
	CC+TT	35 (37)	0 (0)	1.16 (0.96-1.40)	35 (33)	29 (4)	0.68 (0.53-0.86)

Continued…

Table 20. Continued.

Candidate Gene Variant	Genetic Comparison Model	All Injuries			Non-Contact Injuries		
		Injured	Non-Injured	Incidence per year (99% CI)	Injured	Non-Injured	Incidence per year (99% CI)
GDF5	Codominant						
(rs143383)	AA	24 (25)	0 (0)	1.14 (0.92-1.43)	24 (23)	14 (2)	0.66 (0.50-0.89)
	AG	40 (42)	50 (2)	1.22 (1.05-1.43)	39 (37)	50 (7)	0.75 (0.61-0.91)
	GG	36 (38)	50 (2)	1.16 (0.97-1.40)	37 (35)	36 (5)	0.70 (0.55-0.89)
	Dominant						
	AA+AG	64 (67)	50 (2)	1.20 (1.06-1.36)	63 (60)	64 (9)	0.72 (0.61-0.85)
	GG	36 (38)	50 (2)	1.16 (0.97-1.40)	37 (35)	36 (5)	0.70 (0.55-0.89)
	Recessive						
	AA	24 (25)	0 (0)	1.14 (0.92-1.43)	24 (23)	14 (2)	0.66 (0.50-0.89)
	AG+GG	76 (80)	100 (4)	1.20 (1.07-1.35)	76 (72)	86 (12)	0.73 (0.63-0.85)
	Overdominant						
	AG	40 (42)	50 (2)	1.22 (1.05-1.43)	39 (37)	50 (7)	0.75 (0.61-0.91)
	AA+GG	60 (63)	50 (2)	1.16 (1.00-1.33)	61 (58)	50 (7)	0.69 (0.57-0.82)
MMP3	Codominant						
(rs679620)	CC	28 (30)	25 (1)	1.06 (0.87-1.30)	*25 (25)*	*43 (6)*	*0.56 (0.42-0.73)*
	CT	47 (51)	75 (3)	1.26 (1.09-1.45)	*49 (48)*	*43 (6)*	*0.80 (0.67-0.96)*
	TT	25 (27)	0 (0)	1.19 (0.96-1.47)	*25 (25)*	*14 (2)*	*0.74 (0.56-0.96)*
	Dominant						
	CC+CT	75 (81)	100 (4)	1.18 (1.05-1.33)	75 (73)	86 (12)	0.71 (0.61-0.82)
	TT	25 (27)	0 (0)	1.19 (0.96-1.47)	25 (25)	14 (2)	0.74 (0.56-0.96)
	Recessive						
	CC	28 (30)	25 (1)	1.06 (0.87-1.30)	***25 (25)***	***43 (6)***	***0.56 (0.42-0.73)***
	CT+TT	72 (78)	75 (3)	1.23 (1.09-1.39)	***75 (73)***	***57 (8)***	***0.78 (0.67-0.91)***
	Overdominant						
	CT	47 (51)	75 (3)	1.26 (1.09-1.45)	*49 (48)*	*43 (6)*	*0.80 (0.67-0.96)*
	CC+TT	53 (57)	25 (1)	1.12 (0.97-1.29)	*51 (50)*	*57 (8)*	*0.63 (0.52-0.77)*
VDR	Codominant						
(rs2228570)	AA	8 (9)	0 (0)	1.10 (1.75-1.62)	*7 (7)*	*14 (2)*	*0.60 (0.35-1.01)*
	AG	37 (39)	0 (0)	1.31 (1.12-1.54)	*37 (35)*	*29 (4)*	*0.83 (0.68-1.01)*
	GG	55 (58)	100 (4)	1.07 (0.93-1.25)	*56 (54)*	*57 (8)*	*0.64 (0.52-0.77)*
	Dominant						
	AA+AG	*45 (48)*	*0 (0)*	*1.28 (1.10-1.48)*	*44 (42)*	*43 (6)*	*0.79 (0.65-0.95)*
	GG	*55 (58)*	*100 (4)*	*1.07 (0.93-1.25)*	*56 (54)*	*57 (8)*	*0.64 (0.52-0.77)*
	Recessive						
	AA	8 (9)	0 (0)	1.10 (0.75-1.62)	7 (7)	14 (2)	0.60 (0.35-1.01)
	AG+GG	92 (97)	100 (4)	1.17 (1.05-1.31)	93 (89)	86 (12)	0.71 (0.62-0.82)
	Overdominant						
	AG	*37 (39)*	*0 (0)*	*1.31 (1.12-1.54)*	***37 (35)***	***29 (4)***	***0.83 (0.68-1.01)***
	AA+GG	*63 (67)*	*100 (4)*	*1.08 (0.94-1.24)*	***63 (61)***	***71 (10)***	***0.63 (0.53-0.76)***

Note: Genotype frequencies are expressed as a percentage with the number of observed participants (n) in parentheses. CI; Confidence interval. * Indicates genotype frequencies not in Hardy-Weinberg Equilibrium $p < 0.05$. Significant genetic associations with $p \leq 0.01$ are indicated in bold italics and $p < 0.05$ in italics.

Table 21. Genetic association candidate gene analysis with fracture & apophysitis injuries observed in elite male footballers.

Candidate Gene Variant	Genetic Comparison Model	Fracture Injuries			Apophysitis Injuries		
		Injured	Non-Injured	% Chance of Injury (99% CI)	Injured	Non-Injured	% Chance of Injury (99% CI)
ACTN3 (rs1815739)	Codominant						
	TT (XX)	14 (5)	7 (5)	0.50 (0.22-1.15)	9 (3)	8 (4)	0.43 (0.14-1.36)
	TC (XR)	39 (14)	36 (27)	0.34 (0.19-0.60)	34 (11)	37 (19)	0.37 (0.19-0.69)
	CC (RR)	47 (17)	57 (43)	0.28 (0.17-0.49)	56 (18)	56 (29)	0.38 (0.24-0.62)
	Dominant						
	XX+XR	53 (19)	43 (32)	0.37 (0.23-0.60)	44 (14)	44 (23)	0.38 (0.22-0.66)
	RR	47 (17)	57 (43)	0.28 (0.17-0.49)	56 (18)	56 (29)	0.38 (0.24-0.62)
	Recessive						
	XX	14 (5)	7 (5)	0.50 (0.22-1.15)	9 (3)	8 (4)	0.43 (0.14-1.36)
	XR+RR	86 (31)	93 (70)	0.31 (0.21-0.45)	91 (29)	92 (48)	0.38 (0.26-0.55)
	Overdominant						
	XR	39 (14)	36 (27)	0.34 (0.19-0.60)	34 (11)	37 (19)	0.39 (0.25-0.61)
	XX+RR	61 (22)	64 (48)	0.31 (0.20-0.50)	66 (21)	63 (33)	0.37 (0.19-0.69)
COL1A1 (rs1800012)	Codominant						
	TT	0 (0)	5 (4)	0.00 (0.00-0.00)	6 (2)	2 (1)	0.67 (0.23-1.96)
	TG	23 (8)	20 (15)	0.35 (0.16-0.74)	31 (10)	15 (8)	0.56 (0.32-0.97)
	GG	77 (27)	75 (57)	0.32 (0.21-0.49)	63 (20)	83 (43)	0.32 (0.20-0.52)
	Dominant						
	TT+TG	23 (8)	25 (19)	0.30 (0.14-0.64)	*38 (12)*	*17 (9)*	*0.57 (0.35-0.94)*
	GG	77 (27)	75 (57)	0.32 (0.21-0.49)	*63 (20)*	*83 (43)*	*0.32 (0.20-0.52)*
	Recessive						
	TT	0 (0)	5 (4)	0.00 (0.00-0.00)	6 (2)	2 (1)	0.67 (0.23-1.96)
	TG+GG	100 (35)	95 (72)	0.33 (0.23-0.47)	94 (30)	98 (51)	0.37 (0.25-0.54)
	Overdominant						
	TG	23 (8)	20 (15)	0.35 (0.16-0.74)	31 (10)	15 (8)	0.56 (0.32-0.97)
	TT+GG	77 (27)	80 (61)	0.31 (0.20-0.47)	69 (22)	85 (44)	0.33 (0.21-0.53)
COL1A2 (rs412777)	Codominant						
	CC	***17 (6)***	***4 (3)***	***0.67 (0.36-1.24)***	10 (3)	12 (6)	0.33 (0.10-1.16)
	CA	***47 (17)***	***43 (31)***	***0.35 (0.21-0.59)***	36 (11)	45 (23)	0.32 (0.17-0.62)
	AA	***36 (13)***	***53 (39)***	***0.25 (0.13-0.47)***	55 (17)	43 (22)	0.44 (0.27-0.71)
	Dominant						
	CC+CA	64 (23)	47 (34)	0.40 (0.26-0.62)	45 (14)	57 (29)	0.33 (0.18-0.58)
	AA	36 (13)	53 (39)	0.25 (0.13-0.47)	55 (17)	43 (22)	0.44 (0.27-0.71)
	Recessive						
	CC	***17 (6)***	***4 (3)***	***0.67 (0.36-1.24)***	12 (3)	12 (6)	0.33 (0.10-1.16)
	CA+AA	***83 (30)***	***96 (70)***	***0.30 (0.20-0.45)***	88 (28)	88 (45)	0.38 (0.26-0.57)
	Overdominant						
	CA	47 (17)	43 (31)	0.35 (0.21-0.59)	36 (11)	45 (23)	0.32 (0.17-0.62)
	CC+AA	53 (19)	57 (42)	0.31 (0.19-0.51)	64 (20)	55 (28)	0.42 (0.27-0.65)

Continued…

Table 21. Continued.

Candidate Gene Variant	Genetic Comparison Model	Fracture Injuries			Apophysitis Injuries		
		Injured	Non-Injured	% Chance of Injury (99% CI)	Injured	Non-Injured	% Chance of Injury (99% CI)
COL5A1 (rs12722)	Codominant						
	TT	26 (9)	22 (17)	0.35 (0.17-0.70)	19 (6)	27 (14)	0.30 (0.12-0.74)
	TC	40 (14)	43 (33)	0.30 (0.17-0.54)	50 (16)	37 (19)	0.46 (0.28-0.74)
	CC	34 (12)	34 (26)	0.32 (0.17-0.59)	31 (10)	37 (19)	0.34 (0.18-0.68)
	Dominant						
	TT+TC	66 (23)	66 (50)	0.32 (0.20-0.50)	69 (22)	64 (33)	0.40 (0.26-0.62)
	CC	34 (12)	34 (26)	0.32 (0.17-0.59)	31 (10)	37 (19)	0.34 (0.18-0.68)
	Recessive						
	TT	26 (9)	22 (17)	0.35 (0.17-0.70)	19 (6)	27 (14)	0.30 (0.12-0.74)
	TC+CC	74 (26)	78 (59)	0.31 (0.19-0.47)	81 (26)	73 (38)	0.41 (0.27-0.61)
	Overdominant						
	TC	40 (14)	43 (33)	0.30 (0.17-0.54)	50 (16)	37 (19)	0.46 (0.28-0.74)
	TT+CC	60 (21)	57 (43)	0.33 (0.21-0.52)	50 (16)	63 (33)	0.33 (0.19-0.56)
ESR1 (rs2234693)	Codominant						
	CC	17 (6)	23 (17)*	0.26 (0.10-0.66)	27 (8)	17 (9)*	0.47 (0.24-0.93)
	CT	69 (24)	65 (48)*	0.33 (0.22-0.52)	57 (17)	71 (37)*	0.31 (0.19-0.53)
	TT	14 (5)	12 (9)*	0.36 (0.14-0.91)	17 (5)	12 (6)*	0.45 (0.19-1.09)
	Dominant						
	CC+CT	86 (30)	88 (65)	0.32 (0.21-0.47)	83 (25)	89 (46)	0.35 (0.23-0.54)
	TT	14 (5)	12 (9)	0.36 (0.14-0.91)	17 (5)	12 (60	0.45 (0.19-1.09)
	Recessive						
	CC	17 (6)	23 (17)	0.26 (0.10-0.66)	27 (8)	17 (9)	0.47 (0.24-0.93)
	CT+TT	83 (29)	77 (57)	0.34 (0.23-0.50)	73 (22)	83 (43)	0.34 (0.21-0.53)
	Overdominant						
	CT	69 (24)	65 (48)	0.33 (0.22-0.52)	57 (17)	71 (37)	0.31 (0.19-0.53)
	CC+TT	31 (26)	35 (26)	0.30 (0.15-0.58)	43 (13)	29 (15)	0.46 (0.27-0.79)
GDF5 (rs143383)	Codominant						
	AA	*35 (12)*	*17 (13)*	*0.48 (0.28-0.83)*	23 (7)	26 (13)	0.35 (0.16-0.78)
	AG	*44 (15)*	*39 (29)*	*0.34 (0.20-0.59)*	48 (15)	37 (19)	0.44 (0.27-0.73)
	GG	*21 (7)*	*44 (33)*	*0.18 (0.07-0.43)*	29 (9)	37 (19)	0.32 (0.16-0.66)
	Dominant						
	AA+AG	*79 (27)*	*56 (42)*	*0.39 (0.26-0.58)*	71 (22)	63 (32)	0.41 (0.26-0.63)
	GG	*21 (7)*	*44 (33)*	*0.18 (0.07-0.43)*	29 (9)	37 (19)	0.18 (0.07-0.43)
	Recessive						
	AA	*35 (12)*	*17 (13)*	*0.48 (0.28-0.83)*	23 (7)	26 (13)	0.48 (0.28-0.83)
	AG+GG	*65 (22)*	*83 (62)*	*0.26 (0.16-0.42)*	77 (24)	74 (38)	0.39 (0.25-0.59)
	Overdominant						
	AG	44 (15)	39 (29)	0.34 (0.20-0.59)	48 (15)	37 (19)	0.44 (0.27-0.73)
	AA+GG	56 (19)	61 (46)	0.29 (0.18-0.48)	52 (16)	63 (32)	0.33 (0.25-0.59)

Continued…

Table 21. Continued.

Candidate Gene Variant	Genetic Comparison Model	Fracture Injuries			Apophysitis Injuries		
		Injured	Non-Injured	% Chance of Injury (99% CI)	Injured	Non-Injured	% Chance of Injury (99% CI)
VDR (rs2228570)	Codominant						
	AA	8 (3)	8 (6)	0.33 (0.10-1.15)	0 (0)	13 (7)	0.00 (0.00-0.00)
	AG	39 (14)	34 (25)	0.36 (0.20-0.63)	36 (11)	28 (15)	0.42 (0.23-0.79)
	GG	53 (19)	58 (43)	0.31 (0.19-0.51)	64 (20)	59 (31)	0.39 (0.25-0.61)
	Dominant						
	AA+AG	47 (17)	42 (31)	0.35 (0.21-0.59)	36 (11)	41 (22)	0.33 (0.17-0.64)
	GG	53 (19)	58 (43)	0.31 (0.19-0.51)	64 (20)	59 (31)	0.39 (0.25-0.61)
	Recessive						
	AA	8 (3)	8 (6)	0.33 (0.10-1.15)	0 (0)	13 (7)	0.00 (0.00-0.00)
	AG+GG	92 (33)	92 (68)	0.31 (0.19-0.51)	100 (31)	87 (46)	0.40 (0.28-0.58)
	Overdominant						
	AG	39 (14)	34 (25)	0.36 (0.20-0.63)	36 (11)	28 (15)	0.42 (0.23-0.79)
	AA+GG	61 (49)	66 (49)	0.31 (0.19-0.49)	64 (20)	72 (38)	0.34 (0.21-0.56)

Note: Genotype frequencies are expressed as a percentage with the number of observed participants (n) in parentheses. CI; Confidence interval. * Indicates genotype frequencies not in Hardy-Weinberg Equilibrium $p < 0.05$. Significant genetic associations with $p \leq 0.01$ are indicated in bold italics and $p < 0.05$ in italics.

Table 22. Genetic association candidate gene analysis with ligament & tendon injuries observed in elite male footballers.

Candidate Gene Variant	Genetic Comparison Model	Ligament Injuries			Tendon Injuries		
		Injured	Non-Injured	% Chance of Injury (99% CI)	Injured	Non-Injured	% Chance of Injury (99% CI)
ACTN3 (rs1815739)	Codominant						
	TT (XX)	9 (5)	9 (5)	0.50 (0.22-1.15)			
	TC (XR)	37 (21)	37 (20)	0.51 (0.34-0.76)			
	CC (RR)	54 (31)	54 (29)	0.52 (0.37-0.72)			
	Dominant						
	XX+XR	46 (26)	46 (25)	0.51 (0.36-0.73)			
	RR	54 (31)	54 (29)	0.52 (0.37-0.72)		Analysis not planned	
	Recessive						
	XX	9 (5)	9 (5)	0.50 (0.22-1.15)			
	XR+RR	91 (52)	91 (49)	0.51 (0.40-0.66)			
	Overdominant						
	XR	37 (21)	37 (20)	0.51 (0.34-0.76)			
	XX+RR	63 (36)	63 (34)	0.51 (0.38-0.70)			
COL1A1 (rs1800012)	Codominant						
	TT	7 (4)	0 (0)	1.00 (0.74-1.24)			
	TG	21 (12)	20 (11)	0.53 (0.31-0.88)			
	GG	71 (40)	80 (44)	0.48 (0.36-0.65)			
	Dominant						
	TT+TG	29 (16)	20 (11)	0.59 (0.39-0.90)			
	GG	71 (40)	80 (44)	0.48 (0.31-0.88)		Analysis not planned	
	Recessive						
	TT	7 (4)	0 (0)	1.00 (0.74-1.24)			
	TG+GG	93 (52)	100 (55)	0.49 (0.38-0.63)			
	Overdominant						
	TG	21 (12)	20 (11)	0.53 (0.31-0.88)			
	TT+GG	79 (44)	80 (44)	0.50 (0.38-0.66)			
COL5A1 (rs12722)	Codominant						
	TT	25 (14)	22 (12)	0.54 (0.33-0.87)	43 (9)	19 (17)	0.35 (0.17-0.70)
	TC	43 (24)	42 (23)	0.51 (0.35-0.74)	33 (7)	44 (40)	0.15 (0.06-0.37)
	CC	32 (18)	36 (20)	0.47 (0.30-0.74)	24 (5)	37 (33)	0.13 (0.04-0.39)
	Dominant						
	TT+TC	68 (38)	64 (35)	0.52 (0.39-0.70)	76 (16)	63 (57)	0.22 (0.12-0.39)
	CC	32 (18)	36 (20)	0.47 (0.30-0.74)	24 (5)	37 (33)	0.13 (0.04-0.39)
	Recessive						
	TT	25 (14)	22 (12)	0.54 (0.33-0.87)	*43 (9)*	*19 (17)*	*0.35 (0.17-0.70)*
	TC+CC	75 (42)	78 (43)	0.49 (0.37-0.66)	*57 (12)*	*81 (73)*	*0.14 (0.07-0.28)*
	Overdominant						
	TC	43 (24)	42 (23)	0.51 (0.35-0.74)	33 (7)	44 (40)	0.15 (0.06-0.37)
	TT+CC	57 (32)	58 (32)	0.50 (0.36-0.69)	66 (14)	56 (50)	0.22 (0.12-0.41)

Continued…

Table 21. Continued.

Candidate Gene Variant	Genetic Comparison Model	Ligament Injuries			Tendon Injuries		
		Injured	Non-Injured	% Chance of Injury (99% CI)	Injured	Non-Injured	% Chance of Injury (99% CI)
GDF5	Codominant						
(rs143383)	AA	25 (14)	21 (11)	0.56 (0.35-0.89)	19 (4)	24 (21)	0.16 (0.05-0.53)
	AG	41 (23)	40 (21)	0.52 (0.36-0.76)	43 (9)	40 (35)	0.20 (0.09-0.45)
	GG	34 (19)	40 (21)	0.48 (0.31-0.73)	38 (8)	36 (32)	0.20 (0.09-0.46)
	Dominant						
	AA+AG	66 (37)	60 (32)	0.54 (0.40-0.72)	62 (13)	64 (56)	0.19 (0.10-0.36)
	GG	34 (19)	40 (21)	0.48 (0.31-0.73)	38 (8)	36 (32)	0.20 (0.09-0.46)
	Recessive						
	AA	25 (14)	21 (11)	0.56 (0.35-0.89)	19 (4)	24 (21)	0.16 (0.05-0.53)
	AG+GG	75 (42)	79 (42)	0.50 (0.38-0.67)	81 (17)	76 (67)	0.20 (0.11-0.36)
	Overdominant						
	AG	41 (23)	40 (21)	0.52 (0.36-0.76)	43 (9)	40 (35)	0.20 (0.09-0.45)
	AA+GG	59 (33)	60 (32)	0.51 (0.37-0.70)	57 (12)	60 (53)	0.18 (0.09-0.37)
MMP3	Codominant						
(rs679620)	CC	26 (15)	29 (16)	0.48 (0.30-0.79)	14 (3)	31 (28)	0.10 (0.02-0.41)
	CT	49 (28)	47 (26)	0.52 (0.37-0.73)	57 (12)	46 (42)	0.22 (0.11-0.43)
	TT	25 (14)	24 (13)	0.52 (0.32-0.84)	29 (6)	23 (21)	0.22 (0.09-0.57)
	Dominant						
	CC+CT	75 (43)	76 (42)	0.51 (0.38-0.67)	71 (15)	77 (70)	0.18 (0.10-0.33)
	TT	25 (14)	24 (13)	0.52 (0.32-0.84)	29 (6)	23 (21)	0.22 (0.09-0.57)
	Recessive						
	CC	26 (15)	29 (16)	0.48 (0.30-0.79)	14 (3)	31 (28)	0.10 (0.02-0.41)
	CT+TT	74 (42)	71 (39)	0.52 (0.39-0.69)	86 (18)	69 (63)	0.22 (0.13-0.38)
	Overdominant						
	CT	49 (28)	47 (26)	0.52 (0.37-0.73)	57 (12)	46 (42)	0.22 (0.11-0.43)
	CC+TT	51 (29)	53 (29)	0.50 (0.35-0.71)	43 (9)	54 (49)	0.16 (0.07-0.35)

Note: Genotype frequencies are expressed as a percentage with the number of observed participants (n) in parentheses. CI; Confidence interval. * Indicates genotype frequencies not in Hardy-Weinberg Equilibrium $p < 0.05$. Significant genetic associations with $p \leq 0.01$ are indicated in bold italics and $p < 0.05$ in italics.

Table 23. Genetic association candidate gene analysis with non-contact muscle injuries observed in elite male footballers.

Candidate Gene Variant	Genetic Comparison Model	Non-Contact Muscle Injuries		
		Injured	Non-Injured	Incidence per year (99% CI)
ACE (rs1799752)	Codominant			
	II	0 (0)	0 (0)	N/A
	ID	28 (21)	19 (7)	0.33 (0.22-0.50)
	DD	72 (53)	81 (29)	0.35 (0.28-0.43)
	Dominant			
	II+ID	28 (21)	19 (7)	0.33 (0.22-0.50)
	DD	72 (53)	81 (29)	0.35 (0.28-0.43)
	Recessive			
	II	0 (0)	0 (0)	N/A
	ID+DD	100 (74)	100 (29)	N/A
	Overdominant			
	ID	28 (21)	19 (7)	0.33 (0.22-0.50)
	II+DD	72 (53)	81 (29)	0.35 (0.28-0.43)
ACTN3 (rs1815739)	Codominant			
	TT (XX)	7 (5)	14 (5)	0.19 (0.08-0.48)
	TC (XR)	39 (29)	33 (12)	0.36 (0.27-0.49)
	CC (RR)	55 (41)	53 (19)	0.35 (0.27-0.45)
	Dominant			
	XX+XR	45 (34)	47 (17)	0.33 (0.25-0.44)
	RR	55 (41)	53 (19)	0.35 (0.27-0.45)
	Recessive			
	XX	7 (5)	14 (5)	0.19 (0.08-0.48)
	XR+RR	93 (70)	86 (31)	0.35 (0.29-0.43)
	Overdominant			
	XR	39 (29)	33 (12)	0.36 (0.27-0.49)
	XX+RR	61 (46)	67 (24)	0.33 (0.25-0.42)
COL5A1 (rs12722)	Codominant			
	TT	*22 (16)*	*27 (10)*	*0.35 (0.25-0.51)*
	TC	*42 (31)*	*43 (16)*	*0.26 (0.18-0.37)*
	CC	*36 (27)*	*30 (11)*	*0.42 (0.31-0.57)*
	Dominant			
	TT+TC	*64 (47)*	*70 (26)*	*0.30 (0.23-0.38)*
	CC	*36 (27)*	*30 (11)*	*0.42 (0.31-0.57)*
	Recessive			
	TT	22 (16)	27 (10)	0.35 (0.25-0.51)
	TC+CC	78 (58)	73 (27)	0.33 (0.26-0.42)
	Overdominant			
	TC	*42 (31)*	*43 (16)*	*0.26 (0.18-0.37)*
	TT+CC	*58 (43)*	*57 (21)*	*0.39 (0.31-0.49)*
ESR1 (rs2234693)	Codominant			
	CC	21 (15)	22 (8)	0.39 (0.26-0.58)
	CT	64 (47)	69 (25)	0.35 (0.28-0.44)
	TT	15 (11)	8 (3)	0.25 (0.13-0.48)
	Dominant			
	CC+CT	85 (62)	92 (33)	0.36 (0.29-0.44)
	TT	15 (11)	8 (3)	0.25 (0.13-0.48)
	Recessive			
	CC	21 (15)	22 (8)	0.39 (0.26-0.58)
	CT+TT	79 (58)	78 (28)	0.33 (0.27-0.42)
	Overdominant			
	CT	64 (47)	69 (25)	0.35 (0.28-0.44)
	CC+TT	36 (26)	31 (11)	0.33 (0.24-0.47)
GDF5 (rs143383)	Codominant			
	AA	19 (14)	31 (11)	0.26 (0.17-0.42)
	AG	44 (32)	33 (12)	0.35 (0.26-0.47)
	GG	37 (27)	36 (13)	0.37 (0.27-0.51)
	Dominant			
	AA+AG	63 (46)	64 (23)	0.32 (0.25-0.41)
	GG	37 (27)	36 (13)	0.37 (0.27-0.51)
	Recessive			
	AA	19 (14)	31 (11)	0.26 (0.17-0.42)
	AG+GG	81 (59)	69 (25)	0.36 (0.29-0.45)
	Overdominant			
	AG	44 (32)	33 (12)	0.35 (0.26-0.47)
	AA+GG	56 (41)	67 (24)	0.33 (0.25-0.43)

Continued...

Table 23. Continued.

Candidate Gene Variant	Genetic Comparison Model	Non-Contact Muscle Injuries		
		Injured	Non-Injured	Incidence per year (99% CI)
MMP3 (rs679620)	Codominant			
	CC	25 (19)	32 (12)	0.28 (0.19-0.42)
	CT	49 (37)	46 (17)	0.37 (0.29-0.49)
	TT	25 (19)	22 (8)	0.34 (0.23-0.50)
	Dominant			
	CC+CT	75 (56)	78 (29)	0.34 (0.27-0.42)
	TT	25 (19)	22 (8)	0.34 (0.23-0.50)
	Recessive			
	CC	25 (19)	32 (12)	0.28 (0.19-0.42)
	CT+TT	75 (56)	68 (25)	0.36 (0.29-0.45)
	Overdominant			
	CT	49 (37)	46 (17)	0.37 (0.29-0.49)
	CC+TT	51 (38)	54 (20)	0.31 (0.23-0.40)

Note: Genotype frequencies are expressed as a percentage with the number of observed participants (n) in parentheses. CI; Confidence interval. * Indicates genotype frequencies not in Hardy-Weinberg Equilibrium $p < 0.05$. Significant genetic associations with $p \leq 0.01$ are indicated in bold italics and $p < 0.05$ in italics.

Table 24. Participant self-reported ethnicity 1000 Genomes Project super population genotype frequency distribution.

Candidate Gene Variant Genotype	1000G Super Population Ethnicity Code						
	African		European		Mixed		
					African & European	All Participants	
	Observed	1000G	Observed	1000G	Observed	Observed	1000G
ACE (rs1799752)							
II	0 (0)	Unknown	0 (0)	Unknown	0 (0)	0 (0)	Unknown
ID	26 (7)		25 (16)		31 (5)	26 (28)	
DD	74 (20)		75 (48)		69 (20)	74 (82)	
ACTN3 (rs1815739)							
TT (XX)	0 (0)	1	16 (10)	18	0 (0)	9 (10)	18
TC (XR)	14 (4)	22	45 (29)	51	44 (7)	37 (41)	44
CC (RR)	86 (24)	78	40 (25)	31	56 (9)	54 (60)	38
COL1A1 (rs1800012)							
TT	0 (0)	1	6 (4)	4	0 (0)	4 (4)	2
TG	7 (2)	12	29 (18)	29	19 (3)	21 (23)	15
GG	93 (27)	88	65 (41)	67	81 (13)	76 (84)	83
COL1A2 (rs412777)							
CC	7 (2)	11	8 (5)	16	7 (1)	8 (9)	12
CA	21 (6)	46	56 (35)	45	40 (6)	44 (48)	43
AA	71 (20)	43	37 (23)	39	53 (8)	48 (52)	46
COL5A1 (rs12722)							
TT	3 (1)	1	33 (21)	36	19 (3)	23 (26)	16
TC	38 (11)	26	54 (34)	46	13 (2)	42 (47)	38
CC	59 (17)	72	13 (8)	19	69 (11)	34 (38)	46
ESR1 (rs2234693)							
CC	25 (7)	31	21 (13)	18	20 (3)	21 (23)*	20
CT	64 (18)	52	65 (41)	50	67 (10)	66 (72)*	49
TT	11 (3)	17	14 (9)	33	13 (2)	13 (14)*	31
GDF5 (rs143383)							
AA	7 (2)	0	34 (21)	42	6 (1)	23 (25)	29
AG	17 (5)	6	51 (31)	42	44 (7)	40 (44)	33
GG	76 (22)	94	15 (9)	16	50 (8)	37 (40)	38
MMP3 (rs679620)							
CC	52 (15)	45	17 (11)	28	25 (4)	28 (31)	43
CT	41 (12)	45	48 (31)	50	56 (9)	48 (54)	44
TT	7 (2)	11	34 (22)	22	19 (3)	24 (27)	13
VDR (rs2228570)							
AA	0 (0)	4	11 (7)	16	6 (1)	8 (9)	13
AG	25 (7)	31	46 (29)	44	19 (3)	36 (39)	41
GG	75 (21)	66	43 (27)	40	75 (12)	56 (62)	47
Participants (n)	29		64		16	112	

Note: Genotype frequencies are expressed as a percentage with the number of observed participants (n) in parentheses. 1000G represents genotype frequency data from the 1000 Genomes project phase 3. Data are not presented for the East Asian and Mixed East Asian & Ad Mixed American participants (n=3) to protect data anonymity. * Indicates genotype frequencies not in Hardy-Weinberg Equilibrium $p < 0.05$.

5.4 Discussion

The main findings of the present study were the genetic associations observed between the *COL1A2* (rs412777) SNP with fracture risk, the *MMP3* (rs679620) and *VDR* (rs2228570) SNPs with non-contact injury and the *COL5A1* (rs12722) SNP with non-contact muscle injury in elite male football players ($p \leq 0.01$). Genetic associations worthy of further investigation with injury were also indicated between the *GDF5* (rs143383) SNP with fracture risk, the *COL5A1* (rs12722) SNP with tendon injury risk and the *COL1A1* (rs1800012) SNP with apophysitis injury risk ($0.01 < p \leq 0.05$). Both the *GDF5* (rs143383) and *COL5A1* (rs12722) SNPs have been previously associated with fracture (Zhao et al., 2016) and tendon (Altinisik et al., 2015; Mokone et al., 2006; Pabalan et al., 2018; September et al., 2009) injury risk, respectively. These findings suggest that genetic associations are present but, due to the frequency of observations, may be considered with less certainty than those significant with $p \leq 0.01$. It has been argued that hypothesis-driven genetic association studies replicating previously established findings do not need to adjust for multiple testing (Gibbon et al., 2020). However, as the present study completed many independent, hypothesis-driven tests of association, which would not individually need adjustment for multiple testing (Gibbon et al., 2020), it was considered appropriate to acknowledge the differences in confidence when asserting genetic associations with $p < 0.05$ and $p \leq 0.01$. Therefore, the genetic associations observed in the present study between the *GDF5* (rs143383), *COL5A1* (rs12722) and *COL1A1* (rs1800012) SNPs with fracture, tendon and apophysitis injury risk should be considered as potentially influential, but with some caution.

The *COL1A1* (rs1800012) SNP has not been previously associated with apophysitis injury risk but the T allele has been associated with increased risk of fracture during childhood (Blades et al., 2010) and reduced BMD in early puberty (Suuriniemi et al., 2006). The T allele also been associated with increased risk of osteoporotic fracture in the elderly (Mann & Ralston, 2003), although, this has not been replicated in young physically active adults (Cosman et al., 2013; Korvala et al., 2010; Varley et al., 2018), and there may even be a protective effect against fracture in females only, as shown in chapter three. A protective effect of the *COL1A1* (rs1800012) T allele has also been repeatedly observed with reduced risk of ACL rupture in young physically active participants (Ficek et al., 2013; Gibbon et al., 2020; Khoschnau et al., 2008). These findings suggest the genetic penetrance of this SNP changes with age and / or maturation as the influence of the T allele varies between children, adult and elderly individuals (Blades et al., 2010; Korvala et al., 2010; Mann & Ralston, 2003; Suuriniemi et al., 2006). The influence of the T allele may vary with aging as a consequence of altered recovery (Baumert et al., 2016) and T allele carriers reported greater muscle soreness and impaired strength recovery following exercise induced muscle damage in comparison to G allele carriers (Baumert et al., 2018). Muscular strength is linked with reduced fracture risk in the elderly (Alajlouni et al., 2020) and adolescents with apophysitis injuries have been shown to have reduced strength compared with non-injured controls (Rathleff et al., 2020). Furthermore, apophysitis injuries typically occur via non-contact mechanisms, following repeated exposure to exercise loads, on transiently susceptible bone, with insufficient recovery (Faulkner et al., 2006; Wang et al., 2010). Therefore, it is plausible that the risk of apophysitis injury would increase for *COL1A1* (rs1800012) T

allele carriers as a result of impaired recovery (Baumert et al., 2018) and inferior bone strength properties (Suuriniemi et al., 2006) during transient periods of additional bone weakness associated with pubertal growth (Faulkner et al., 2006; Wang et al., 2010) in addition to the exposure of an elite youth football training programme.

Blades et al. (2010) examined the association between the *COL1A1* (rs1800012) and *COL1A2* (rs412777) SNPs with fracture risk in physically active children aged 4-16 years but observed directly contradictory results to those of the present study with *COL1A2* (rs412777) CC individuals, demonstrating a significantly reduced risk of fracture compared with A allele carriers (OR = 0.45 [95% CI = 0.24-0.82], p = 0.01). Chapter three shows that Blades et al. (2010) was the only study to achieve the highest quality classification. However, the only other paper to investigate the genetic association between *COL1A2* (rs412777) and fracture risk also found an increased risk of fracture in prepubertal girls who possessed a C allele compared to those who did not (OR = 4.0 [95% CI = 1.4-11.8], p < 0.05) (Suuriniemi et al., 2003), which is similar to the findings of the present study. Blades et al. (2010) discussed the potential that these contrasting results could be attributable to differences in the genetic background of included participants, or from misidentification of the SNP due to analytical limitations of the study by Suuriniemi et al. (2003) to genotype the participants. However, like Blades et al. (2010), the present study specifically identified the *COL1A2* (rs412777) SNP and includes a broadly similar participant ethnicity group, yet still found opposing results. The authors also discuss how the physiological mechanism by which the *COL1A2* (rs412777) SNP influences fracture risk remains unclear (Blades et al., 2010) and this synonymous A to C SNP may be in strong linkage disequilibrium with a different, currently unknown, causal variant. If this is true, then ethnicity dependent variations could still explain the divergent results but, nevertheless, those possessing the *COL1A2* (rs412777) C allele, and specifically the CC genotype in the present study, appear to demonstrate a near two-fold greater risk of fracture injury than A allele carriers.

The meta-analysis in chapter three examining candidate gene association studies with fracture risk in physically active participants observed substantial heterogeneity between studies and no significant associations could be established from the overall analyses with different genetic comparison models of the *COL1A1* (rs1800012), *COL1A2* (rs412777), *ESR1* (rs2234693) and *VDR* (rs2228570) SNPs. The meta-analyses identified one paper which observed a significantly increased risk of fracture for male military T allele carriers of the *GDF5* (rs143383) SNP (Zhao et al., 2016). However, this association had not been replicated and was not able to be included in the quantitative analysis. The *GDF5* (rs143383) SNP is a non-coding variant in the 5' untranslated region of *GDF5* resulting in a significant reduction in mRNA transcript production linked with the T allele (Southam et al., 2007). The GDF5 protein appears to play an important role in appendicular skeletal development (Buxton et al., 2001; Storm & Kingsley, 1999) and the T allele has been repeatedly associated with osteoarthritis in the knee (Liu et al., 2013; Valdes et al., 2011). The TT genotype of the *GDF5* (rs143383) SNP has also been associated with increased risk of all injuries in football (McCabe & Collins, 2018), although this was not replicated for injury rate, severity or recovery time (Pruna et al., 2016), hamstring muscle injuries (Larruskain et al., 2018) by other, nor observed in the present study. The T allele has also been linked with an increased risk of ACL rupture (Chen et al., 2015), meniscal injury (Ge et al., 2014) and tendinopathy (Posthumus et al., 2010a) but these findings were also not replicated in the present study. Nevertheless, if the results of the present study are included in the

quantitative analysis of the meta-analysis of fracture risk in chapter three and combined with those of Zhao et al. (2016), the C allele of the *GDF5* (rs143383) SNP would demonstrate a significant reduction in fracture risk across all three models ($p < 0.001$) with very little heterogeneity between studies. Completing this analysis indicates that fracture risk is significantly lower for C allele homozygotes than CT and TT individuals (OR = 0.33 [95% CI = 0.18-0.61], $p = 0.0004$, $I^2 = 0\%$). Therefore, although this genetic association observed in the present study was only significant at $p < 0.03$, it is likely that the *GDF5* (rs143383) T allele is associated with an increased risk of fracture injury in elite male footballers, consistent with previous findings in male military recruits (Zhao et al., 2016).

The *VDR* (rs2228570) SNP was considered to be a likely candidate to influence fracture risk in elite male footballers (Table 18) but did not reach statistical significance in the meta-analysis of described in chapter three using a random effects model ($p = 0.06$). The random effects meta-analysis model was required due to significant heterogeneity observed between studies which rebalanced the contribution of included studies. However, using a fixed effects model would have observed a significant increased risk of fracture association with the *VDR* (rs2228570) T allele. Indeed, the T allele of this SNP has been repeatedly associated with reduced BMD (Nakamura et al., 2002b; Strandberg et al., 2003) and increased fracture risk (Chatzipapas et al., 2009; Varley et al., 2018). Nevertheless, these results are inconsistent, and some have observed improved BMD in adolescent footballers (Diogenes et al., 2010) and no association with fracture risk (Korvala et al., 2010), as reported in the present study. However, the *VDR* (rs2228570) SNP was significantly associated with the incidence of non-contact injuries in the present study, with heterozygotes appearing to be at greater risk than homozygotes of either allele ($p = 0.01$). The *VDR* (rs2228570) GA genotype was also associated with an increased risk of all time-loss injuries ($p < 0.04$). However, this is thought to result from the increase in non-contact injuries specifically. A similar study in elite male footballers found no significant difference in muscle injury incidence between the three *VDR* (rs2228570) genotypes but only the codominant genetic model was assessed (Massidda et al., 2015b). Nevertheless, the injury incidence was also highest in GA heterozygotes (0.37 ± 0.62 injuries per 1000 hours) of the *VDR* (rs2228570) SNP when compared to homozygotes (GG = 0.30 ± 0.60 injuries per 1000 hours and AA = 0.20 ± 0.60 injuries per 1000 hours) although this was not statistically significant (Massidda et al., 2015b). Vitamin D interacts with VDRs to activate transcriptional regulation of target cells mediating the effects of vitamin D in the body (Wang et al., 2005), including muscle repair, muscle function, immunity, cardiac function and bone homeostasis (Dahlquist et al., 2015; Owens et al., 2018). The transcriptional activation of vitamin D appears to be 1.7 times greater for the C than T allele of the *VDR* (rs2228570) SNP (Arai et al., 1997) and CC individuals have shown better strength performance (Windelinckx et al., 2007). Non-contact injuries are common in football (Le Gall et al., 2006; Price et al., 2004; Read et al., 2018) and effective recovery is important to deal with the demands of congested weekly competition schedules (Arruda et al., 2015; Carling et al., 2016). Therefore, although it is unclear why incidence would be greatest for CT individuals, the *VDR* (rs2228570) SNP may influence non-contact injury in elite male footballers because of an effect on physical performance. Alternatively, these differences may be due to changes in VDR activity, which has multiple physical performance and exercise related functions within the body, and their influence on recovery (Dahlquist et al., 2015; Owens et al., 2018). It could be that homozygotes of each allele possess some advantage or sufficient compensatory

mechanisms that reduce non-contact injury incidence, which might not be as pronounced in heterozygotes.

The present study also found a significant association between *COL5A1* (rs12722) and non-contact muscle injury ($p \leq 0.01$). As with *VDR* (rs2228570) the incidence of injury was significant for the *COL5A1* (rs12722) SNP using an overdominant genetic comparison model, although, the risk of non-contact muscle injury was significantly lower for *COL5A1* (rs12722) TC heterozygotes. The *COL5A1* (rs12722) SNP is located in the 3' untranslated region of *COL5A1*, increasing stability of the mRNA transcript and suggesting greater production of COL5A1 with the T allele (Laguette et al., 2011). Increased COL5A1, in turn, appears to reduce the diameter of type I collagen fibres (Birk et al., 1990). Whilst the *COL5A1* (rs12722) C allele has been repeatedly associated with a reduced risk of tendon (Altinisik et al., 2015; Mokone et al., 2006) and ligament injury (O'Connell et al., 2015; Posthumus et al., 2009b), the T allele has also been linked with improved endurance performance (Brown et al., 2011a; Posthumus et al., 2011b). In the present study, C allele carriers appeared to be less than half as likely to suffer a tendon injury than T homozygotes (RR = 0.41 [99% CI = 0.15-1.10], $p = 0.02$) but no association was observed with ligament injury. The CC genotype has been associated with a reduced risk of exercise associated muscle cramping in endurance athletes (O'Connell et al., 2013) but no significant association has previously been observed between the COL5A1 (rs12722) SNP and muscle injuries in football (Larruskain et al., 2018) or mixed athlete groups (Miyamoto-Mikami et al., 2019). The TC genotype has also been associated with an increase in muscle injury severity in footballers (Pruna et al., 2016). However, the HWE was not examined in this study and no TT homozygotes were observed in the predominantly Caucasian sample, which is certainly unexpected, as indicated in the European supergroup of Table 24. Therefore, the previous association between the *COL5A1* (rs12722) TC genotype and muscle injury severity in footballers (Pruna et al., 2016) is questionable. Nevertheless, the association of the *COL5A1* (rs12722) TC genotype with reduced incidence of non-contact muscle injury in the present study is a novel finding, and further investigation is required to confirm this association.

Muscle-tendon stiffness and its associated benefits to running economy were originally hypothesised as the mechanism by which the *COL5A1* (rs12722) T may enhance endurance performance following previous associations with joint range of motion (Brown et al., 2011b; Collins & Posthumus, 2011). However, subsequent research does not support this theory, with no significant differences observed in the volume or elasticity of tendons (Foster et al., 2014) nor running economy (Bertuzzi et al., 2014) between genotypes. Nevertheless, type V collagen plays a crucial role in the regulation of fibrilogenesis in non-cartilaginous tissue (Collins & Posthumus, 2011) and it may be that the *COL5A1* (rs12722) SNP influences endurance performance via an alternative mechanism. Mice models indicate that reduced *col5a1* mRNA production - associated with the C allele of the human *COL5A1* (rs12722) SNP - decreases the compliance and tensile strength of the aorta (Wenstrup et al., 2006). Higher compliance of the aorta is associated with endurance training and greater stroke volume (Tarumi et al., 2021), which may provide an alternative mechanism for enhanced endurance performance, despite the reduced diameter of type I fibres, resulting from increased type V collagen abundance (Birk et al., 1990), associated with the T allele. Therefore, although associations between the *COL5A1* (rs12722) T allele and endurance performance remain equivocal, TC individuals may have an advantage over homozygotes, resulting from a balance

between improved endurance performance and tissue loading capacity, which are both considered protective against non-contact muscle injury (Gabbett, 2016; Kalkhoven, 2021). Alternatively, this protective effect of the TC genotype may result from a balanced trade-off between the increased non-contact muscle injury susceptibility observed in the present study for CC homozygotes and increased tendon injury incidence observed for TT homozygotes. Mechanical stress at the musculotendinous junction in *COL5A1* (rs12722) T and C allele homozygotes may subsequently be more susceptible to injury than TC heterozygotes, which could be expressed by increased observation of non-contact muscle injury.

In addition to *COL5A1* (rs12722) and *VDR* (rs2228570), the *MMP3* (rs679620) SNP was also associated with non-contact injury incidence in the present study, with CC individuals suffering significantly fewer non-contact injuries than CT & TT individuals. This finding aligns with previous research investigating the *MMP3* (rs679620) SNP in football (Larruskain et al., 2018), although, no significant associations were found with tendon or ligament injury specifically in the present study, as has been repeatedly reported by others (Briški et al., 2021; El Khoury et al., 2016; Nie et al., 2019). The MMP3 molecule stimulates activation of other metalloproteinases (Nagase et al., 2006; Toth et al., 2003), which regulate the extra-cellular matrix by catalytically degrading structural proteins (Birkedal-Hansen et al., 1993; Somerville et al., 2003). MMPs are critical to the initial healing processes following tissue damage (Somerville et al., 2003) but elevated MMP levels can disrupt long-term healing and appear in chronic injuries (Bullen et al., 1995). Therefore, the catalytic activity of MMPs are tightly controlled, amongst other mechanisms, by gene transcription (Löffek et al., 2011). The *MMP3* (rs679620) SNP is a missense coding variant and although the physiological consequence of this is unknown, it could interfere with the physiological function of the resultant MMP3 protein (Raleigh et al., 2009). Therefore, the C allele may be protective against non-contact injury due to a reduction in the sustained catalytic action of MMPs in breaking down the extra-cellular matrix following damage which is important for mechanical support and force transmission (Somerville et al., 2003).

Non-contact injuries are typically considered to be more preventable than contact injuries and more directly attributable to the loading exposure of the athlete at that moment (Gabbett, 2016; Kalkhoven et al., 2021). Therefore, understanding the interaction that loading exposure may have on the risk of non-contact injuries in football between genotypes of the *COL5A1* (rs12722), *VDR* (rs2228570) and *MMP3* (rs679620) SNPs may prove informative to guide the practical application and understanding of genetic predisposition to injury. The VDR, COL5A1 & MMP3 proteins all appear to be involved in the regulatory or adaptive functions of the musculoskeletal system (Birk et al., 1990; Nagase et al., 2006; Owens et al., 2018). The general adaptation syndrome provides a conceptual model to understand the acute response and fatigue following exercise which, with adequate recovery, can stimulate an adaptive response via genetic mechanisms, resulting in a supercompensation effect and improved performance (Cunanan et al., 2018; Wackerhage & Woods, 2002). However, if insufficient recovery is achieved before the next fatiguing exercise, then performance can further reduce and, if repeated, lead to reduced tissue-specific loading capacity and increased injury susceptibility (Cunanan et al., 2018; Meeusen et al., 2006). Therefore, it is plausible that genetic variants within the *COL5A1*, *VDR* and *MMP3* genes may influence the incidence of non-contact injuries due to the potentially far-reaching consequences on tissue-specific

loading capacity, recovery, and adaptation. Furthermore, the interaction between SNPs was not explored in the present study but could provide further insight and the *COL5A1* (rs12722) T allele has been shown to interact with the C allele of the *MMP3* (rs679620) SNP to cause an even greater risk of tendinopathy than each SNP in isolation (Raleigh et al., 2009). Recovery is particularly important in football and genetic variants which affect an individual's ability to withstand a high training and competitive load could be supported to guide individualised applied decisions which could support players to remain injury free through more targeted / bespoke training exposure.

The *ACTN3* (rs1815739), *ACE* (rs1799752) and *ESR1* (rs2234693) SNPs each showed no significant influence on the incidence or risk of injury in elite male footballer despite previous genetic associations (Kumagai et al., 2018; Massidda et al., 2019, 2020). However, despite the *ACTN3* (rs1815739) genotype frequencies appearing in HWE, there was a greater number of RR individuals than might be expected in the present study. This appeared most evident in the European population group and a similar subtle skew is observed in the Caucasian football subgroup of (Clos et al., 2019) and is common in football (McAuley et al., 2020). Indeed, all the XX individuals in the present study were of European ancestry and it is possible that other unknown ethnicity-dependent genetic factors may have contributed to this unexpected observation. The XX genotype of the *ACTN3* (rs1815739) SNP has been repeatedly associated with increased risk of non-contact muscle injuries (Lim et al., 2021; Zouhal et al., 2021) and R allele homozygotes have consistently been shown to suffer a significantly lower degree of muscle damage than X allele carriers following exercise (Belli et al., 2017; Del Coso et al., 2016). Similar observations have been made between the *ACE* (rs1799752) genotypes and II homozygotes appear to suffer greater muscle damage following exercise (Sierra et al., 2019) and are at greater risk of muscle injury in football (Massidda et al., 2020). However, this was not observed in the present study and is likely affected by the absence of any II individuals in the participant group who may be expected to demonstrate the clearest influence should it exist. The D allele appears to be overrepresented in youth football (age 15-21 years) (McAuley et al., 2020). The authors take care to make clear that such observations suggest, at best, a very minor advantage for *ACE* (rs1799752) D allele carriers in youth football, which does not appear to translate to significant differences between professionals and controls in their meta-analysis (McAuley et al., 2020). Nevertheless, this may explain the underrepresentation of II individuals in the present study, which includes a substantial number of youth players.

The complex and multifactorial nature of football performance success (Reilly et al., 2000), in addition to, the existence of alleles which may confer both a physical performance advantage and an increased injury risk - e.g., *COL5A1* (rs12722) - mean the results of the present study should not be considered deterministic for future success or prohibitive to football participation for any individual at this time. Our current understanding of the role that genetics plays on physical performance and injury susceptibility remains extremely limited and rudimentary. Therefore, although some genetic associations appear to have been observed in the present study, these should be considered, at best, predispositional and not deterministic for injury occurrence. The ethical considerations for the study were particularly important as concern has been expressed around the use of genetic testing to examine the risk of exercise-related injury in children and young people (Vlahovich et al., 2016; Williams et al., 2016). The purpose of the present study was to improve the current understanding of genetic factors contributing to an individual's inherent injury susceptibility to help protect all players

from harm and facilitate the long-term development and wellbeing of every individual. Fracture and apophysitis injuries are particularly prevalent in youth footballers which frequently result in severe time loss injury (Le Gall et al., 2006; Light et al., 2021; Price et al., 2004; Read et al., 2018). Therefore, inclusion of youth players was considered appropriate to investigate the genetic association with injury in the present study. However, the results of this study cannot be used with any validity to discriminate or exclude individuals from participation in sport as our understanding of the complex interaction between genetic and environmental factors is completely insufficient to assert such long-term performance implications.

The findings should be acknowledged within the present study population and considered along with its limitations. Recruitment of an adequate numbers of cases within a highly specific disease / injury definition is often difficult and less specific definitions are frequently introduced to make up sufficient numbers in an attempt to attain a certain level of power (Zondervan & Cardon, 2007). A balance between practical applicability and clinical specificity was sought in the present study by targeting tissue-specific pathologies. However, these would have included tissue injuries from various locations around the body (e.g., ligament sprains in the knee and ankle were combined). Therefore, whilst a gain in power may have been achieved statistically, a loss of power may have occurred in reality if this resulted in a substantial increase in causal heterogeneity (Zondervan & Cardon, 2007). Furthermore, although the longitudinal nature of the study is arguably a strength, data were retrieved retrospectively for individuals across a range of seasons and time periods, which only included injuries observed during registration at the football club and injuries occurring prior to or following the injury observation period were not examined and could plausibly affect the genetic association with injury findings.

The results of this paper should not be considered as a replacement for direct measurement and monitoring of physical performance qualities, which remain the most effective evaluation of the tissue-specific load capacity of an individual at that moment (Wackerhage & Schoenfeld, 2021). Nevertheless, considering the inherent genetic predisposition of individual tissue-specific load tolerance capacity/ injury susceptibility may support targeted training interventions to reduce players risk of injury. This may be even more important for youth athletes for whom the future physical performance and injury susceptibility may be mitigated with training and or nutritional interventions (Anderson et al., 2017; Larruskain et al., 2021; Lemes et al., 2021). The results of the present study suggest that the *GDF5* (rs143383) and *COL1A2* (rs412777) SNPs may be candidates for consideration to inform targeted intervention strategies to mitigate inherent fracture injury risk in elite male footballers. The *COL5A1* (rs12722), *VDR* (rs2228570) and *MMP3* (rs679620) SNPs appear to influence the risk of non-contact injury and understanding the dynamic interaction that football loading exposure has in combination with the different genotypes, and combined interactions of, these SNPs may be informative to support the individualisation of training programmes to support long term development and performance. Furthermore, although worthy of scepticism, genetic associations observed between the *GDF5* (rs143383), *COL5A1* (rs12722) and *COL1A1* (rs1800012) SNPs with fracture, tendon and apophysitis injury risk should not be completely disregarded and future research on the influence of these genetic variants with injury risk may be warranted.

5.5 Conclusion

The *COL1A2* (rs412777), *MMP3* (rs679620), *VDR* (rs2228570), *COL5A1* (rs12722), *GDF5* (rs143383) and *COL1A1* (rs1800012) SNPs all appear to influence the interindividual risk of injury in elite male football players. The *COL1A2* (rs412777) and *GDF5* (rs143383) SNPs appear to play a significant role in the risk of fracture injury and may prove informative to guide individualised interventions designed to mitigate this risk and support the long-term player development. The *MMP3* (rs679620) and *VDR* (rs2228570) SNPs appear to influence the incidence of all non-contact injuries, with the *COL5A1* (rs12722) and *COL1A1* (rs1800012) SNPs specifically related to non-contact muscle and apophysitis injuries, respectively. Non-contact injuries are more preventable and attributable to variations in training load and / or maturation than contact injuries (Gabbett, 2016; Kalkhoven et al., 2021; Kemper et al., 2015; Read et al., 2018). Therefore, the relationship between loading exposure and / or maturational status with the *MMP3* (rs679620), *VDR* (rs2228570), *COL5A1* (rs12722) and *COL1A1* (rs1800012) SNPs may provide more granular information, which could support the practical application of genetically individualised training programmes to reduce the risk of injury and future studies to explore these relationships could be beneficial. A greater understanding of how genetic variants influence tissue-specific injury susceptibility could guide tissue-specific injury prevention strategies to support the long-term development of elite male football players, which vary depending on age, growth and maturation.

CHAPTER 6: Total Genotype Score, Growth or Maturation and Loading Exposure as Risk Factors for Injury in Elite Male Youth Football

This chapter examines the potential interaction of growth or maturation and training load variables with candidate genes and TGS, based on the findings of previous associations, with joint, tendon, non-contact, non-contact muscle, apophysitis and fracture injuries throughout the elite male football development pathway. Possession of the *COL1A1* (rs1800012) T allele and growth rate >0.6 $cm.m^{-1}$ were associated with significantly greater risk of apophysitis injury than GG individuals and growth rate <0.6 $cm.m^{-1}$, respectively. Furthermore, an interaction effect was observed, whereby TT+TG individuals with growth rate >0.6 $cm.m^{-1}$ were associated with the greatest incidence of apophysitis injury. Others have observed that high calcium intake can stimulate greater bone mass accrual pubertal in *COL1A1* (rs1800012) T allele carriers. Therefore, the *COL1A1* (rs1800012) genotype could inform genetically individualised calcium supplementation interventions during periods of rapid growth in elite male youth footballers to reduce apophysitis injury risk. However, none of the other models were significantly associated with injury incidence and the findings with apophysitis injury will need to be replicated in an independent sample of elite male youth footballers to validate the findings of the present study. Therefore, further research is required to identify genetic variants, which influence tissue specific injury risk in elite male football, before genetically individualised injury prevention strategies may be implemented in applied practise with confidence at this time despite some promising observations.

6.1 Introduction

Time loss injury is detrimental to senior football team performance success (Hägglund et al., 2013) and can impede the long-term development of youth footballers (Jones et al., 2019; Larruskain et al., 2021). Musculoskeletal tissue injury occurs when tissue-specific loading exposure exceeds the threshold tolerance of that individual at that time (Kalkhoven et al., 2020, 2021). Training load monitoring allows inferences of tissue-specific loading to be estimated in an attempt to identify individuals at risk of injury and appropriate progressive overload (Bowen et al., 2017, 2020; Buckthorpe et al., 2019). This information is then combined with other athlete monitoring, medical screening, and injury history data to guide future training decisions, with the aim of reducing injury and supporting successful performance (Brink et al., 2010; Buckthorpe et al., 2019; Gabbett, 2018; Halson, 2014). These practical interventions are informed by the prognostic understanding interpreted from the data, which may include reducing players' training and competition exposure and, therefore, also the risk of experiencing injurious tissue-specific loading, which the player is not expected to tolerate. Muscle, bone, ligament, and tendon are tissues prone to injury in football, which predominantly occur in the lower limbs (Le Gall et al., 2006; Price et al., 2004; Read et al., 2018b). Bone fracture and apophysitis are more common in children and adolescent footballers, which typically occur prior to, and during, the peak adolescent growth spurt. On the other hand, muscle, ligament, and tendon injuries tend to increase into adulthood (Johnson et al., 2020; Le Gall et al., 2006, 2007; Light et al., 2021; Materne et al., 2020; Price et al., 2004; Read et al., 2018; Rumpf & Cronin, 2012; Wik et al., 2020a). Injury occurrence is complex and multifaceted (Bittencourt et al.,

2016; Tee et al., 2020) and, although prognostic assessments show promise in identifying individuals at increased risk of injury in football, more information is needed to improve the accuracy of individual risk estimations (Hughes et al., 2020).

Genetic differences between individuals have repeatedly been shown to influence tissue-specific injury incidence risk (Blades et al., 2010; Diogenes et al., 2010; Kim et al., 2017a; Larruskain et al., 2018; Massidda et al., 2015a, 2020; Posthumus et al., 2010b; Varley et al., 2018), and heritable differences are estimated to account for 40-69% of the variability in ACL rupture risk, fracture, and tendinopathy injury risk (Andrew et al., 2004; Hakim et al., 2003; Magnusson et al., 2020). Understanding which genetic variants contribute to the heritable variability in injury risk of elite male football players could allow individuals to be stratified based on specific-tissue injury susceptibility. Training and nutritional strategies which aim to reduce the risk of individual injury susceptibility also vary depending on the targeted musculoskeletal tissue (Roessler et al., 2014). Therefore, targeting individual injury susceptibility may be more effective than traditional general injury prevention programmes (Fuller, 2019). Consequently, understanding how genetic variations influence individual susceptibility to tissue-specific injury could provide more bespoke designed injury prevention interventions (Fanchini et al., 2020; Helgerud et al., 2001; Soomro et al., 2016). In this way, genetically individualised programmes have the potential to improve the specificity, efficiency, and effectiveness of training interventions (Pickering & Kiely, 2019).

A comprehensive understanding of the heritable factors that influence individual susceptibility to tissue-specific injury remains unclear. As a complex trait, tissue-specific injury susceptibility is thought to be influenced by the contribution of a large number of heritable factors, which may individually have only a small effect (Gibson, 2009). Furthermore, accounting or controlling for other contributory genetic and environmental factors is challenging and / or expensive so large sample sizes are often needed to establish clear associations in one of the potentially thousands of heritable factors influencing injury risk. Nevertheless, the interaction and cumulative contribution of numerous genetic variants, of both large and small effect, is considered likely to predispose, rather than determine, injury occurrence (Gibson, 2016; Zondervan & Cardon, 2007). In acknowledgment of the substantial interindividual variability in the response to training and exercise (Baumert et al., 2016; Bouchard & Rankinen, 2001; Hubal et al., 2005), an increased emphasis has been placed on the need to individualise monitoring, development, and injury prevention strategies (Halson, 2014; Pickering & Kiely, 2019). Therefore, some have sought to develop TGSs, which aim to estimate the relative inherent injury risk of an individual based on the combination several genetic variants previously associated with injury risk (Goodlin et al., 2015; Posthumus et al., 2011a). These TGSs have shown some promise in differentiating the response to aerobic training (Pickering et al., 2018) and all injury incidence (Hall et al., 2022) of male youth footballers, in addition to, endurance / power related phenotypes in athletes of various sports (Ahmetov et al., 2009; Ben-Zaken et al., 2015; Grealy et al., 2015; Ruiz et al., 2009, 2010; Santiago et al., 2010). However, due to the current limited understanding, TGSs for injury often use data collected from non-elite, inactive, elderly, or diseased participants to inform the risk profile development (Goodlin et al., 2015). While the use of clinical and epidemiological data should be considered when examining the plausibility of a genetic variant for inclusion in TGS, the influence of genetic variants appear to change with sex and physiological stages of growth, maturation, and aging as shown in chapter three and by others (Cooper et al.,

2013; Hall et al., 2022). Hall et al. (2022) investigated if nine SNPs - including *ACTN3* (rs1815739), *COL1A1* (rs1800012), *COL5A1* (rs12722) and *MMP3* (rs679620) - influenced injury risk in 402 Caucasian youth footballers at different stages of maturation in one season. Single SNP maturation and tissue-specific genetic associations were observed but no tissue-specific TGS was conducted, despite injured players showing greater TGS than non-injured players (Hall et al., 2022). Therefore, tissue-specific injury risk TGS estimations in elite male youth football could be improved by using genetic variants associated with injury in young, healthy, physically active male participants and provide more applicable information on individual susceptibility.

Considering the dynamic aetiology of musculoskeletal injury presented by Kalkhoven et al. (2020, 2021), an individual's tissue-specific threshold tolerance may be mediated by genetic factors to predispose injury occurrence. However, loading exposure at the moment of injury is the injurious force that results in the tissue damage. Consequently, coaches aim to manage training and competition exposure to progressively overload at a safe rate to support performance adaptation and reduce injury risk (Bowen et al., 2020; Cunanan et al., 2018; Dalen-Lorentsen et al., 2021; Fanchini et al., 2020). Including both TGS and loading exposure experienced by players into a risk model could improve coaches' ability to identify when individuals are at significantly greater risk of injury. Indeed, Grealy et al. (2015) highlight how important training and environmental factors are in understanding how TGSs affect physical phenotypes and suggest that more sophisticated genetic models accounting for these variables could improve the accuracy of TGS predictions. Aging appears to increase the risk of muscle, ligament, and tendon injuries from childhood into adulthood, but apophysitis injuries spike around PHV during which time a temporal weakness in rapidly growing long bones occurs (Light et al., 2021; Materne et al., 2020; Wang et al., 2010; Wik et al., 2020a). Consequently, tissue-specific loading capacity appears to be influenced by growth and maturation (Kemper et al., 2015; Light et al., 2021; Wang et al., 2010; Wik et al., 2020a), which may interact with genetic predisposition and loading exposure to substantially differentiate the risk of injuries in elite male youth footballers. Therefore, the aim of the present study was to examine if genetic predisposition, estimated using tissue-specific TGSs, loading exposure, and maturation or growth status interact to influence the individual risk of injury in elite male youth football development.

6.2 Methods

6.2.1 Study design

The present study was designed in accordance with published guidelines for genetic association studies (Little et al., 2009; Romero et al., 2002) and consensus statements for injury research in sport (Bahr et al., 2020) and football (Fuller, 2006). A retrospective, observational, case-control genetic association experimental design was adopted to investigate the aim of the study at an elite male football club (Fulham Football Club). The study received institutional ethical approval and was registered at ClinicalTrials.gov (NCT05220969).

6.2.2 Participants

Convenience sampling identified 157 eligible candidates, who were invited to participate in the study as elite male football players, aged between 10 and 35 years at the time of recruitment. Of these, 113 players, and parents of those under 18 years of age, provided written informed consent to participate in the study. Fourteen goalkeepers were excluded from the analysis because of the substantial differences in the nature of their activity compared with outfielders. One participant was excluded as >50% of his genotype profile could not be clearly determined using the saliva sample provided and could not be resampled. Three players were excluded because they did not have at least one season of loading and injury surveillance data. Therefore, a total of 95 participants were included in the study. Participant self-identified ethnicity was categorised into the 1000 Genomes Project five continental super population groups: African, ad mixed American, East Asian, European and South Asian (The 1000 Genomes Project Consortium, 2015). Players were considered elite, as they were registered players at an English Premier League Category One Football Academy or professional players competing in the top two tiers of English football (English Football League Championship or English Premier League Divisions). This study only includes data collected from participants who were a registered player within the under-12 to under-23 academy and senior male First teams at the football club between November 2015 and June 2021. Fracture and apophysitis injuries are particularly prevalent in youth footballers and can result in severe time-loss injury during important physiological and football pathway, particularly in development periods when growth and maturation may influence the risk of injury (Larruskain et al., 2021; Light et al., 2021; Read et al., 2018b). For this reason, the inclusion of youth players was considered important to allow specific investigation of apophysitis injuries.

6.2.3 Injury surveillance

Injury data were retrospectively collected from injury data archives. These data were prospectively recorded as part of normal working practice at an elite football club between November 2015 and June 2021. Initial injury diagnoses were performed by club medical staff, who were blinded to the genotype data, with their assessments verified by second opinion and / or scans, when these were considered practically appropriate in the presence of reasonable doubt. Club's medical staff recorded all complaints requiring attention, in accordance with consensus guidance on injury definition and data collection procedures in football (Fuller, 2006). Time loss injury was defined as tissue damage or disruption to normal physical function, occurring from football or related training activities (e.g., gym-based strength or field-based conditioning sessions), resulting in at least one day of missed training or competition (Bahr et al., 2020; Fuller, 2006). Data for each time-loss injury included incidence date, occurrence type (match or training), onset (acute, overuse, or mixed), mechanism (contact or non-contact), full-training return date, diagnosis and location - using the Orchard Sports Injury and Illness Classification System (OSIICS) version 10 (Rae & Orchard, 2007). The type of occurrence, OSIICS code, onset and mechanism of each injury were verified by cross-reference with injury rehabilitation and / or management notes and, where possible, with video confirmation from match analysis archives. Incidence of injury was calculated as the number of injuries per 1000 hours of recorded training and matches during the injury surveillance and load

monitoring period for included participants (Bahr et al., 2020). Illnesses were excluded from this analysis.

6.2.4 Injury phenotype classification

Time loss injuries were classified into six groups for statistical analysis: non-contact injuries, non-contact muscle injuries, fracture injuries, apophysitis injuries, joint injuries and tendon injuries. These injury classifications were adopted as they provide clear and accurately definable phenotypes with clinical relevance to guide applied decision making, as recommended by Zondervan and Cardon (2007). Non-contact injuries include all time-loss injuries occurring via non-contact mechanisms, and non-contact muscle injuries those specific to muscle tissue. Non-contact mechanisms of injury were considered to have occurred without any direct or indirect contact from another player and included injuries resulting from the injured player striking a ball. Fracture injuries included both contact and non-contact mechanisms and were always confirmed by X-ray or magnetic resonance imaging scans. Joint injuries encompassed both contact and non-contact time loss injuries to the ligaments and cartilaginous tissue of joints. Tendon injuries were time-loss injuries determined to be predominantly affecting tendinous tissue. However, tendinopathy injuries frequently display complex aetiology, which can result from referral of other symptoms, such as adductor tendinopathy from pubic bone oedema, for example. Apophysitis injuries were gradual onset and non-contact time-loss injuries, including Severs disease, Osgood-Schlatter's disease and hip apophysitis. The apophysitis injury analysis included a subset of 71 participants who had injury surveillance, loading exposure, and growth and maturation data between the ages of 10 and 16 years of age, after which apophysitis injury risk dramatically declines, as explored in Chapter four (de Loës, 1995; Hall et al., 2020; Le Gall et al., 2006; Price et al., 2004; Read et al., 2018b; Rumpf & Cronin, 2012).

6.2.5 Genetic testing

Participants were asked to provide a 2 mL saliva sample into a collection vial (SalivaGene Collection Module II; Stratec Molecular GmbH), adhering to manufacturers guidelines and under the supervision of the main investigator (ERM). This method was selected as a non-invasive method allowing collection of ample genetic material for genotyping and a more appropriate one for use with children (Romero et al., 2002). A stabiliser solution provided by the manufacturer (SalivaGene Collection Module II; Stratec Molecular GmbH) was then mixed with the saliva sample, and the container was subsequently sealed and labelled with an anonymous identification code, known only to the main investigator of the study. Sealed and labelled samples were then each placed into an individual grip seal plastic envelope, which was also labelled with the participant identification code. This was then transported to and stored within a Human Tissue Authority certified laboratory, following certification guidelines at room temperatures. Stored saliva samples were transferred into 1.6 mL screw-top tubes, which were labelled and shipped to LGC genomics (LGC Limited, United Kingdom), who were blinded to the injury surveillance data. LGC Genomics (LGC Limited, United Kingdom) extracted DNA from participant samples, designed KASP™ assays for, and completed the genotyping services. Primer sequences were validated by LGC Genomics prior to the analysis of

experimental DNA samples for the seven genetic variants of interest, which are all in linkage equilibrium with each other. KASP is a homogeneous, fluorescence-based genotyping technology based on allele-specific oligo extension and fluorescence resonance energy transfer for signal generation (Semagn et al., 2014). Variant specific primer and master mix were added to 10 ng of DNA according to the manufacturer's instructions. PCR amplification was performed (Veriti 384 thermal cycler, Applied Biosystems) and fluorescent signals read out using the 7900HT Fast Real-Time PCR System (Applied Biosystems) and converted to genotype information with the SDS program (version 2.3 Applied Biosystems). All assays tested gave greater than 94.6% call rate on the samples tested and a negative control was included on each plate to check for non-specific amplification. Genotypes were not automatically assigned to samples which failed to amplify consistently with the rest of the cluster. Genotype classifications were independently confirmed by visual inspection of cluster plots and agreed by consensus opinion by two of the main investigators (YM and ERM) using SNPViewer2 (KBiosciences UK Ltd., Hoddesdon, Herts, UK) as recommended (Semagn et al., 2014).

6.2.6 Total genotype score generation

Genetic variants were identified as candidates for inclusion in the TGS for each injury classification based on the presence of a plausible mechanism of effect and existent associations, and those observed in Chapter five, with the tissue of injury in young healthy physically active male populations as shown in Table 25. This process identified seven genetic variants - *ACTN3* (rs1815739), *COL1A1* (rs1800012), *COL1A2* (rs412777), *GDF5* (rs143383), *MMP3* (rs679620), *COL5A1* (rs12722) and *VDR* (rs2228570) – with evidence of an influence on injury analysis classifications. The TGS for each participant and injury classification was calculated in line with previous research using a simple additive model, mathematically transformed to a 0-100 scale for ease of interpretability (Williams & Folland, 2008). Low-risk genotypes were given a risk score of 0, medium risk a score of 1 and high risk a score of 2 although the apophysitis injury analysis was dichotomised as low or how risk based on the presence of the *COL1A1* (rs1800012) T allele only using a dominant genetic association model (TT *vs.* GT + GG). Considering the genotype risk scores GS_1, GS_2, up to GS_X, and the maximum possible TGS (TGS_{MAX}), an individual's TGS was calculated as:

$$TGS = \left(\frac{100}{TGS_{MAX}}\right) \times (\, GS_1 + GS_2 + \cdots + GS_x)$$

Participants were then categorised as possessing a low (TGS ≤ 33), medium (TGS > 33 ≤ 66) or high (TGS > 66) inherent genetic risk of injury for each injury analyses, similar to previous stratifications of TGS to differentiate categories (Baumert et al., 2022; Pickering et al., 2018; Williams & Folland, 2008).

Table 25. Total genotype risk score variant allocation for tissue-specific injury in elite male football.

Injury Model	Included variants	Genotype Risk Score Allocation			Genetic Associations	References
		High = 2	Medium = 1	Low = 0		
Non-contact injury	*ACTN3* (rs1815739)	XX	XR	RR	X allele ↑ muscle damage and non-contact injury	(Clos et al., 2019; Lim et al., 2021; Zouhal et al., 2021)
	MMP3 (rs679620)	TT	CT	CC	T allele ↑ non-contact and tendon injury	(Briški et al., 2021; Larruskain et al., 2018)
	VDR (rs2228570)	AG	AA	GG	G allele ↑ back pain A allele ↑ stress fracture and ↓ bone mineral density	(Cauci et al., 2017; Nakamura et al., 2002; Varley et al., 2018) Chapter five
Non-contact muscle injury	*ACTN3* (rs1815739)	XX	XR	RR	X allele ↑ muscle damage and non-contact muscle injury	(Clos et al., 2019; Lim et al., 2021; Zouhal et al., 2021)
	MMP3 (rs679620)	TT	TC	CC	T allele ↑ non-contact muscle injury	(Briški et al., 2021; Larruskain et al., 2018)
	COL5A1 (rs12722)	TT	CC	TC	C allele ↓ muscle cramping but T allele ↑ muscle injury severity	(O'Connell et al., 2013; Pruna et al., 2016) Chapter five
Joint injury	*ACTN3* (rs1815739)	XX	XR	RR	X allele ↑ muscle damage, ankle joint injury incidence and severity	(Shang et al., 2015; Zouhal et al., 2021) (Hall et al., 2022)
	COL1A1 (rs1800012)	GG	GT	TT	G allele ↓ joint laxity and ↑ anterior cruciate ligament injury	(Ficek et al., 2013)
	COL5A1 (rs12722)	TT	TC	CC	T allele ↑ joint injury	(Lulińska-Kuklik et al., 2018) (Pabalan et al., 2018)
Fracture injury	*COL1A2* (rs412777)	CC	CA	AA	C allele ↑ fracture risk and ↓ bone mineral density	(Blades et al., 2010) Chapter five
	GDF5 (rs143383)	AA	AG	GG	A allele ↑ fracture injury	(Zhao et al., 2016) Chapter five
	VDR (rs2228570)	AA	AG	GG	A allele ↑ fracture risk and ↓ bone mineral density	(Chatzipapas et al., 2009; Nakamura et al., 2002; Varley et al., 2018)
Tendon injury	*MMP3* (rs679620)	CC	CT	TT	C allele ↑ tendon injury	(Briški et al., 2021; Larruskain et al., 2018)
	COL5A1 (rs12722)	TT	TC	CC	T allele ↑ tendon injury	(El Khoury et al., 2016; Nie et al., 2019) Chapter five
Apophysitis injury	*COL1A1* (rs1800012)	T allele carriers (TT+TG)		GG	T allele ↑ fracture and apophysitis injury	(Blades et al., 2010) Chapter five

Note: T allele carriers of the *COL1A1* (rs1800012) variant considered high risk.

6.2.7 Growth and maturation

Stature and body mass were scheduled for measurement every 5th, 6th, or 7th week of the standard football season, dependent on training schedule but including pre-season (July – May), for participants in the Under-12 to Under-16 academy teams. Frequency of measurement for these age

groups was determined by evidence that their growth rates are expected to change rapidly (Abbassi, 1998). Participants in the Under-18, Under-23 and First team groups were measured once a season, as growth rate dramatically slow into adulthood (Abbassi, 1998). Stature was measured using a free-standing portable stadiometer (Seca 213 portable stadiometer; Seca, Birmingham, United Kingdom) to the nearest 0.1 cm by four different football club sports science staff, who followed the International Society for the Advancement of Kinanthropometry (ISAK) recommended procedures (Norton, 2018). The majority of measurements were collected by an ISAK Level one certified Anthropometrist and the primary author of the study (ERM). Participants were asked to stand erect on the base of the stadiometer, without shoes and with their head in the Frankfort horizontal plane (Norton, 2018) for measurement. Measurements were repeated, and a third measure collected, if the difference between stature measures was > 0.5 cm. The two closest measures within this range were accepted and averaged to establish the recorded stature. Body mass was measured using portable digital scales (Seca 875 flat scale; Seca, Birmingham, United Kingdom) to the nearest 0.5 kg, with the final measurements stored in a secured central anthropometric database. Participants who missed measurement within a scheduled window were measured as soon as possible over the two subsequent weeks. The intra-tester reliability of anthropometric measurements showed a technical error of measurement of 0.2 cm for stature equating to a relative technical error of measurement of 0.1%.

Growth rate was calculated as the linear change in stature divided by the days from previous measurement and dichotomised as high growth (>0.6 $cm.m^{-1}$ = >7.2 $cm.y^{-1}$) or not (<0.6 $cm.m^{-1}$ = <7.2 $cm.y^{-1}$). These growth rates were selected to align with previously identified thresholds significantly associated with increased injury risk (Kemper et al., 2015; Wik et al., 2020a) and variation around normal growth through childhood and PHV (Abbassi, 1998). Maturation at each measurement was estimated based on the percentage attainment of predicted adult height (PAH). Self-reported biological parent statures, participant measured stature, and body mass were combined with age specific coefficients to estimate PAH (Khamis & Roche, 1994). Participants were subsequently classified as Pre-, Circa- or Post-PHV when PAH attainment was <89%, 89-95% or >95% respectively, similar to previously established thresholds aligned with PHV (Beunen et al., 1997; Johnson et al., 2020; Malina et al., 2007a; Parr et al., 2020). Percentage attainment of PAH has been used as a non-invasive alternative estimate of maturation / pubertal status (Malina et al., 2015) and was previously validated in youth American Football players (Malina et al., 2007a). Participants aged 18 years and older were automatically classified as Post-PHV because the Khamis-Roche method does not extend beyond 18 years of age, by which time 99% of boys are expected to be at least 96% of PAH (Beunen et al., 1997; Khamis & Roche, 1994; Tanner et al., 1966).

6.2.8 Loading exposure

The football training and match loading data were retrospectively collected from load monitoring databases, which has been prospectively recorded between November 2015 and June 2021, thus spanning six football seasons. On-pitch training and competition total distance and high-intensity distance above 5.5m/s for the Under-18, Under-23 and First teams were recorded using

portable 10 Hz global positioning system units (Catapult Sports, Melbourne, Australia; S5 Optimeye prior to the 2018/19 season then S7 Vector) or semi-automated multiple camera (Tracab®, Chyronhego Optical Tracking, New York, USA) tracking systems, which can be used interchangeably (Buchheit et al., 2014; Taberner et al., 2020). The same measures were collected from the Under-12 to Under-16 teams using a foot-mounted inertial tracking system (Playermaker 1.3.17, Playermaker™, London, UK). The acute distance for each participant was calculated as the sum of distances covered in the most recent seven-day period, including the present day. The chronic distance for each participant was calculated as the sum of distances covered in the twenty-one-day period prior to the acute period divided by three, to provide a weekly average. The absolute difference between the acute and chronic distances was than calculated for each player and every session to provide the acute-chronic difference for both total distance (TD_{ACDiff}), and high-intensity distance (HID_{ACDiff}). For example, TD_{ACDiff} = acute total distance – chronic total distance. The z-score for both TD_{ACDiff} and HID_{ACDiff}, was then calculated and loading exposure classified based on the mean and standard deviation of each team's TD_{ACDiff} and HID_{ACDiff} within each season. Therefore, the differences in total distance and high-intensity distance, which might be expected between these tracking systems (Waldron et al., 2021), should not affect the loading exposure analysis, as each device was only compared with itself. The loading exposure was then discretised using the z score as follows: ≥1 was considered a period of substantial overload; between -1 and 1 was considered to normal variation in load exposure; and ≤-1 a substantial de-load.

6.2.9 Data integrity and quality assurance

The genotype call rate of all samples tested was greater than 94.6% and a negative control was included on each plate to check for non-specific amplification. Genotypes were not automatically assigned to samples which failed to amplify consistently with the rest of the cluster and genotype classifications were confirmed by visual inspection of cluster plots by two of the main investigators (YM and ERM) using SNPViewer2 (KBiosciences UK Ltd., Hoddesdon, Herts, UK) as recommended (Semagn et al., 2014). The intra-tester reliability of anthropometric measurements showed a technical error of measurement of 0.2 cm for stature and 0.3 cm for seated height, equating to a relative technical error of measurement of 0.1% and 0.4% respectively. The type of occurrence, OSIICS code, onset and mechanism of each injury were verified by cross-reference with injury rehabilitation / management notes and, where possible, video confirmation from match analysis archives. Any discrepancies in injury data, which could not be clarified with medical notes, were resolved and confirmed via consultation with medical staff who recorded the initial assessment and/or injury management notes.

6.2.10 Statistical analysis

Initially, one model was produced for each tissue injury classification to examine the relationship between TGS, loading, growth or maturation variables, and injury incidence using generalised linear mixed models with PROC GLIMMIX in SAS OnDemand for Academics (SAS Institute Inc., Cary, NC, USA). Injury occurrence, dichotomously coded as either injured or non-

injured, was entered as the dependent variable. The TGS, growth or maturation, and loading exposure variables were entered, in that order, as independent variables and participant identification number was included as a random effect. The loading and growth or maturation variables were entered into the model based on expert opinion and previous literature, as shown in Table 26. Risk ratios (RR) and 95% confidence intervals (CIs) between TGS risk (low, medium, and high), growth (>0.6 cm.m^{-1} and <0.6 cm.m^{-1}), maturation (Pre-, Circa-, and Post-PHV) and loading groups (TD_{ACDiff} or HID_{ACDiff}) were used to evaluate differences in injury risk with statistical significance set as $p \leq 0.05$. No adjustment for multiple testing was selected due to the hypothesis-driven approach to the experimental design (Gibbon et al., 2020). The Hardy–Weinberg Equilibrium (HWE) of overall genotype frequencies (cases and controls) of included variants and those within each ethnicity group were evaluated using $\chi 2$ tests.

6.3 Results

The descriptive data of included participants, based on age at recruitment, and the injuries they experienced is shown in Table 27. Injuries occurred across multiple age groups throughout the surveillance period. The total injury surveillance period included 38,907 elite football training and match exposure hours across six seasons, with an overall injury incidence of 6.9 injuries per 1000 hours. The injury incidence varied by age group as shown in Table 28, which also shows the injuries recorded in each age group and may include repeated participants injuries and loading data. Figures 8 and 9 show the duration of each participant's injury surveillance period by age and across seasons, respectively. Figure 9 also highlights the seventy-one participants with data included in the apophysitis risk period. The successful genotyping call rate of saliva samples ranged from 97.3 to 100% for included genetic variants. The genotype frequencies of included variants were in HWE ($p > 0.10$) except for the *GDF5* rs143383 variant ($p = 0.04$) for all participants as shown in Table 29 along with genotype frequencies based on participant self-identified ethnicity. The average injury incidence for the different injury analyses in each genetic injury risk categorisation are provided in Table 30.

No significant tissue-specific injury risk model was observed when including TGS, maturation, and loading factors for non-contact, non-contact muscle, fracture, joint and tendon injuries ($p > 0.16$) (Table 26). However, *COL1A1* (rs1800012) T allele carriers and growth rate >0.6 cm.m^{-1} were significantly associated with apophysitis injury incidence when TD_{ACDiff} was also included in the model ($p < 0.04$). Those with the *COL1A1* (rs1800012) TT or TG genotype and those with growth rate >0.6 cm.m^{-1} were associated with a significantly greater risk of apophysitis than G allele homozygotes (RR = 4.03 [95% CI 1.10-14.68], $p = 0.03$) and those growing <0.6 cm.m^{-1} (RR = 3.57 [95% CI 1.25-10.16], $p = 0.02$). No significant difference was observed between TD_{ACDiff} groups in this model ($p = 0.71$). When added into the model, the interaction between *COL1A1* (rs1800012) genotype and high growth rate category was not significant ($p = 0.2931$) but high growth rate ($p = 0.03$) and *COL1A1* (rs1800012) genotype ($p = 0.05$) remained significant and TD_{ACDiff} non-significant ($p = 0.72$). Nevertheless, post-hoc analysis indicated that GG individuals with high growth rate were nearly six times as likely to sustain an apophysitis injury than GG Individuals at low growth rate (RR = 5.80 [95% CI 1.47-22.89], $p = 0.01$). The greatest difference in injury risk was observed between T allele

carriers experiencing rapid growth and G homozygotes at low growth (RR = 11.92 [95% CI 2.00-70.72], p = 0.006) but T allele carriers were also still at greater risk of apophysitis injury when both groups experience low growth (RR = 6.27 [95% CI 1.40-28.10], p = 0.02).

Table 26. Injury incidence risk model in elite male football players.

Injury Model	AIC	Fixed Factors	F-Value	P-Value	Absolute Injury Risk Estimate Per Player Per Session (%)
Non-contact injury	2793	TGS injury risk category	1.02	0.36	Overall = 0.64 (0.48-0.86) High = 0.78 (0.51-1.18) Medium = 0.58 (0.41-0.81) Low = 0.59 (0.41-0.86)
		Maturation status	0.38	0.68	Post-PHV = 0.57 (0.45-0.72) Circa-PHV = 0.69 (0.43-1.09) Pre-PHV = 0.68 (0.39-1.19)
		HID_{ACDiff}	0.05	0.95	Overload = 0.67 (0.44-1.01) Normal training = 0.63 (0.48-0.82) Deload = 0.64 (0.42-0.98)
Non-contact muscle injury	1604	TGS injury risk category	0.74	0.48	Overall = 0.22 (0.14-0.36) High = 0.18 (0.09-0.35) Medium = 0.24 (0.14-0.39) Low = 0.26 (0.15-0.45)
		Maturation Status	0.56	0.57	Post-PHV = 0.28 (0.20-0.38) Circa-PHV = 0.24 (0.12-0.48) Pre-PHV = 0.17 (0.06-0.46)
		HID_{ACDiff}	0.68	0.51	Overload = 0.27 (0.15-0.49) Normal training = 0.23 (0.15-0.36) Deload = 0.18 (0.09-0.35)
Joint injury	918	TGS injury risk category	0.38	0.69	Overall = 0.10 (0.04-0.23) High = 0.13 (0.07-0.27) Medium = 0.12 (0.06-0.23) Low = 0.06 (0.01-0.45)
		Maturation Status	0.42	0.65	Post-PHV = 0.11 (0.05-0.24) Circa-PHV = 0.06 (0.02-0.24) Pre-PHV = 0.12 (0.04-0.42)
		TD_{ACDiff}	0.35	0.70	Overload = 0.09 (0.03-0.26) Normal training = 0.11 (0.05-0.26) Deload = 0.09 (0.03-0.26)
Tendon injury	2056	TGS injury risk category	0.43	0.65	Overall = 0.00 (0.00-0.00) High = 0.00 (0.00-0.00) Medium = 0.00 (0.00-0.00) Low = 0.00 (0.00-0.00)
		Maturation Status	0.27	0.76	Post-PHV = 0.03 (0.01-0.10) Circa-PHV = 0.01 (0.00-0.13) Pre-PHV = 0.00 (0.00-0.00)
		TD_{ACDiff}	0.27	0.76	Overload = 0.00 (0.00-0.00) Normal training = 0.00 (0.00-0.00) Deload = 0.00 (0.00-0.00)
Fracture injury	328	TGS injury risk category	0.43	0.48	Overall = 0.03 (0.01-0.11) High = 0.03 (0.00-0.21) Medium = 0.04 (0.01-0.17) Low = 0.02 (0.00-0.09)
		Maturation Status	1.81	0.16	Post-PHV = 0.01 (0.00-0.05) Circa-PHV = 0.04 (0.01-0.22) Pre-PHV = 0.04 (0.01-0.24)
		TD_{ACDiff}	0.54	0.59	Overload = 0.04 (0.01-0.19) Normal training = 0.04 (0.01-0.14) Deload = 0.01 (0.00-0.13)
Apophysitis injury	323	*COL1A1* T Allele Carriers	4.46	*0.03*	Overall = 0.07 (0.02-0.21) TT+TG = 0.14 (0.04-0.52) GG = 0.03 (0.01-0.11)
		High Growth Rate	5.68	*0.02*	Rapid growth >0.6cm^{-1} = 0.13 (0.03-0.49) Non-rapid growth = 0.04 (0.01-0.11)
		TD_{ACDiff}	0.34	0.71	Overload = 0.05 (0.01-0.24) Normal training = 0.07 (0.03-0.21) Deload = 0.09 (0.03-0.35)

Note: p-values in italics indicate significant fixed effects at $p < 0.05$. * indicates significant pairwise difference between fixed effect groups $p < 0.05$. AIC: Akaike information criterion. TGS: Total genotype risk score; TD: Total distance; HID; High-intensity distance (>5.5m/s); PHV: Peak height velocity; ACDiff: Acute to chronic loading difference. *COL1A1* is the *COL1A1* Sp1 (rs1800012) G→T single nucleotide polymorphism. % risk representative of per player per session.

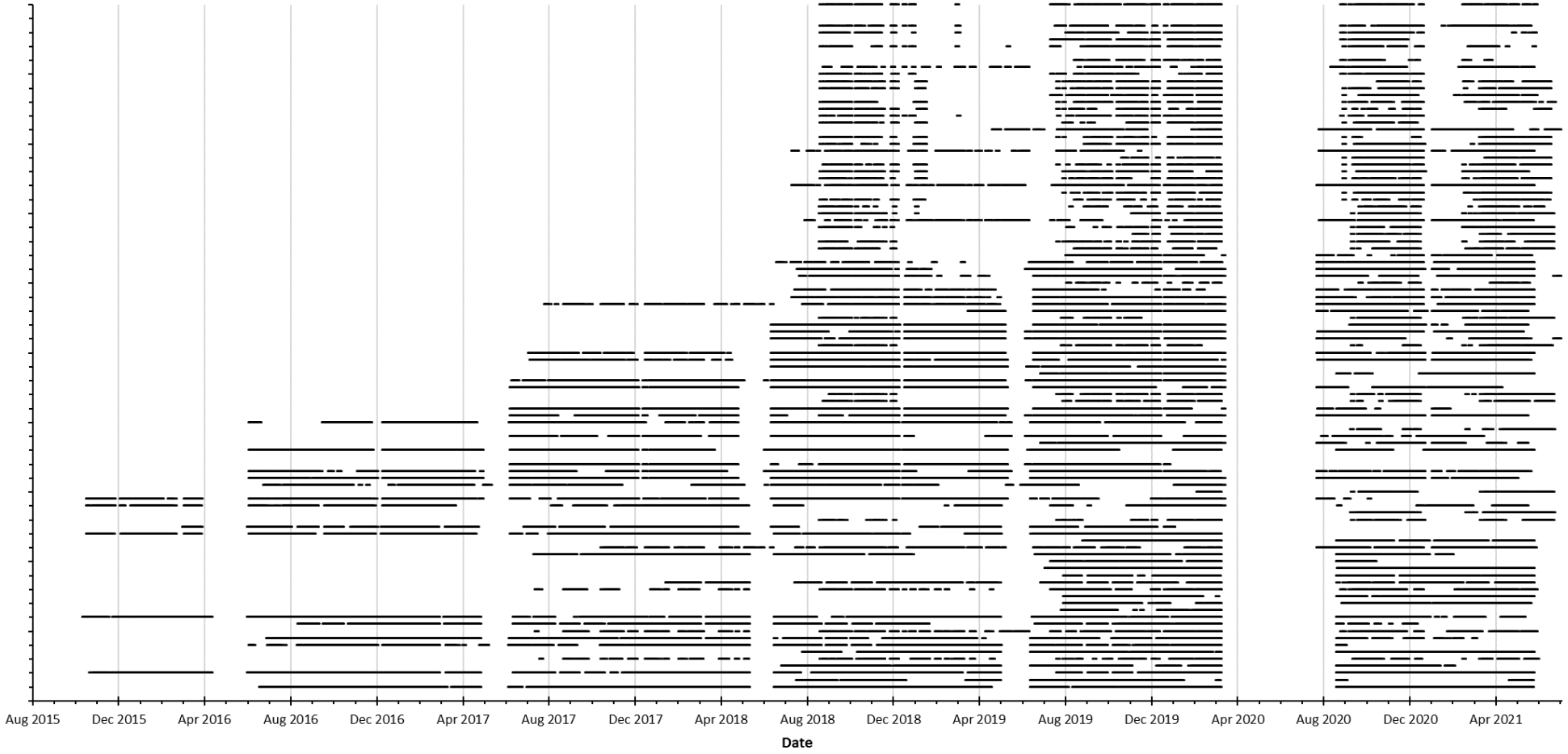

Figure 16. Sessions recorded throughout the surveillance period for each participant by date. Note: each line on the y-axis represents an individual anonymous participant.

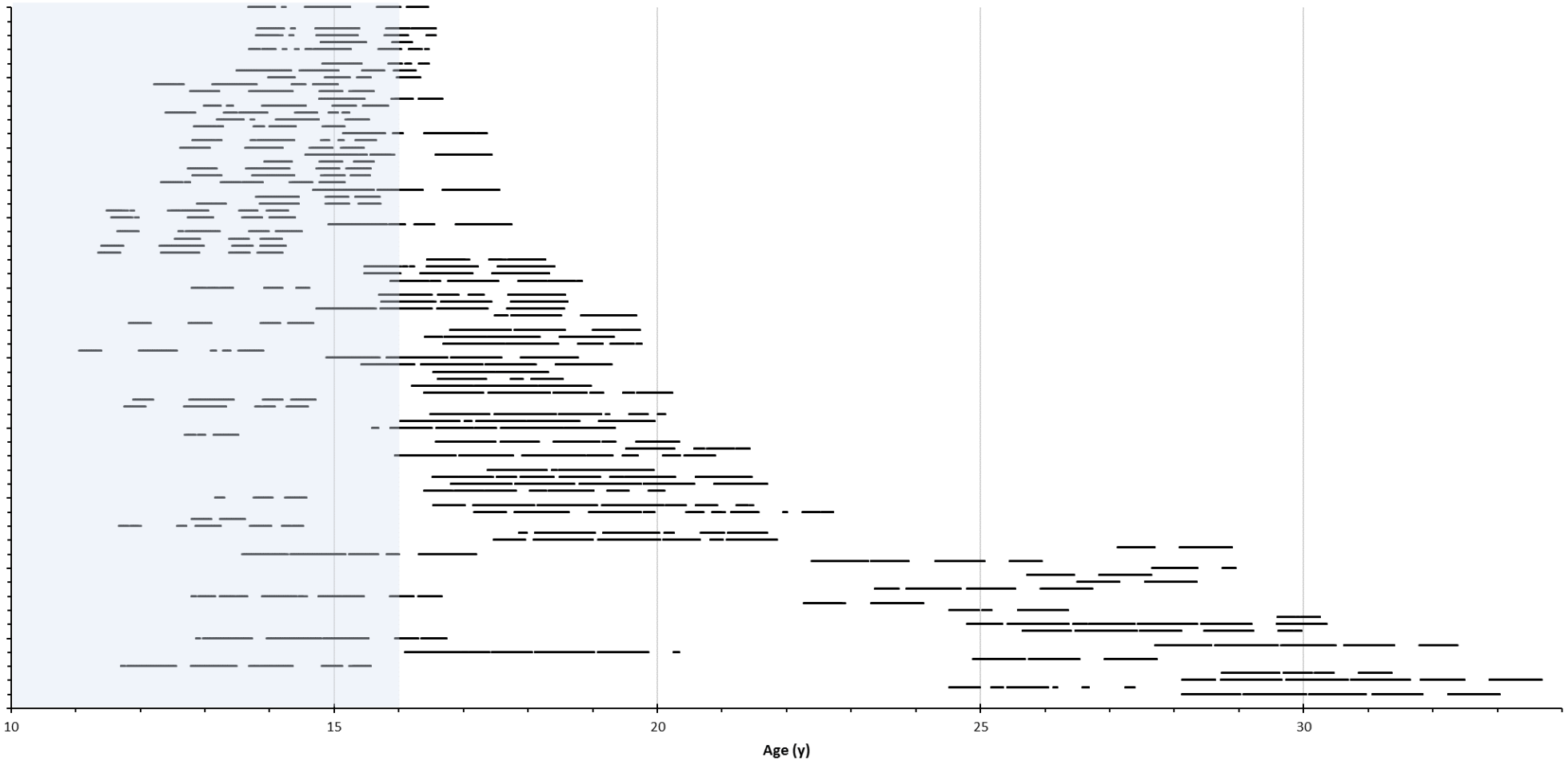

Figure 17. Participant recorded sessions by age during the surveillance period. Note: shaded region includes participants with data in the apophysitis injury risk period.

Table 27. Descriptive statistics of included participants and injury summary by age group at recruitment for elite male youth footballers.

Age group	Under-12	Under-13	Under-14	Under-15	Under-16	Under-18	Under-23	First Team	Total
Included participants (*n*)	2	13	14	12	5	15	16	18	95
All injuries count (*n*)	4	21	31	22	11	60	114	76	339
Injuries per player (n)	2 ± 0	2 ± 1	2 ± 1	2 ± 2	2 ± 1	4 ± 1	7 ± 4	7 ± 4	4 ± 3
Total Recorded Football (hours)	113 ± 29	235 ± 63	260 ± 47	301 ± 114	515 ± 92	681 ± 189	932 ± 242	597 ± 338	516 ± 321
Player injuries per 1000 hours	18.4 ± 4.7	7.4 ± 5.8	8.8 ± 5.9	6.1 ± 6.5	4.4 ± 2.9	6.2 ± 2.3	7.8 ± 4.1	8.7 ± 7.2	7.7 ± 5.6
Non-contact injuries count (*n*)	2	14	18	19	7	33	76	63	232
Non-contact injuries per player (n)	1 ± 0	1 ± 1	1 ± 1	2 ± 2	1 ± 1	2 ± 1	5 ± 3	4 ± 3	2 ± 2
Player non-contact injuries per 1000 hours	9.2 ± 2.3	4.8 ± 5.6	5.2 ± 5.2	5.3 ± 5.6	2.8 ± 2.3	3.6 ± 2.4	5.3 ± 3.4	7.7 ± 7.2	5.4 ± 5.1
Non-contact muscle injuries count (*n*)	1	2	8	9	2	16	46	37	121
Non-contact muscle injuries per player (n)	1 ± 1	0 ± 0	1 ± 1	1 ± 1	0 ± 1	1 ± 1	3 ± 2	2 ± 2	1 ± 2
Player non-contact muscle injuries per 1000 hours	3.8 ± 5.3	0.5 ± 1.3	2.3 ± 2.6	2.2 ± 2.8	0.9 ± 1.3	1.5 ± 1.1	3.2 ± 2.0	4.3 ± 4.4	2.4 ± 2.9
Tendon injuries count (*n*)	0	1	0	0	1	7	10	8	27
Tendon injuries per player (*n*)	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 1	1 ± 1	0 ± 1	0 ± 1
Player tendon injuries per 1000 hours (*n*)	0.0 ± 0.0	0.3 ± 1.0	0.0 ± 0.0	0.0 ± 0.0	0.3 ± 0.7	1.0 ± 1.8	0.7 ± 1.1	0.7 ± 1.8	0.5 ± 1.3
Joint injuries (*n*)	1	4	3	2	2	13	28	10	63
Joint injuries per player (*n*)	1 ± 1	0 ± 0	0 ± 0	0 ± 0	0 ± 1	1 ± 1	2 ± 2	1 ± 1	1 ± 1
Player joint injuries per 1000 hours	3.8 ± 5.3	1.6 ± 2.7	0.8 ± 1.6	0.8 ± 2.2	0.8 ± 1.0	1.2 ± 1.0	1.8 ± 1.6	1.4 ± 3.2	1.3 ± 2.2
Fracture injuries (*n*)	1	2	5	0	2	4	3	2	19
Fracture injuries per player (*n*)	1 ± 1	0 ± 1	0 ± 0	0 ± 0	0 ± 1	0 ± 0	0 ± 1	0 ± 0	0 ± 0
Player fracture injuries per 1000 hours	5.4 ± 7.7	0.8 ± 2.8	1.5 ± 2.1	0.0 ± 0.0	0.9 ± 1.9	0.5 ± 0.8	0.2 ± 0.6	0.1 ± 0.4	0.6 ± 1.8
Apophysitis injuries (*n*)	1	6	5	8	1	0	0	0	21
Apophysitis injuries per player (*n*)	1 ± 1	0 ± 1	0 ± 1	1 ± 1	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Player apophysitis injuries per player 1000 hours	5.4 ± 7.7	2.1 ± 3.2	1.3 ± 2.4	2.3 ± 4.1	0.4 ± 1.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.9 ± 2.5
Age at start of observation (years)	12.8 ± 0.1	11.9 ± 0.6	12.8 ± 0.6	13.8 ± 0.7	14.6 ± 0.6	16.0 ± 0.8	16.8 ± 0.9	25.5 ± 3.2	16.5 ± 4.9
Age at end of observation (years)	13.5 ± 0.1	14.4 ± 0.2	15.5 ± 0.2	16.4 ± 0.3	17.4 ± 0.2	18.8 ± 0.5	20.7 ± 1.0	28.5 ± 3.3	19.4 ± 5.1
Recorded Sessions (***n***)	76 ± 20	155 ± 40	188 ± 40	219 ± 91	374 ± 71	503 ± 148	733 ± 182	545 ± 306	404 ± 270

Note: Data presented as means ± standard deviations of the measurements within the age group of participants at recruitment which may include injuries sustained in other age groups included in the observation period.

Table 28. Training and match exposure time, injury count and injury per 1000 hours recorded in each age group team in elite male youth footballers.

Age group	Under-12	Under-13	Under-14	Under-15	Under-16	Under-18	Under-23	First Team	Total
Training exposure hours	553	1986	3137	2803	1839	11404	7259	9927	38907
Match exposure hours	101	298	431	835	759	3039	2085	2353	9902
Total exposure hours	653	2284	3569	3641	2618	14529	9405	12283	48983
Training injuries	4	7	19	19	7	48	46	44	194
Match injuries	1	4	5	16	11	40	28	39	144
Total Injuries	5	11	24	35	18	89	74	83	339
Training injuries per 1000 hours	7.2	3.5	6.1	6.8	3.8	4.2	6.3	4.4	5.0
Match injuries per 1000 hours	9.9	13.4	11.6	19.2	14.5	13.2	13.4	16.6	14.5
Total injuries per 1000 hours	7.7	4.8	6.7	9.6	6.9	6.1	7.9	6.8	6.9
Non-contact injuries	2	8	17	24	12	47	52	70	232
Non-contact injuries per 1000 hours	3.1	3.5	4.8	6.6	4.6	3.2	5.5	5.7	4.7
Non-contact muscle injuries	0	1	6	12	4	26	30	42	121
Non-contact muscle injuries per 1000 hours	0.0	0.4	1.7	3.3	1.5	1.8	3.2	3.4	2.5
Tendon injuries	0	0	1	0	1	9	8	8	27
Tendon injuries per 1000 hours	0.0	0.0	0.3	0.0	0.4	0.6	0.9	0.7	0.6
Joint injuries	0	1	6	4	2	19	19	12	63
Joint injuries per 1000 hours	0.0	0.4	1.7	1.1	0.8	1.3	2.0	1.0	1.3
Fracture injuries	1	2	3	2	2	5	1	3	19
Fracture injuries per 1000 hours	1.5	0.9	0.8	0.5	0.8	0.3	0.1	0.2	0.4
Apophysitis injuries	2	5	6	5	3	0	0	0	21
Apophysitis injuries per 1000 hours	3.1	2.2	1.7	1.4	1.1	0.0	0.0	0.0	0.4
Total distance acute (km)	12.3 ± 5.3	15.9 ± 7.2	16.7 ± 7.5	18.7 ± 8.3	20.7 ± 9.0	30.6 ± 9.3	29.1 ± 8.8	25.5 ± 8.8	25.8 ± 10.0
Total distance chronic (km)	8.0 ± 5.3	10.1 ± 6.5	11.6 ± 6.9	13.7 ± 7.8	16.0 ± 8.6	26.0 ± 10.1	25.0 ± 10.1	21.1 ± 9.0	21.2 ± 10.5
Total distance acute-chronic difference (km)	4.3 ± 7.4	5.8 ± 7.3	5.1 ± 7.3	5.1 ± 8.4	4.7 ± 9.6	4.6 ± 11.4	4.1 ± 10.8	4.4 ± 11.6	4.6 ± 10.7
High-intensity distance acute (m)	49 ± 51	195 ± 206	379 ± 322	667 ± 495	960 ± 542	1751 ± 744	1813 ± 818	1416 ± 762	1379 ± 867
High-intensity distance chronic (m)	33 ± 32	116 ± 131	261 ± 250	501 ± 405	763 ± 464	1508 ± 674	1575 ± 749	1189 ± 624	1165 ± 766
High-intensity distance acute-chronic difference (m)	16 ± 59	78 ± 157	118 ± 254	166 ± 421	197 ± 522	244 ± 815	244 ± 888	229 ± 867	215 ± 769
Recorded Sessions (***n***)	345	1591	2477	2772	1948	10715	7332	11198	38378

Note: Data presented as means ± standard deviations of that recorded in each age group team which may include multiple injuries sustained by the same individual over time and different age groups. High-intensity distance > 5.5 m/s.

Table 29. Participant self-reported ethnicity 1000 Genomes Project super population genotype frequency distribution.

Candidate Gene Variant Genotype	1000G Super Population Ethnicity Code						
	African		European		Mixed		
					African & European	All Participants	
	Observed	1000G	Observed	1000G	Observed	Observed	1000G
ACTN3 (rs1815739)							
TT (XX)	0 (0)	1	14 (7)	18	0 (0)	7 (7)	18
TC (XR)	14 (4)	22	46 (23)	51	46 (6)	36 (34)	44
CC (RR)	86 (24)	78	40 (20)	31	54 (7)	56 (53)	38
COL1A1 (rs1800012)							
TT	0 (0)	1	6 (3)	4	0 (0)	3 (3)	2
TG	7 (2)	12	26 (13)	29	15 (2)	18 (17)	15
GG	93 (27)	88	68 (34)	67	85 (11)	79 (75)	83
COL1A2 (rs412777)							
CC	7 (2)	11	8 (4)	16	8 (1)	9 (8)	12
CA	21 (6)	46	57 (28)	45	38 (5)	43 (40)	43
AA	71 (20)	43	35 (17)	39	54 (7)	48 (45)	46
COL5A1 (rs12722)							
TT	3 (1)	1	36 (18)	36	23 (3)	24 (23)	16
TC	38 (11)	26	52 (26)	46	0 (0)	39 (37)	38
CC	59 (17)	72	12 (6)	19	77 (10)	37 (35)	46
GDF5 (rs143383)							
AA	7 (2)	0	35 (17)	42	0 (0)	22 (20)*	29
AG	17 (5)	6	46 (22)	42	38 (5)	35 (33)*	33
GG	76 (22)	94	19 (9)	16	62 (8)	43 (40)*	38
MMP3 (rs679620)							
CC	52 (15)	45	16 (8)	28	31 (4)	29 (28)	43
CT	41 (12)	45	50 (25)	50	54 (7)	48 (46)	44
TT	7 (2)	11	34 (17)	22	15 (2)	22 (21)	13
VDR (rs2228570)							
AA	0 (0)	4	10 (5)	16	0 (0)	6 (6)	13
AG	25 (7)	31	46 (23)	44	23 (3)	35 (33)	41
GG	75 (21)	66	44 (22)	40	77 (10)	59 (55)	47
Participants (n)	29		50		13	95	

Note: Genotype frequencies are expressed as a percentage with the number of observed participants (n) in parentheses. 1000G represents genotype frequency data from the 1000 Genomes project phase 3. Data are not presented for the East Asian and Mixed East Asian & Ad Mixed American participants (n=3) to protect data anonymity. * Indicates genotype frequencies not in Hardy-Weinberg Equilibrium $p<0.05$.

Table 30. Injury incidence by total genotype risk score category.

Injury	Total Genotype Injury Risk Score Category					
	High Genetic Risk		Medium Genetic Risk		Low Genetic Risk	
	Participants (n)	Injuries per 1000 hours	Participants (n)	Injuries per 1000 hours	Participants (n)	Injuries per 1000 hours
Non-contact injury	19	6.4 ± 5.3	44	5.2 ± 5.3	32	4.9 ± 4.7
Non-contact muscle injury	19	1.6 ± 2.2	48	3.0 ± 3.4	28	1.9 ± 2.1
Joint injury	31	1.6 ± 2.5	60	1.2 ± 2.1	4	0.8 ± 1.7
Tendon injury	24	0.3 ± 0.9	40	0.6 ± 1.5	31	0.5 ± 1.2
Fracture injury	12	0.3 ± 0.7	39	0.9 ± 2.0	44	0.5 ± 1.8
Apophysitis injury	16	2.7 ± 4.5			55	0.8 ± 1.9

Note: Apophysitis analysis includes sub-group of participants who had injury and loading data at age <16 years.

6.4 Discussion

The main finding of the present study was that no significant tissue-specific injury risk model was observed when including TGS, growth or maturation and loading factors ($p > 0.16$), except for apophysitis injury ($p < 0.04$). Possession of the *COL1A1* (rs1800012) T allele (TT & TG individuals) showed a four-fold increase in apophysitis injury risk ($p = 0.03$) compared to GG individuals with growth rate >0.6 $cm.m^{-1}$ and appeared to increase the risk of injury more than three-and-a-half times that of slower growth rates ($p = 0.02$). Type 1 collagen is the major structural protein of bone tissue and is typically formed of subunits composed of two collagen type I α1 (COL1A1) and one collage type I α2 (COL1A2) procollagen sub-units (Ghosh, 2002; Myllyharju & Kivirikko, 2001; Tzaphlidou, 2008). The *COL1A1* rs1800012 T allele is associated with greater Sp1 binding affinity, increased production of COL1A1 and, consequently, type 1 procollagen formed exclusively of three COL1A1 polypeptides (Mann et al., 2001). Type 1 collagen formed with a high proportion of COL1A1 only procollagen subunits is thought to be weaker than the COL1A1 / COL1A2 combination, as the T allele has repeatedly been associated with increased risk of osteoporotic fracture in the elderly (Mann et al., 2001; Mann & Ralston, 2003). However, the mechanisms behind this effect remain unclear as the T allele has also been linked with protection against ACL ruptures (Ficek et al., 2013; Khoschnau et al., 2008; Posthumus et al., 2009a) and fracture risk in young physically active females in chapter three. Nevertheless, the T allele has also been associated with increased risk of ACL rupture (Stępień-Słodkowska et al., 2013) and bone fractures in pre-pubertal children (Blades et al., 2010), with others observing no influence on fracture risk in physically active adults (Cosman et al., 2013; Korvala et al., 2010; Varley et al., 2018).

The divergent injury risk associations observed for the *COL1A1* (rs1800012) T allele appear to vary by physiological ageing and developmental processes. The T allele has been associated with lower bone mineral density and reduced bone formation relative to resorption in early pubertal females (Suuriniemi et al., 2006) and greater bone loss in elderly post-menopausal women (Brown et al., 2001). Therefore, it is possible that the formation of type 1 collagen with increased proportions of COL1A1 only procollagen subunits, associated with the T allele, affects fracture risk by influencing bone homeostasis and regulation. This could explain why the effect of the T allele on fracture risk varies with age. Peak bone mass and size is achieved between ages 25 to 30 years (Abrams, 2003), with peak acquisition occurring between 15-20 years (Raisz & Seeman, 2001), when bone fragility is naturally at its lowest in healthy individuals (Abrams, 2003). Bone mineralisation increases prior to these ages and decreases after as bone turnover slows and resorption begins to increase with aging (Demontiero et al., 2012). Therefore, the reduced bone formation relative to resorption associated with the *COL1A1* (rs1800012) T allele (Brown et al., 2001; Suuriniemi et al., 2006) may not be sufficiently detrimental to significantly impact the individual threshold tolerance of bone tissue, with the exception of young and elderly populations for whom bone tissue is already weaker than adults. Furthermore, the *COL1A1* (rs1800012) T allele may still be protective against ligament injury in young physically active participants if the three COL1A1 procollagen type 1 collagen tissue is stronger when incorporated into ligamentous tissue. Ligament tissue may be less susceptible to the influence of tissue turnover with aging as it appears to have a higher tissue-specific protein synthesis rate (Smeets et al., 2019). However, the T allele has also been linked with greater muscle soreness and impaired strength recovery following exercise induced muscle damage compared to G allele

carriers (Baumert et al., 2018). This may also suggest tissue weakness and the protective effect which has been observed for the *COL1A1* (rs1800012) T allele may be due to an alternative, currently unknown mechanism.

Apophysitis injuries are generally insidious non-contact injuries resulting from repeated microfractures at the apophysis of bone, which is unable to withstand shear forces exerted by the tendon during exercise (Arnold et al., 2017; Gholve et al., 2007; Holden & Rathleff, 2020; Ogden & Southwick, 1976). A transient period of increased bone weakness occurs as bone mineral density declines prior to PHV and then rebounds after PHV (Faulkner et al., 2006; Wang et al., 2010; Wang & Seeman, 2009) when bone mass accrual is at its peak (Yilmaz et al., 2005). Therefore, increased apophysitis injury risk may occur due to the interaction between the puberty dependent reduction of bone strength, and T allele carriers lower bone mineral density and formation relative to resorption. The results of the present study suggest that the risk of apophysitis injury was not significantly influenced by loading exposure. Therefore, it could be that genetic susceptibility and growth is sufficient to predispose youth footballers to apophysitis injury with regardless of the level of exposure to football activity. Indeed, the T allele has been previously associated with increased risk of fracture and reduced bone mineral density which was greatest during periods of appendicular bone growth during adolescence (Blades et al., 2010). Nevertheless, Brown et al. (2001) found the influence of the *COL1A1* (rs1800012) T allele was significantly affected by dietary calcium intake. Those with low calcium intake lost significantly more bone than GG homozygotes, but those of high calcium intake actually gained significantly more bone (Brown et al., 2001). Others have shown that bone collagen synthesis can occur rapidly and is stimulated by nutritional stimulation (Babraj et al., 2005) and that increased dietary calcium intake for TT and TG individuals may be protective against apophysitis injury in elite male youth footballers without the need to restrict their football practice time. Nevertheless, further replication of the *COL1A1* (rs1800012) T allele association with apophysitis injury risk is required to assert confidence of the effect. Future research could explore the influence of dietary calcium intake on apophysitis injury risk in elite male youth footballers with the genotypes of the *COL1A1* (rs1800012) SNP.

Apophysitis injuries have been suggested to be more prevalent around PHV (Price et al., 2004; Read et al., 2018b; Rumpf & Cronin, 2012) and significant associations between rapid annual growth and apophysitis injury have been observed Wik et al. (2020b). However, sensitively identifying individuals for whom rapid growth will significantly increase injury risk more frequently (every 3 to 12 weeks) is challenging as shown in Chapter four. The findings of the present study suggest that inclusion of *COL1A1* (rs1800012) T allele possession and TD_{ACDiff} into the model improves the sensitivity of risk identification. The results of the present study suggest that growth rate >0.6 $cm.m^{-1}$ increases the risk of apophysitis injury more than three times that of slower growth rates ($p = 0.02$). In absolute terms, this may only translate to a 0.09% increase in apophysitis injury risk per player per session. However, considering a typical under-14 age group, with twenty players and four sessions per week, this increase in risk equates to approximately three more apophysitis injuries for those of rapid growth (~4 injuries) than those not (~1 injury) over a 42-week season. Apophysitis injuries were present in the under-12, despite a smaller sample representation, to under-16 age group in the present study. Therefore, as support staff often manage multiple youth age groups in an applied elite youth football setting, this increase in injury incidence is practically meaningful.

As with varied growth rates and loading exposure, individuals possessing risk alleles may not become injured, while others without risk alleles can still experience injury, indicating how the likelihood of injury resulting from possession of a risk allele varies depending on other factors and highlighting the complexity of injury aetiology (Bittencourt et al., 2016; Lander & Schork, 1994; Tee et al., 2020). Indeed, in the present study, apophysitis injury risk was significantly different between combinations of growth (>0.6 $cm.m^{-1}$ or not) and *COL1A1* (rs1800012) genotype (TT+TG *vs.* GG) groups. The greatest difference in injury risk was observed for T allele carriers experiencing rapid growth who were almost twelve times more likely to suffer an apophysitis injury than GG homozygotes at low growth ($p = 0.006$). Even among GG homozygotes, those experiencing growth >0.6 $cm.m^{-1}$ were almost six times more likely to experience apophysitis injury compared with those of lower growth ($p = 0.01$). However, T allele carriers were six-times more likely to suffer apophysitis injury even when both they and G allele homozygotes were experiencing low growth ($p = 0.02$). Together, these findings support the association observed between the *COL1A1* (rs1800012) T allele and apophysitis injury risk and may provide greater sensitivity to identify individuals at risk of injury when stature is measured more frequently as in chapter four.

Aside from that described above, there were no other models in this study significantly associated with injury risk ($p > 0.16$). Others have found a significantly higher TGS for injured than non-injured youth footballers using three SNPs included in the present study (*ACTN3*, *COL5A1* & *MMP3*) and four alternative SNPs (*EMILIN1*, *IL6*, *MYLK* & *VEGFA*) (Hall et al., 2022). However, this was found for the incidence of all observed injuries, which may restrict the specificity of applied interventions and did not account for differences in training exposure, growth or maturation (Hall et al., 2022). Nevertheless, Hall et al. (2022) did report several maturation-dependent genetic associations, including an increased risk of all injuries and ligament injury for pre-PHV *COL5A1* (rs12722) CC homozygotes compared with T allele carriers. The *COL5A1* (rs12722) CC genotype was considered low risk for joint injuries in the present study, based on previous research (Lulińska-Kuklik et al., 2018; Pabalan et al., 2018) and no significant association was found between TGS category nor maturation status in the joint injury model. The different findings discussed may result from a greater number of total, and younger, players included by Hall et al. (2022) (age range 9-23 years) to that of the current study. it may be that CC individuals were incorrectly defined as the low-risk genotype for joint injury in the present study for all maturation groups. Nevertheless, in the current study the incidence of joint injury did appear to increase stepwise from low-to-high genetic risk of injury groups (Table 30). Hall et al. (2022) suggest that the genetic associations with injury observed in Pre-PHV players indicated an increased importance of genes responsible for the structure and mechanical integrity of muscle, tendon, and ligament compared with post-PHV players for whom the inflammatory response also appears important. Therefore, it is possible that age / maturation based TGS may need to be developed, as the interaction and understanding of the genetic penetrance of genetic variants with injury risk progresses.

The other TGS risk models in the present study did not appear to significantly influence tissue-specific injury incidence. The tissue-specific genetic associations of variants included in each TGS have each been independently replicated. Therefore, it may be that the TGS risk allocation or calculation was unable to capture the combined influence of the included variants. The aim of the present study was to examine how TGS for tissue-specific injury may interact with physical loading

and growth or maturation. However, the influence of genetic penetrance was not considered when developing the TGSs. Risk genotype allocation was specific to each tissue injury classification and based on previous findings in similar participants to that in the present study, but it may be that the genetic penetrance of included variants also needs to be considered in the formation of TGS. Consider, for example, exercise-induced muscle damage, which is a complex trait, mediated by several genetic variants (Del Coso et al., 2017b). The X allele of the *ACTN3* (rs1815739) SNP has been associated with increased muscle damage following exercise (Vincent et al., 2010) and appears to influence non-contact muscle injury in adults (Zouhal et al., 2021). However, this effect may only be deleterious in adult athletes for whom recovery from exercise may be further reduced due to age-related changes in adaptation (Falk & Dotan, 2006; Hebestreit et al., 1993). Thus, the *ACTN3* (rs1815739) XX genotype may not influence non-contact injury in children and adolescence for whom recovery is more rapid (Falk & Dotan, 2006; Hebestreit et al., 1993). However, adult and older players may be at greater risk when training load and match competition demands increase, and thus muscle damaging forces, as indicated in Table 28. Therefore, if RR adults are protected from exercise-induced muscle damage, compared with X allele carriers, and have improved adaptive signalling and physical performance recovery (Belli et al., 2017; Del Coso et al., 2016, 2017a; Pimenta et al., 2012; Vincent et al., 2010), then RR adult footballers would be expected to be protected from non-contact injury compared with X allele carriers, while no association is observed in children. Therefore, it may be that the X allele is only associated with variations in injury risk above a certain age. Consequently, genetic penetrance may need to be considered when genotype risk allocations are determined for TGS to improve specificity with injury incidence. Nevertheless, the results of the current study suggest that the TGS for tissue-specific injury used in the present study are not associated with tissue-specific injury incidence in elite male youth footballers. Nevertheless, the incidence of some injuries, such as those to tendon tissue, appeared to be so rare in the present cohort of footballers that this study was not able to achieve an adequate sample of cases within each group to draw robust results. Therefore, it is possible that significant association may have been observed due to a lack of injury cases.

Another study exploring the influence of TGS with injury risk in twenty-seven elite male footballers claimed that their model, including the *ACTN3* (rs1815739) and *COL5A1* (rs12722) variants, predicted the number of injuries a player suffered with 97% accuracy (Montagna et al., 2019). However, the genotype risk score assignment and calculation were not reported with the genetic injury model retrofitted to match injury data (Montagna et al., 2019). The authors acknowledge that due to their low sample size this may result in overfitting and suggest that further validation is needed despite not reporting the model used. The results of Montagna et al. (2019) should, therefore, be treated with some caution and future studies should look to be as transparent as possible to allow validation attempts to be made. Nevertheless, TGS have shown promise in differentiating the proportion of fatigue resistance slow-twitch muscle fibres and maximal oxygen consumption (Ahmetov et al., 2009), aerobic training response (Pickering et al., 2018) and all injury incidence (Hall et al., 2022) in footballers. Despite the emergence of supportive research, little attention appears to have been made on developing or exploring TGS for tissue-specific injuries. Although this study was unable to establish a significant TGS model including multiple genetic variants, the use of TGS for injury risk management has a strong theoretical foundation. As our understanding of the genetic determinants of tissue-specific injury increase, more complex models

that use information from multiple sources to truly individualise the training and management of elite athletes could be designed (Pickering & Kiely, 2019). Indeed, a thorough paper by Grealy et al. (2015) evaluated the influence of 7 SNPs previously associated with endurance but found that TGS alone was unable to differentiate performance and highlighted the need for inclusion of training and environmental factors to improve the model. Nevertheless, studies exploring the influence of environmental factors such as training load on injury risk should also begin to consider the importance of interindividual genetic differences. Indeed, developing individualised intensity and loading threshold capacities may improve the sensitivity of prognostic monitoring tools to detect the at-risk individual. Furthermore, greater understanding is required on how tissue-specific loading and exercise intensity thresholds may be used to reduce injury incidence in adolescent athletes depending on growth, maturation and aging.

The results of this study should be considered alongside its limitations. Nutritional factors appear to play a significant role in the influence of genetic variants for injury risk, but no nutritional controls or monitoring was performed in the present study. This was considered beyond the original scope of the research, which already included both loading and growth / maturation factors. However, nutritional differences could have influenced the findings if significant variability in calcium or vitamin D consumption occurred between groups. Furthermore, despite the observation period extending over multiple years, the incidence of some injuries (tendon, fracture and apophysitis) were particularly low. Therefore, although significant genetic associations were still detected for these injuries the number of potential comparisons observed was low and true genetic associations with small effect may not have been detected. Consequently, the results of this study should not be used with any validity to discriminate or exclude individuals from participation in sport due to the complex interaction between genetic and environmental factors, which are not known with enough confidence to assert such long-term implications. Nevertheless, interventions could be inferred from the results of this study in elite male football, which protect all players from harm and support long-term development, and wellbeing, for every individual.

The genetic variants chosen for the TGSs in the present study were consistently shown to be associated with, and / or result in a plausible physiological consequence to mediate, injury risk. Nevertheless, continued research into the genetic determinants of injury in elite male football appears necessary before TGS for tissue-specific injury may be successfully and comprehensively used to support players' long-term development through protection from injury. Injury aetiology is a multifactorial emergent occurrence (Bittencourt et al., 2016; Tee et al., 2020). Nevertheless, the individual threshold tolerance and injury susceptibility can be both developed with training (Lemes et al., 2021; Malone et al., 2019) and mediated by heritable factors (Andrew et al., 2004; Hakim et al., 2003; Magnusson et al., 2020). Therefore, bespoke training interventions based on individual needs are likely to be more effective at reducing injury occurrence when informed by a greater understanding of the genetic and environmental factors that interact to affect tissue-specific injury susceptibility. The results of the present study suggest that our current ability to sensitively identify individuals at significantly increased risk of tissue-specific injury using TGSs, loading exposure and growth or maturation remains questionable. Nevertheless, the *COL1A1* (rs1800012) SNP appears to significantly influence the risk of apophysitis injury, allowing us to differentiate those at increased risk of injury and for whom nutritional interventions may be effective.

6.5 Conclusion

Possession of the *COL1A1* (rs1800012) T allele and growth rate >0.6 cm.m^{-1} were associated with significantly greater risk of apophysitis injury when compared to CC individuals and those experiencing slower growth when including TD_{ACDiff} in the model. Lower bone mineral density and formation relative to resorption associated with the T allele may interact with a temporal reduction in bone strength around puberty to significantly increase the risk of apophysitis injury during rapid growth. Dietary calcium intake could mitigate and even reverse this increased susceptibility. However, further research and replication of the findings in the present study are required to confirm these findings and hypothesis. Consideration of the dynamic change in the genetic influence of variants (genetic penetrance based on age / maturation etc.) may improve TGSs in future research. Further investigation is needed to develop accurate TGS for tissue-specific injury in elite male football. However, integration of genetic and environmental factors may improve the sensitivity of individual tissue-specific injury risk models, such as apophysitis injury risk, as inherent underlying susceptibilities may only become apparent during different periods of physiological aging, growth, development and maturation.

CHAPTER 7: Discussion

This chapter aims to discuss and critically evaluate key findings of the research and its implications. The aims of the thesis were: 1) To identify candidate genetic variants with potential utility of application to reduce injury risk in elite male youth football player development; 2) To explore how physiological development and candidate genetic variants independently influence the risk of injury in elite male football player development; and 3) To develop and evaluate potential applications of combining genetic, physiological, and environmental information to reduce injury risk in elite male youth football. Chapter three identified a sex-specific protective effect of the *COL1A1* (rs1800012) T allele in young healthy physically active females from the pooled results of published literature. This is in contrast to previous associations with osteoporotic fracture in the elderly and highlights the potential affect that sex and ageing may have on the influence of genetic variants with injury risk. The results of chapter four challenge the applicability of previous research identifying significant associations between growth rate and injury incidence using measurements collected every 5-7 weeks in elite male youth football. This highlights a potential area for application of genetics as growth rate alone was unable to sensitively identify individuals at increased risk of injury. Chapter five examined the relationship between previously identified genetic variants associated with tissue-specific injury incidence in elite male football. The *COL1A2* (rs412777) and *GDF5* (rs143383) SNPs were associated with fracture injury risk, while *MMP3* (rs679620) and *VDR* (rs2228570) were associated with incidence of non-contact injuries. The *COL5A1* (rs12722) SNP was related to non-contact muscle and tendon injury, and *COL1A1* (rs1800012) associated with apophysitis injuries. The results of previous chapters were combined in chapter six to examine the associations between tissue-specific TGS, growth or maturation, and loading exposure on injury incidence. The aim of this final study was to examine how these factors may interact to explore their potential to inform applied interventions and protect players from injury. Possession of the *COL1A1* (rs1800012) T allele and growth rate >0.6 $cm.m^{-1}$ significantly increased apophysitis injury risk. Previous research has indicated that bone strength can be augmented for T allele carriers with dietary calcium intake, which could reduce the risk of apophysitis injury in these individuals. The effectiveness of this genetically individualised intervention was unable to be evaluated in the current project due to the research project deadline and lack of time for further investigation.

7.1 Research-practitioner evaluation of genetic factors to support elite male youth football player development

Completing the research project whilst working as a practitioner within an elite male youth football development context framed the research from an applied perspective. This research-practitioner perspective guided the evolution of the project to explore potentially viable applications of genetic information to support long-term development in elite male football as discussed by Burden et al. (2022). Specifically, as research specific knowledge developed, potential applications for innovation began to become apparent by also drawing upon context specific understanding and identified applied problems.

7.1.1 Context of the applied problem

Considerable research has explored the influence of growth, maturation and aging on long-term athletic development and injury risk in elite male youth footballers (Kemper et al., 2015; Rumpf & Cronin, 2012; Ryan et al., 2018). Pubertal changes appear to impact the risk of different types of injury. The risk of muscle, tendon and ligament injuries increase with age and maturation as shown in chapter four and by others (Monasterio et al., 2021a). However, fracture injuries appear to be more prevalent in childhood and adolescence, prior to peak bone mass accrual during puberty (Parfitt, 1994). Meanwhile, the risk of apophysitis injuries appears to peak in the Under-13 and Under-14 age groups (Read et al., 2018b). This peak in apophysitis injury coincides with the average age of PHV in males (Abbassi, 1998; Tanner et al., 1966) leading authors to hypothesis that these injuries are "growth-related" (Monasterio et al., 2021a; Read et al., 2018b). Subsequently, it has been suggested that training load should be altered during PHV to lower the risk of apophysitis injuries (Arnold et al., 2017; DiFiori et al., 2014). Nevertheless, few studies appear to have directly explored the relationship between growth and apophysitis injuries (Kemper et al., 2015; Wik et al., 2020b) despite this training alteration appearing to become part of the status quo in long-term youth development research (Ryan et al., 2018).

7.1.2 Limitations of growth and maturation only as a risk identification tool

The adolescent PHV is typically observed between 85-96% of PAH (Parr et al., 2020), which broadly align with the Tanner Stages of Sexual Maturity (Malina et al., 2005b; Marshall & Tanner, 1970). The Tanner stages provide an objective classification system for the development of puberty in children but requires genital inspection by trained clinicians or self-assessment using sample pictures (Marshall & Tanner, 1970). This represents an invasive and time-consuming measurement process that is unrealistic to conduct in an applied setting, with hundreds of players, periodically. However, because the original work by Marshall and Tanner (1970) showed that PHV typically occurred later in the maturational development of males, during Tanner genitalia stage 4. The use of somatic measures of height and height growth velocity became considered as viable, non-invasive, alternatives of estimating maturation stage (Malina et al., 2005b, 2007a, 2012). However, the work of Marshall and Tanner (1970) was originally based on white British children and ethnicity dependent variations in pubertal maturation are well known (Sun et al., 2002). Furthermore, even in the original study 22% of boys were found to have experienced PHV during Stage 5 (Marshall & Tanner, 1970) and others have shown that around 30% of boys (who were majority white) still had not experienced PHV even by Tanner stage 5 (Granados et al., 2015). The development of sexual maturity evaluated in the Tanner stages is primarily driven by testosterone in males (Marshall & Tanner, 1970; Rogol et al., 2002). Testosterone also plays a significant role in the development of physical performance capabilities during puberty. Therefore, although useful for practical consideration, somatic estimations of maturation are limited to confidently identify PHV for any given individual (Teunissen et al., 2020). Consequently, caution should be used when advocating for training programme modifications for individuals based on somatic maturation estimates alone.

Periodic measurements of stature and weight could supplement somatic maturity estimates by attempting to directly measure PHV and PWV. However, this requires continuous and longitudinal

measurement to indicate the different stages of pubertal maturation. Furthermore, it is worth acknowledging that even if directly measured, PHV may not fully indicate pubertal development due to the differences observed between PHV and Tanner stages highlighted previously (Granados et al., 2015; Marshall & Tanner, 1970). Indeed, these observations likely reflect the different influence of various hormones responsible for pubertal development, which interact but also exert independent effects (Cutler, 1997; Rogol et al., 2002). The anabolic influence of testosterone, which appears to drive sexual maturation and muscle growth in boys, can interact with growth hormone and estrogen, which in turn effect pubertal stature growth, skeletal maturation and fusion of epiphyseal plates (Cutler, 1997; Rogol et al., 2002). The practical relevance of these relationships is important to consider when estimating somatic maturation, or attempting to directly measure PHV, to inform individual physical development and injury prevention interventions. If PHV is primarily driven by growth hormone, skeletal maturation by estrogen, and muscle mass by testosterone (Cutler, 1997; Rogol et al., 2002), then estimation or measurement of PHV alone may not accurately identify when players are at increased risk of apophysitis injury. However, this relationship works the other way as maturation assessed with Tanner stages may more clearly indicate testosterone activity. Therefore, creating a clear picture of pubertal development and identifying individuals at current risk, for whom training interventions may be beneficial, with isolated measures is challenging.

Even if maturation could be accurately determined with confidence, the physiological stages of pubertal maturation occur over months and years (Cutler, 1997; Rogol et al., 2002). Indeed, if players were classified as circa-PHV using the 85 to 96% of PAH range previously cited (Parr et al., 2020), then the average player in chapter four would be classified as circa-PHV from his Under-12 to Under-15 year. As the results of chapter four indicate, although the incidence of apophysitis injury was greater in these age groups, changes may not be required for all individuals through this time. Previous literature has indicated that players are at an increased risk of injury during puberty and when experiencing rapid growth (Kemper et al., 2015; Read et al., 2018b; Wik et al., 2020b). However, as players experience rapid changes during puberty, measurements are required more frequently to remain up to date with the players current status (Lampl & Johnson, 1993). Therefore, although similar patterns of injury prevalence were observed across age groups in chapter four to previous findings (Kemper et al., 2015; Read et al., 2018b; Wik et al., 2020b), growth rate alone was unable to differentiate between those who became injured and those who did not. Large scale research, or those using annual measurements of growth, may be able to observe a significant group level associations between growth rate and injury incidence, but this is not observed at a practically meaningful sensitivity to validate training interventions in individuals. When implemented in a practical setting a high number of false positive and negative observations become a reality as many circa-PHV players growing rapidly (a theoretically high risk of injury state) suffer no apophysitis injuries, while those pre-PHV of steady growth (a theoretically low risk of injury state) become injured. Furthermore, the large timeframes in which players may experience pubertal maturation changes make interventions difficult to rationalise without confident and clear supportive evidence. It may be that measurements collected over larger intervals result in spurious associations related to other hormonal changes during puberty. For example, if growth hormone regulates stature growth and estrogen controls epiphyseal growth plate ossification then it may be that the relationship or interaction between these two physiological processes affects apophysitis injury risk rather than growth rate alone. Therefore, despite being well intentioned, this makes training intervention

decisions based on growth rate and maturation estimates alone an ineffective tool to mitigate injury risk.

7.1.3 Growth and / or maturation appear to affect injury risk for some

Despite the results of chapter four indicating that frequent measurement of growth were not associated with apophysitis injury, a transient weakness in cortical bone strength has been observed during rapid pubertal growth (Wang et al., 2010). This provides a plausible mechanism supporting the influence of rapid pubertal growth on apophysitis injury incidence but only one study has shown a direct relationship between growth rate and apophysitis injury incidence using annual growth (Wik et al., 2020b). However, the baseline percentage of PAH (92.5 ± 5.6%) and age (13.3 ± 0.9 years) of participants included by Wik et al. (2020b) indicate an early-maturing group with a low average absolute stature growth of 3.4 $cm.y^{-1}$. The equivalent level of maturation in chapter four was not observed until the Under-14 age group (92.2 ± 2.6% and 14.1 ± 0.4 years) with a faster average growth rate (5.2 $cm.y^{-1}$) across similar ages. Wik et al. (2020b) also observed that rapid growth in limb length was associated with injury incidence, which, precedes rapid axial growth (Kelly & Diméglio, 2008; Malina et al., 2004a), as observed in chapter four. Therefore, the association between leg length and injury risk observed by Wik et al. (2020b) likely represents the injuries occurring in later maturing individuals. This may also explain why Wik et al. (2020b) found stature growth >0.7 $cm.m^{-1}$ (calculated from annual measurements) to be associated with increased bone and growth plate injury incidence. The incidence of apophysitis injury declines with skeletal maturation and the association between growth and apophysitis injury could just reflect the injuries which occurred in boys who were still circa-pubescent and growing, rather than resulting directly from growth. Others have observed a significant increase in all injuries during rapid growth using monthly measurements (Kemper et al., 2015). However, this was not specific to apophysitis injuries, and no association was observed between all injuries and growth rate in chapter four. Furthermore, it is unclear how rapid growth rate alone is directly causal to an increased injury incidence without an accompanying increase in any particular injury (Swain et al., 2018). Some have indicated that adolescent motor awkwardness may result in a reduced ability to avoid potentially injurious contact situations during puberty and, therefore, an increased risk of traumatic injuries (Van Der Sluis et al., 2014). However, a systematic review of research on adolescent awkwardness has that it is unclear how it could affect injury risk (Quatman-Yates et al., 2012). These findings further question the current applicability of growth rate as an injury risk identification tool in elite male youth football.

The peak incidence of apophysitis injury in chapter four occurred in the Under-12 age group, which aligns more closely with the average age of take-off (i.e. when growth rate begins accelerating) than PHV (Abbassi, 1998). Growth rate at this time was steady (0.5 ± 0.3 $cm.m^{-1}$) and similar to younger age groups. Others have similarly observed the incidence of apophysitis injuries to be greater in the Under-9 to Under-13 age groups (Hall et al., 2020), expected to be predominantly Pre-PHV. However, the majority of studies observe peak apophysitis injury incidence in the Under-13 / 14 age groups (Materne et al., 2020; Price et al., 2004; Read et al., 2018b). The peak incidence of apophysitis injury in chapter four was in the pre-PHV category ($p < 0.03$) with an average attainment of PAH of 84.4 ± 2.2%. Therefore, this result is not thought to be caused by a skew sample including

early-maturing players in the Under-12 group, but maturation was not controlled for, and it may be that misclassification influenced this result. This may further highlight the potential limitations of using somatic estimations of maturation to identify players at risk of apophysitis injury and / or challenge the recommendations that training interventions to mitigate apophysitis injury are needed only around PHV. Bone mass accrual is at its peak during adolescence, which affects life-long bone strength (Khan et al., 2000; Parfitt, 1994). The accumulation of bone can be augmented during puberty with appropriate and progressive musculoskeletal loading, particularly during Tanner stages 2 and 3 (Khan et al., 2000). Therefore, the blanket modification of training load during periods of pubertal maturation may in fact be detrimental to long-term injury resilience if bone mass accrual optimisation is impeded. This may be particularly important for *GDF5* (rs143383) T and *COL1A2* (rs412777) C allele carriers who appear to have an inherent increased risk of fracture, based on other findings in the research, and for whom augmented bone mass accrual could provide long-term injury prevention. This issue is compounded if we are unable to sensitively identify the at-risk individual and take a one-size-fits-all approach using a method of classification, which may be wrong for a given individual over a prolonged period. Therefore, modifications in training during puberty should be specifically targeted at those at-risk individuals for whom the time-lost to training and competition would be greater if injured than that lost to training modifications to result in a net increase in overall availability.

Despite its limitations, which should be acknowledged, somatic estimation of maturation using percentage attainment of PAH provide an easily applicable, time efficient and repeatable method of maturational assessment. Additionally, it is apparent that some individuals suffer substantially with apophysitis or growth-related injuries between 9 and 16 years of age, which increase around puberty alongside growth rate (Table 8). Data included in chapters four to six indicate that of those who experience apophysitis injury, around half suffered multiple occurrences and 31% of injuries were observed bilaterally, which is aligns with previous observations (Circi et al., 2017; Le Gall et al., 2006). Others have indicated that, contrary to common opinion, there may also be some long-term implications of apophysitis injury which extend beyond puberty (Guldhammer et al., 2019; Kaya et al., 2013). These observations support the potential viability of a genetic influence on apophysitis injury risk as the body repeatedly struggles to withstand the demands of elite youth football, and injury is experienced concurrently at two different sites within the body. The results of chapter four, and others, show that the majority of apophysitis injuries are severe, resulting in more than 28 days of training time loss (Le Gall et al., 2006; Light et al., 2021). Apophysitis injuries are often considered self-limiting, and some players are able to continue training and performing in matches with some restrictions (Circi et al., 2017). However, others have suggested that apophysitis injuries are not self-limiting and result in clear functional consequences on performance (Guldhammer et al., 2019). Therefore, care should be taken when allowing youth players to manage a “self-limiting” injury, which may be damaging to player wellbeing if they are forcing themselves to play through pain (Rathleff et al., 2020). Furthermore, inhibition from pain and reductions in force production, which have been observed in children with apophysitis injuries (Rathleff et al., 2020), could result in further injury. Therefore, some players suffer substantially from apophysitis injury and need protecting, while many others do not, and our current ability to sensitively identify the individual players at risk of apophysitis injury using growth rate and maturation is ineffective.

Subjective observations of the injury surveillance data indicated a general trend for players who suffered multiple apophysitis injuries to initially experience pain distally, which move proximally with age and maturation. For example, players would experience Sever's disease (apophysitis injury at the calcaneal attachment of the Achilles tendon) in the Under-11 and Under-12 age groups; then maybe Osgood Schlatter's Disease (apophysitis injury at the tibial tuberosity attachment of the patella tendon) in the Under-13 and Under-14 age groups; and then apophysitis of the hip potentially in any of several locations in the Under-15 and Under-16 age groups (Achar & Yamanaka, 2019). This was not reported in chapter four but the average age of injury incidence for Sever's disease (12.1 ± 1.3 years), Osgood Schlatter's Disease (12.8 ± 1.1 years), and hip apophysitis (14.4 ± 1.5 years) also matched this pattern, with similar findings noted by others (Arnold et al., 2017; Materne et al., 2021; Monasterio et al., 2021a). This pattern mirrors increases in growth velocity, which occurs first in the appendicular limb bones (Kelly & Diméglio, 2008), and ossification at the epiphyseal growth plates (Circi et al., 2017; Elengard et al., 2010). Therefore, one simple practical application is to ensure accurate injury history records are maintain and individuals who have previously suffered apophysitis injury are highlighted to be at greater risk of injury prior to adulthood. However, this requires longitudinal monitoring of a player and can only occur once an injury has been experienced. Therefore, genetic information could further improve the identification of individuals at increased risk of apophysitis injury in elite male youth football prior to injuries occurring.

7.1.4 Musculoskeletal loading and injury incidence

Chapter four also demonstrates similar injury patterns to those seen previously across football academy age groups as the risk of all, non-contact and non-contact muscle injuries increase with increasing age and maturation. This data also indicates jumps in the incidence of injury that coincide with phase transitional age groups (Tables 8, 19 & 28) - foundation development phase (Under-9 to Under-12) to youth development phase (Under-13 to Under-16) and professional development phase (Under-18 to Under-23). These age group phase transitions see an increase in training and competitive match demands, as exemplified by the increased acute total distances between the Under-12 to Under-13 and Under-16 to Under-18 age groups (Table 28). Others have seen similar spikes in the Under-13 and Under-18 age groups (Hall et al., 2020; Light et al., 2021; Materne et al., 2020). The Premier League Elite Player Performance Plan outlines a structured progression for the English football academy system, which offers a framework for the development of players in accordance with the individual club's philosophy (Premier League, 2012). Therefore, although external organisational structures are imposed, individual clubs have relative autonomy over the training and development of their players. The injuries per 1000 h data, reported in chapter four, indicate that the transition from Foundation Phase to the Youth Development Phase (from Under-12 into Under-13) had little influence on training injury incidence but match injury incidence increased. However, the transition to the Professional Development Phase (from Under-16 into Under-18) showed in increase in injury incidence for both training and matches, similar to previous findings (Le Gall et al., 2006; Read et al., 2018; Rumpf & Cronin, 2012). Age group phase transitions represent substantial changes to the training and competitive match demands as players experience new regulatory constraints. Specifically, as players enter the youth development phase they begin to play on larger pitches and full 11 *vs.* 11 formats, rather than 9 *vs.* 9, and when players become scholars

as Under-18s they begin to train full-time during the day and compete in 90-min matches, rather than 80-min (Read et al., 2018). Therefore, spikes in non-contact muscle injury incidence observed in the present study likely reflect the increased exposure that a new training phase and / or maturational development represents. Attention focused on managing the progressive overload and conditioning to game exposure based on growth and maturation between phases may provide beneficial reductions in injury incidence.

In addition to organisational changes, increased age and subsequent maturation results in substantial physical development. Increased size, mass, strength, power and endurance during puberty (Hansen et al., 1999; Philippaerts et al., 2006; Saavedra et al., 1991) confer players with a greater ability to exert and experience potentially injurious forces from themselves and others. Therefore, it is unsurprising that research has consistently reported the risk of all injuries, non-contact, muscle, ligament and tendon injuries to increase across academy football age groups (Hall et al., 2020; Le Gall et al., 2006; Light et al., 2021; Materne et al., 2020; Read et al., 2018). Indeed, the incidence of non-contact muscle injuries appeared to jump up in the Under-14 / 15 age groups (Tables 8, 19 & 28), which broadly align with pubertal development and the average age of PWV (Abbassi, 1998). Non-contact injuries are considered to be more preventable and attributable to loading exposure (Bowen et al., 2017; Gabbett, 2016). However, non-contact injuries can occur to all the different musculoskeletal tissues which may result from different loading exposure patterns and tissue-specific susceptibility (Kalkhoven et al., 2021; Nielsen et al., 2018). Therefore, because the current understanding of tissue-specific loading is currently limited (Nielsen et al., 2018), understanding the individual tissue-specific susceptibility to injury may allow for more targeted prevention strategies from whole-system loading.

7.2 Genetic associations with fracture risk

7.2.1 Systematic review and meta-analyses of candidate gene association studies

Conclusively identifying genetic variants which influence complex traits like injury risk is challenging and initial findings are often difficult to replicate (Ioannidis et al., 2001; Salanti et al., 2005). Therefore, the initial aim of chapter three was to complete a systematic review and meta-analysis of genetic variants associated with musculoskeletal injury risk in young healthy physically active participants. However, the genetic influence of physical performance and exercise related traits represents the integration of multiple disciplines and intersecting research areas. Consequently, developing appropriate systematic review search criteria to identify relevant articles, while limiting inclusion of irrelevant ones, was challenging. Furthermore, as the cost of conducting genetic association studies has reduced the number of genetic variants analysed has increased. This, combined with the unique nature of genetic research and the broad scope of potential injuries available for association, meant that such an all-encompassing meta-analysis would not be feasible to complete in the current research project. Furthermore, the specificity required to complete meta-analyses would mean that a single article including all candidate gene associations with musculoskeletal injury risk in young physically active participants would demand multiple meta-analyses to be completed.

Initially broad systematic searches identified clusters or papers which examined the influence of genetic variants on tissue-specific injuries. At the time of searching in 2017, twenty papers were identified that explored the genetic association with Achilles tendinopathy, fourteen examined that of ACL rupture and eleven investigated the genetic association with fracture risk. Fracture and apophysitis injuries, both affecting bone tissue, were noted to be particularly prevalent and often severe injuries in elite male youth footballers (Read et al., 2018). However, specifically identifying individuals at risk of injury was challenging and genetic predisposition was considered as a potentially viable avenue to support long-term player development with bespoke interventions. Very few studies were found to have explored the influence of genetic associations with ankle ligament or muscle injury and others had already completed meta-analysis for ACL rupture and tendinopathy injuries. Therefore, the focus of the systematic review and meta-analysis shifted to specifically focus on fracture injuries to identify genetic variants which would be valid candidates for further investigation in later research.

Meta-analyses of candidate gene association studies require additional consideration to account for genetic inheritance models and risk of bias due to the observational study designs (Lee, 2015; Sterne et al., 2016). Indeed, although the outcome variable of fracture injury incidence can be the same, combining the effect of different SNPs into one meta-analysis could misrepresent or fail to highlight the influence of individual SNPs at the overall level. Sub-group analyses within one overarching meta-analysis could resolve this issue but many studies also lack homogeneous groups, due to the need for large sample sizes in genetic research. Therefore, exploring the influence of individual SNPs using meta-analytical techniques and sub-group analysis of different study groups was considered to be more informative to understand the genetic influence on fracture risk. Indeed, the results of chapter three indicate that sex-specific genetic associations exist, despite several studies using combined analysis of both male and female participants, with a disproportional number of male to female participants. Overall, there was more than five times the number of male than female participants. However, the prevalence of fracture cases was greater in females (27%) than males (21%), which aligns with previous research indicating that physically active females are at greater risk of fracture than males (Waterman et al., 2016; Wentz et al., 2011). Therefore, the mediating effect of autosomal genetic variants may be greater, because the absolute risk of fracture appears greater, in young female than male athletes. Furthermore, autosomal variants may interact with sex chromosome genes to mediate the genetic influence on fracture injury risk. Therefore, separate analyses should also be conducted and reported when including male and female participants and / or sex accounted for in statistical analyses.

7.2.1.1 Meta-analysis study heterogeneity, quality assessment and VDR (rs2228570)

The *VDR* (rs2228570) SNP was found to be amongst the most frequently replicated SNPs with fracture risk in chapter three, included in three studies (Chatzipapas et al., 2009; Korvala et al., 2010; Varley et al., 2018). However, as heterogeneity between studies was significant and high in the allele contrast model ($p = 0.006$, $I^2 = 76\%$) a random effects meta-analysis model was selected. This heterogeneity may be attributable to unknown ethnicity dependent variations of included participants as ethnicity was not fully reported (Chatzipapas et al., 2009; Korvala et al., 2010; Varley

et al., 2018). A random effects meta-analysis model redistributes the contribution of each study to the overall effect across all included studies to be more balanced than that of a fixed effects model (Sterne et al., 2011). This has the potential to increase the relative contribution of smaller low-quality studies on the overall effect (Sterne et al., 2011). Therefore, although a fixed effects model was not considered appropriate for the *VDR* (rs2228570) analysis, a significant trivial increase of fracture risk would have been evident with the C allele using a fixed effects model (OR = 1.37, 95% CI = 1.03 – 1.81, p = 0.03, d = 0.07). However, when the results of chapter five are included into the *VDR* (rs2228570) fracture risk meta-analysis calculations of chapter three, the results remain significantly heterogeneous (p = 0.007, I^2 = 72%) and the fixed effects meta-analysis model non-significant (OR = 1.27, 95% CI = 0.98 – 1.64, p = 0.07, d = 0.13) for the allele contrast model. Nevertheless, the influence of unknown ethnicity dependent variation is particularly important when considering the *VDR* gene as a paradoxical relationship is observed between ethnicity, vitamin D concentration and fracture risk (Owens et al., 2018). Black and Hispanic men have been shown to have an increased risk of vitamin D deficiency but a reduced risk of osteoporosis, rapid bone loss, and fractures when compared to Caucasians (Engelman et al., 2008; Hannan et al., 2008).

As proximal portions of the DNA are frequently inherited together, the frequency of alleles can vary within the population and significantly differ between ethnicities (Gibson, 2016; Lewontin, 1964). Numerous genetic variants can be present in a single gene and if different alleles demonstrate non-random inheritance patterns, they are considered to be in linkage disequilibrium (Lewontin, 1964). Genetic variants which show linkage disequilibrium can create common groups of alleles, known as haplotypes, which are often inherited together (International HapMap Consortium, 2005). Consequently, reporting the ethnicity of included participants is important to understand the potential influence that unknown ethnicity dependent causal variants may have on the genetic influence with injury occurrence (Pruna et al., 2015). If the inheritance pattern is strong, then candidate variants may still be used as a proxy to estimate the genetic association with injury risk in that group, but this should be acknowledged. Indeed, a recent study in elite male footballers found significant differences in the genetic association of candidate variants with tissue-specific injuries between different ethnicity groups (Pruna et al., 2015).

No overall effect was observed from the pooled results of candidate gene association studies in chapter three ($p > 0.06$). However, discussion of findings indicated the *VDR* (rs2228570), *COL1A2* (rs412777), *COL1A1* (rs1800012), and *GDF5* (rs143383) SNPs as potential candidates for further exploration. It was hypothesised that further research on the *VDR* (rs2228570) SNP could reduce the heterogeneity between studies and indicate the existence of any ethnicity dependent relationships with fracture risk if clearly reported. The influence of the *COL1A2* (rs412777) SNP with fracture risk was only examined in two studies but both independently observed strong significant, but directly contradictory results in children 4 to 16 years of age (Blades et al., 2010; Suuriniemi et al., 2003). These results cancelled each other when combined in the meta-analyses of chapter three but further investigation in the present research project, and future research, appeared warranted. The results of chapter five supported those of (Suuriniemi et al., 2003) who observed the C allele of the *COL1A2* (rs412777) SNP to be significantly associated with increased risk of fracture. However, Blades et al. (2010) was the only study to achieve the highest study quality assessment in chapter three and when the results of chapter five are included in the *COL1A2* (rs412777) meta-analysis of

chapter three, no significant overall effect is observed (p > 0.17). Furthermore, each study used a different participant group, Blades et al. (2010) reported a combined analysis, while Suuriniemi et al (2003) used only females, and chapter five only males, so no sex-specific sub-group analysis could be conducted.

The systematic review and meta-analysis qualitative assessment also highlighted several areas for improvement in future research using the quality of genetic association studies (Q-Genie) assessment tool (Sohani et al., 2016). The Q-Genie tool is valuable as one of few assessment methods specifically designed for genetic association studies, which require particular considerations to ensure study quality is achieved (Gibson, 2016). However, there is limited explanation of how to determine the score of each evaluation criteria. While this allows for efficient evaluation of numerous studies, discrepancies between authors were apparent due to room for individual interpretation in the scoring system based on the assessment criteria. This was particularly evident for the sample size and power criteria, which had two questions: 1) “was the sample size appropriate?”, and 2) “was an *a priori* power analysis conducted?”, to determine a score of 1 one to seven (Sohani et al., 2016). Consequently, the quality assessment result for several of the studies evaluated with the Q-Genie tool in the systematic review and meta-analysis of chapter three were largely determined by this question. The research team felt that for consistency and objectivity of the evaluation process to be maintained, with such little evaluation criteria, only marks of one / two or six / seven were appropriate based on this criteria. Nevertheless, the quality assessment process indicated a distinct absence in the clear reporting of participant ethnicity.

7.2.2 Genetic penetrance

7.2.2.1 Growth and maturation and COL1A2 (rs412777)

Interestingly, both Suuriniemi et al (2003) and Blades et al. (2010) observed significant differences in BMD that supported their respective findings, which are specific to pre- or circa-pubertal children. Suuriniemi et al (2003) observed that Tanner stage 1 & 2 (equivalent to pre-PHV in the present study) female *COL1A2* (rs412777) CC individuals had a near five times greater relative risk of fracture than AA individuals (RR = 4.9 [95% CI = 1.4 - 17.4], p = 0.015). Overall, no significant difference were found in anthropometric measurements, physical activity, or bone mass between females with fractures and non-injured controls. However, lumbar spine and distal radius BMD was significantly lower in the fractured individuals (Suuriniemi et al., 2003). On the other hand, Blades et al (2010) found that *COL1A2* (rs412777) CC individuals in Tanner stages 1-3 (equivalent to pre- and circa-PHV in the current study) demonstrated reduced risk of fracture (OR = 0.38 [95% CI = 0.19 – 0.79], p = 0.01) and increased lumbar spine BMD. In comparison the results of chapter five found that elite male footballers including those at all stages of maturation with the *COL1A2* (rs412777) CC genotype had a significantly greater risk of fracture injury and A allele carriers were more than half as likely to sustain a fracture injury (RR = 0.45 [99% CI: 0.22-0.94], p = 0.005). However, when combining the *COL1A2* (rs412777), *GDF5* (rs143383) and *VDR* (rs2228570) SNPs into a TGS for fracture risk in chapter six, no significant association was observed with maturation (p = 0.16). Nevertheless, if the pre- and circa-PHV categories were combined the risk of fracture would likely have been significantly influenced by maturation as the frequency of fracture injury risk appeared

higher for pre-PHV (Injury risk = 0.04% [95% CI = 0.01 - 0.24]) and circa-PHV (Injury risk = 0.04% [95% CI = 0.01 - 0.22]) than the post-PHV groups (Injury risk = 0.01% [95% CI = 0.00 - 0.05]) as may be expected (Cooper et al., 2004; Parfitt, 1994; Tinkle & Wenstrup, 2005).

7.2.2.2 Sex and COL1A1 (rs1800012)

Blades et al (2010) also observed that the *COL1A1* (rs1800012) T allele was associated with three times greater odds of fracture in prepubertal children (OR = 3.1 [95% CI = 1.43 – 6.61], p = 0.004), although this was not accompanied with any bone strength associations. The meta-analysis in chapter three identified a novel, sex-specific, significant but trivial reduction in fracture risk associated with the *COL1A1* (rs1800012) T allele in young physically active females (OR = 0.48 [95% CI = 0.25 – 0.91], p = 0.03, d = *-0.18*). These findings highlighted the substantial influence that that both age and sex can have of the genetic penetrance of variants associated with injury risk. Indeed, a study including 603 participants, which aimed to determine if genetic variants linked with adult BMD were associated with that of children and adolescents, concluded that the direction and magnitude of associations often only became evident when accounting for sex and maturation (Mitchell et al., 2015). The *COL1A1* (rs1800012) T allele was also associated with apophysitis injury in chapter five (p = 0.03) and seven (p = 0.03) and has been associated with increased bone injury in prepubescent children of both sexes (Blades et al., 2010) and repeatedly in postmenopausal women (Jin et al., 2009; Mann & Ralston, 2003). Indeed, in another study by Suuriniemi et al. (2006) also found that the *COL1A1* (rs1800012) T allele was associated with reduced bone strength properties in pubertal girls (Suuriniemi et al., 2006). Nevertheless, the same *COL1A1* (rs1800012) T allele is also associated with reduced risk of ACL rupture in young healthy adult populations (Ficek et al., 2013; Khoschnau et al., 2008). Furthermore, no significant association was observed between *COL1A1* (rs1800012) and fracture risk in chapter five. Additionally, no difference in the overall effect was observed when adding the results of chapter five to the *COL1A1* (rs1800012) meta-analysis calculations reported in chapter three. Therefore, similar to previous findings, no significant influence was detected between *COL1A1* (rs1800012) and fracture risk in young physically active males (Cosman et al., 2013; Korvala et al., 2010; Varley et al., 2018).

7.2.2.3 Ethnicity and GDF5 (rs143383)

Polygenic phenotypes like injury susceptibility are complex with the variability in observable traits influenced, but not solely determined, by the interaction of numerous genetic variants (Wackerhage, 2014). Therefore, injury incidence can often be an emergent event and any single genetic variant may be influential to injury susceptibility but also mitigated by other heritable factors in some individuals. The T allele of the *COL1A1* (rs1800012) SNP is rarer than the other minor alleles included in the present research, observed in approximately 9% of the overall population (The 1000 Genomes Project Consortium, 2015). This appears to vary by ethnicity and the T allele of the *COL1A1* (rs1800012) occurs in 6% of those with African, and 19% European, ancestry based on the superpopulation categories of the 1000Genomes project, which averages at around 2% in the overall population (The 1000 Genomes Project Consortium, 2015). Homozygotes of the T allele are even more infrequent, found in only 1% of African and 4% European ancestry individuals, which averages

at around 2% in the overall population (The 1000 Genomes Project Consortium, 2015). At the start of the research *a priori* sample size calculations using G*Power software (Faul et al., 2007) suggested that 128 participants would be needed to achieve 80% statistical power using a two-tailed test of difference with medium effect size (Cohen's D = 0.5), significance set at 0.05, and minor allele frequency equal to 0.45. Therefore, we sought to recruit approximately 150 elite male footballers from the Fulham Football Club academy and first team squads between the ages of 8-40 years. One-hundred-and-twelve participants were included in the genetic association study of chapter five of which ninety-five were also included in the TGS, loading and maturation or loading analysis of chapter six. Therefore, although the results of the statistical analysis may be considered to be underpowered in the present study a significant genetic association between the infrequent *COL1A1* (rs1800012) T allele and apophysitis injury risk was still observed with a significance of p = 0.03. Consequently, although this result should be considered with healthy scepticism it is likely that there is a genetic association between *COL1A1* (rs1800012) and apophysitis injury. This finding is supported by observations of other studies, which identify an influence on BMD during childhood and adolescence (Blades et al., 2010; Suuriniemi et al., 2006), in combination with the mechanistic observations of apophysitis injury incidence resulting from a transient weakness in bone strength during rapid growth (Wang et al., 2010).

Only one study was found to have investigated the influence of the *GDF5* (rs143383) SNP with fracture risk in young male military recruits (Zhao et al., 2016). Zhao et al. (2016) observed a significant increased risk of fracture injury associated with the A allele, which has been previously associated with fracture risk in the elderly (Liu et al., 2013; Valdes et al., 2011). However, no meta-analysis was able to be completed for this SNP as no replication had been attempted in young physically active participants. The results of chapter five found that the *GDF5* (rs143383) SNP was associated with fracture injury using the recessive (p = 0.0311), dominant (p = 0.034) and codominant (p = 0.0451) models. Furthermore, when the results of chapter five are combined with those of Zhao et al. (2016), a significant small-to-moderate overall effect was found, indicating a reduced risk of fracture injury associated with the G allele (p < 0.0004, *d* = -0.3 to -0.6). This was evident for the allele contrast (OR = 0.49 [95% CI = 0.38 – 0.63], *p* < 0.0001, *d* = *-0.39*), recessive (OR = 0.53 [95% CI = 0.39 – 0.71], *p* < 0.0001, *d* = *-0.35*) and homozygote (OR = 0.33 [95% CI = 0.18 – 0.61], *p* = 0.0004, *d* = *-0.61*) comparison models with minimal heterogeneity between studies (I^2 = 0%, *p* > 0.48).

The ethnicity of the Chinese participants included by Zhao et al. (2016) was not reported. However, the 1000Genomes East Asian super population includes the Dai Chinese in Xishuangbanna, Han Chinese in Beijing China and Han Chinese South China 1000Genome ethnicity groups (The 1000 Genomes Project Consortium, 2015). The majority of participants in the present research would be included in European (57%) or African (26%) 1000Genomes superpopulation ethnicity categories (The 1000 Genomes Project Consortium, 2015). The frequency of the G allele is much greater in African ethnicity superpopulation (97%) than East Asian (29%) and European (37%) groups. Therefore, it is worth acknowledging the potential influence that unknown ethnicity dependent variants could have on the observed protective effect of the *GDF5* (rs143383) G allele. However, the population ethnicity of Zhao et al. (2016) is likely not as diverse as that of chapter five and both studies showed similar findings. Therefore, it is considered appropriate to conclude that the

GDF5 (rs143383) G allele appears protective for fracture risk injury in young healthy physically active males.

7.2.3 COL1A1 (rs1800012), COL1A2 (rs412777), VDR (rs2228570) and GDF5 (rs143383) SNPs

Overall, the findings of chapters three, five and six progressed the research project and contribute new knowledge on the genetic influence of fracture and apophysitis injury in several ways: Firstly, the *COL1A1* (rs1800012), *COL1A2* (rs412777), *VDR* (rs2228570) and *GDF5* (rs143383) SNPs were identified as viable candidates for continued investigation with fracture risk due to observations from previous genetic association studies in young healthy physically active participants. Secondly, the importance of genetic penetrance in understanding the influence of candidate genes on musculoskeletal injuries, particularly bone injuries, was discussed. Thirdly, and relating to genetic penetrance, the importance of clear reporting on participant ethnicity was highlighted to define and delineate the influence of candidate genes with fracture risk. Finally, a novel genetic association was observed between the *COL1A1* (rs1800012) T allele and increased apophysitis injury risk, which is supported by other mechanistic evidence. In addition to, indicating a protective effect of the *GDF5* (rs143383) G allele on fracture risk from the pooled results of young healthy physically active males.

7.3 Genetic associations with non-contact, non-contact muscle, and tendon injury risk in elite male football

To identify genetic variants capable of improving the sensitivity of tissue-specific injury susceptibility models, chapter five examined previously identified genetic associations, with common injuries in elite male football development. Several variants showed significant associations with injury incidence with $p < 0.01$, which were considered with greater confidence than those significant at $p < 0.05$ due to the potential influence of multiple independent tests conducted (Tables 20-23). The *COL1A2* (rs412777) SNP was associated with fracture risk using both the recessive ($p = 0.005$) and codominant ($p = 0.01$) models. The *GDF5* (rs143383) SNP was also associated with fracture injury using the recessive ($p = 0.03$), dominant ($p = 0.03$) and codominant ($p = 0.045$) models. The *COL1A1* (rs1800012) SNP was associated with apophysitis injury using the dominant model ($p = 0.03$). The *COL5A1* (rs12722) overdominant ($p = 0.01$), dominant ($p = 0.03$) and codominant (p = 0.03) models were all associated with non-contact muscle injury, with the recessive model associated with tendon injury ($p = 0.02$). The *MMP3* (rs679620) recessive ($p = 0.006$), codominant ($p = 0.006$) and overdominant ($p = 0.006$) models, along with *VDR* (rs2228570) dominant ($p = 0.04$), codominant ($p = 0.03$) and overdominant models ($p = 0.01$) were significantly associated with the incidence of all non-contact injuries. This subsequently appeared to influence the overall risk of time loss injury for the *VDR* (rs2228570) dominant ($p = 0.03$) and overdominant ($p = 0.02$) models.

7.3.1 VDR (rs2228570) and non-contact injury

The genetic associations observed between the *COL1A2* (rs412777), *GDF5* (rs143383) and *COL1A1* (rs1800012) SNPs with fracture and apophysitis risk have been discussed earlier in this chapter. These SNPs were all identified, along with *VDR* (rs2228570), as potential candidates for further investigation with fracture risk as part of the conclusions of chapter three. The *VDR* (rs2228570) SNP was the only of these not to be subsequently associated with a bone tissue injury in later chapters. However, the GA genotype of *VDR* (rs2228570) SNP was associated with greater non-contact injury than homozygotes in chapter five. The investigation of candidate variants in the present study was hypothesis-driven and based on the findings of previous genetic association studies. However, at the time of analysis for chapter five, only one other study appeared to have explored the influence of the *VDR* (rs2228570) SNP on injury in football (Massidda et al., 2015b). Massidda et al. (2015b) found no significant association of muscle injury incidence or severity between the *VDR* (rs2228570) genotypes using a codominant model only. Nevertheless, the incidence of muscle injury was higher for heterozygotes but no overdominant comparison was conducted to examine the presence of significant differences with homozygotes (Massidda et al., 2015b). Others have observed increased risk of lower back pain for G allele homozygote athletes (Cauci et al., 2017), which has been consistently observed as an ethnicity dependent risk factor for intervertebral disc degeneration (Pekala et al., 2019). Interestingly, the G allele appears to reduce the risk of intervertebral disc degeneration in Caucasians, but increase risk in Hispanic populations (Pekala et al., 2019). This supports the hypothesis that ethnicity dependent variability may have contributed to the substantial heterogeneity observed in the meta-analyses of the VDR (rs2228570) SNP with fracture risk in chapter three.

Male adolescent footballers who are *VDR* (rs2228570) heterozygotes have, however, also been associated with increased BMD and insulin-like growth factor 1 (IGF-1) (Diogenes et al., 2010). IGF-1 is a principal factor in mediating the anabolic and linear growth effect of growth hormone (Rogol et al., 2002) and, others have shown a significantly greater increase in BMD following resistance training for *VDR* (rs2228570) heterozygotes (Rabon-Stith et al., 2005). However, these findings do not explain how the risk of non-contact injury may be greater for (rs2228570) heterozygotes as was observed in chapter five. A recent study exploring the response of cells to vitamin D in basal and inflamed states found that GG cells promoted synthesis of matrix proteins, while downregulating extracellular matrix catabolism with the opposite effect observed in GA cells (Colombini et al., 2021). This supports the conclusions of Baumert et al. (2022) who observed faster range of motion recovery following exercise induced muscle damage. Indeed, vitamin D plays numerous functions around the body from muscle repair and function to immunity and bone homeostasis, via activation of VDR proteins on target cells (Dahlquist et al., 2015; Owens et al., 2018). The activation of vitamin D is significantly greater for the G than A allele of the VDR (rs2228570) SNP (Arai et al., 1997) and it is unclear why the G allele is associated with increased spinal injury (Pabalan et al., 2017). Therefore, VDR (rs2228570) heterozygotes may be at greater susceptibility to non-contact injury in elite male football due to an impaired recoverability associated with the A allele, while also suffering from increased susceptibility resulting from unknown factors which contribute to the increased risk of back pain associated the G allele.

7.3.2 MMP3 (rs679620) and COL5A1 (rs12722) with non-contact muscle and tendon injury

Baumert et al. (2022) also observed a significant influence of the *MMP3* (rs679620) and *COL5A1* (rs12722) genotypes on recovery from exercise induced muscle damage. The TT genotype of the *MMP3* (rs679620) SNP was associated with improved strength recovery while *COL5A1* (rs12722) T allele homozygotes showed increased range of motion prior to, and muscle soreness 48 hours after, exercise induce muscle damage (Baumert et al., 2022). Conversely, in the present study possession of the *MMP3* (rs679620) CT or TT genotypes was associated with a significantly increased frequency of non-contact injury per season compared with CC homozygotes (RR = 1.40 [99% CI: 1.03-1.92], p = 0.006). Although these results may appear contradictory to those of Baumert et al. (2022), others have also observed a reduced risk of hamstring injury associated with the *MMP3* (rs679620) C allele in elite male footballers (Larruskain et al., 2018). Furthermore, Hall et al. (2022) observed that the *MMP3* (rs679620) T allele was associated with greater time lost from knee injury in football players. The MMP3 protein activates other metalloproteinases (Toth et al., 2003), which regulate the extra cellular matrix by catalytically degrading structural proteins including various collagens (Birkedal-Hansen et al., 1993; Somerville et al., 2003). The *MMP3* (rs679620) SNP is a missense coding variant resulting in a glutamic acid in place of a lysine codon in the amino acid sequence of MMP3. The functional consequences of this change are unclear, however, associations with injury are beginning to become evident. It may be that the improved strength recovery with the T allele observed by Baumert et al. (2022) is misaligned with the recovery of other structural tissues following exercise. This would potentially explain how the C allele could be protective for non-contact injury (chapter five) and hamstring injury (Larruskain et al., 2018) despite showing slower recovery of strength if less force is able to be exerted on weaker surrounding structures. Indeed, the potential for impaired recovery of structural tissues is supported by the findings of Hall et al. (2022). Nevertheless, this is speculative as improved strength recovery would normally be thought to be protective from non-contact injury.

Chapter five also observed that *COL5A1* (rs12722) homozygotes (TT and CC) experienced significantly greater frequency of non-contact muscle injuries per season than heterozygotes (RR = 1.49 [99% CI 0.98-2.26], p = 0.01). This finding was also apparent using the codominant and dominant models which indicated a reduced risk of non-contact muscle injury for T allele carriers compared with the CC genotype (RR = 0.71 [99% CI 0.48-1.06], p = 0.03). Again, this differs to what may be expected based on the results of Baumert et al. (2022) who observed that *COL5A1* (rs12722) TT homozygotes had greater range of motion at baseline but increased muscle soreness 48 hours after exercise induce muscle damage, indicating impaired recovery. The *COL5A1* (rs12722) SNP is a noncoding SNP within the 3' UTR is thought to produce a more stable mRNA transcript (Laguette et al., 2011). This suggests that the T allele results in more COL5A1 protein, which may result in more type V collagen. The diameter of type I collagen fibres appears to decrease with an increased abundance of type V collagen (Birk et al., 1990) and the COL5A1 (rs12722) C allele has been repeatedly associated with a reduced risk of tendon and ligament injury (Pabalan et al., 2018). Indeed, the risk of tendon injury appeared substantially reduced for those with the *COL5A1* (rs12722) TC or CC genotypes compared with TT individuals (RR = 0.41 [99% CI 0.15-1.10], p = 0.02) in

chapter five, similar to previous findings (Altinisik et al., 2015; Mokone et al., 2006; September et al., 2009), although this was not observed for ligament injuries.

Despite the relationship between type V collagen abundance and type I collagen fibre diameter (Birk et al., 1990) a direct causal mechanism for the protective effect of the C allele has yet to be established in vivo. One study investigating potential mechanisms found that *COL5A1* (rs12722) CC individuals had more extensible tendons than T allele carriers (Kubo et al., 2013). Others were unable to replicate this finding (Foster et al., 2014) but the C allele has been repeatedly associated with increased flexibility with differences between genotypes increasing with age (Brown et al., 2011b; Collins et al., 2009). These findings differ to the observations of Baumert et al. (2022) who observed that *COL5A1* (rs12722) TT homozygotes had greater knee flexion and extension range of motion, although others have also observed the CT genotype to be associated with reduced mobility in adolescent team sport athletes compared with TT individuals (Stastny et al., 2019). Furthermore, others have observed that the *COL5A1* (rs12722) T allele was associated with improved endurance performance (Brown et al., 2011a; Posthumus et al., 2011b). Running economy was originally hypothesised as the mechanism by which the *COL5A1* (rs12722) T allele may enhance endurance performance following the previous associations with joint range of motion (Brown et al., 2011b; Collins & Posthumus, 2011). However, no significant differences were observed between genotypes in the volume or elasticity of tendons (Foster et al., 2014) or running economy (Bertuzzi et al., 2014). Nevertheless, type V also collagen plays a crucial role in the regulation of fibrilogenesis in non-cartilaginous tissue (Collins & Posthumus, 2011) and it may be that the *COL5A1* (rs12722) SNP influences endurance performance via an alternative mechanism.

Mice models indicate that reduced *col5a1* mRNA production - associated with the C allele of the human *COL5A1* (rs12722) SNP - decreases the compliance and tensile strength of the aorta (Wenstrup et al., 2006). Higher compliance of the aorta is associated with endurance training and greater stroke volume (Tarumi et al., 2021), which may provide an alternative mechanism for enhanced endurance performance, despite the reduced diameter of type I fibres, resulting from increased type V collagen abundance (Birk et al., 1990), associated with the T allele. Indeed, Pickering and Kiely (2017a) suggest that the previously observed associations between CC homozygotes and reduced muscle cramping during endurance running (O'Connell et al., 2013) may result from a reduced muscle exertion / fatigue inducing capability, rather than being directly protective. Therefore, although associations between the *COL5A1* (rs12722) T allele and endurance performance remain equivocal, heterozygotes may have an advantage over homozygotes due to a balance between improved endurance performance (attributable to the T allele) and tissue loading capacity (attributable to the C allele), which are both considered protective against non-contact muscle injury (Gabbett, 2016; Kalkhoven, 2021).

Alternatively, the protective effect of the TC genotype may result from a balanced trade-off between the increased non-contact muscle injury susceptibility observed in the present study for C homozygotes and increased tendon injury incidence observed for T homozygotes. Mechanical stress at the musculotendinous junction in *COL5A1* (rs12722) T and C allele homozygotes may subsequently be more susceptible to injury than TC heterozygotes, which could be expressed by increased observation of non-contact muscle injury. Nevertheless, the influence of the *COL5A1* (rs12722) SNP on flexibility and endurance performance remain unclear despite the T allele

consistently increasing soreness from exercise induced muscle damage (Baumert et al., 2016, 2022) and risk of tendon and ligament injury (Pabalan et al., 2018). Therefore, it remains unclear why the incidence of non-contact muscle injury was lower for CT heterozygotes in chapter five and the findings of Baumert et al. (2016, 2022) indicate that the CC genotype would be expected to be protective for non-contact muscle injury. Interestingly, the *MMP3* (rs679620) C allele and *COL5A1* (rs12722) T allele have both been independently associated with increased tendinopathy risk (Altinisik et al., 2015; Gibbon et al., 2016), and to interact to further increase tendinopathy risk (Raleigh et al., 2009). Indeed, the *COL5A1* (rs12722) T allele was also associated with an increased risk of tendinopathy in chapter five. However, when the *MMP3* (rs679620) and *COL5A1* (rs12722) SNPs were combined to create a TGS for tendon injury no significant association was found when accounting for loading exposure and maturation in the model (Table 26).

7.3.3 COL1A1 (rs1800012) and apophysitis injury

Despite the replication of several previously observed genetic associations in chapter five the only model to be significantly associated with injury in chapter six for the developed TGS included only one genetic variant (*COL1A1* [rs1800012]). Considering the nature of the observations with injury incidence observed in chapter four, occurring in predominantly non-contact injuries, the addition of loading exposure and maturation was thought to hypothesised to improve the identification of at-risk individuals. Non-contact injuries are considered to be more preventable and attributable to loading exposure than contact injuries (Bowen et al., 2017; McCall et al., 2018). Additionally, maturation had been shown to influence the risk of non-contact injuries in chapter four, with others indicating that growth rate may be a significant factor for injury risk (Kemper et al., 2015; Wang et al., 2010). Therefore, the interaction between loading exposure, growth and / or maturation status with the *MMP3* (rs679620), *VDR* (rs2228570), *COL5A1* (rs12722) and *COL1A1* (rs1800012) SNPs, which had been associated with non-contact injuries in chapter five, were hypothesised to improve the sensitivity to identify elite male youth footballers at increased risk of injury which could guide the practical application of individualised training programmes to reduce injury risk and support long term development.

Chapter six sort to explore the potential interaction that tissue-specific TGSs, calculated based on previous findings and those of chapter five, may have when considering maturation status or growth rate along with loading exposure categories. However, the apophysitis injury model ($p < 0.04$) was the only significant tissue-specific injury risk model observed ($p > 0.16$). This suggested that possession of the *COL1A1* (rs1800012) T allele (TT & TG individuals) showed a four-fold increase in apophysitis injury risk ($p = 0.03$) compared to GG individuals and that growth rate above 0.6 $cm.m^{-1}$ also increased the risk of injury more than three-and-a-half times that of slower growth rates ($p = 0.02$). This was not observed in chapter four using high growth rate categorical alone despite approaching significance ($p = 0.06$). Indeed, an interaction effect indicated that apophysitis injury risk was significantly different between combinations of growth (>0.6 $cm.m^{-1}$ or not) and *COL1A1* (rs1800012) genotype (TT+TG *vs.* GG) groups. This indicated that T allele carriers experiencing rapid growth (>0.6 $cm.m^{-1}$) were almost twelve times more likely to suffer an apophysitis injury than GG homozygotes at low growth ($p = 0.006$). Rapid growth was still associated with an

almost six times increased risk of apophysitis injury compared with low growth considering GG individuals only (p = 0.01). However, when both experiencing low growth T allele carriers were six-times more likely to suffer apophysitis injury than G allele homozygotes (p = 0.02). These findings indicate that the *COL1A1* (rs1800012) genotype may be an important risk factor for apophysitis injury in elite male youth footballers. Furthermore, this may allow more sensitive identification of those at risk of apophysitis injury generally and especially during rapid growth.

This is the first study to identify a genetic association between the *COL1A1* (rs1800012) SNP and apophysitis injury. However, the results of previous genetic association studies indicate a valid mechanism that may explain the observed increase in injury risk associated with the observed with the T allele. Chapter three identified that the *COL1A1* (rs1800012) SNP appears to be particularly influenced by both sex and aging in relation to bone fracture injuries. Apophysitis injuries are considered to be insidious non-contact injuries resulting from repeated microfractures at the apophysis of bone, which is unable to withstand shear forces exerted by the tendon during exercise (Arnold et al., 2017; Gholve et al., 2007; Holden & Rathleff, 2020; Ogden & Southwick, 1976). A transient period of increased bone weakness has been shown to occurs as BMD declines prior to PHV, which then rebounds during puberty (Faulkner et al., 2006; Wang et al., 2010; Wang & Seeman, 2009) when bone mass accrual is at its peak (Yilmaz et al., 2005). Indeed, the results of chapter four indicated the incidence of apophysitis injury actually peaked in the pre-PHV group. The *COL1A1* (rs1800012) SNP T allele has been previously associated with reduced bone formation relative to resorption in during early puberty in females (Suuriniemi et al., 2006). In addition to, an increased risk of fracture and reduced BMD, which was greatest during periods of limb bone growth during adolescence (Blades et al., 2010). Therefore, the T allele may increase apophysitis injury risk by further impairing bone strength during puberty, which may already be reduced due to pubertal hormones and rapid growth. The genetic influence on apophysitis injuries may also be supported by observations that between 15 to 25 % of players who suffer apophysitis injuries also experience them bilaterally (Le Gall et al., 2006) indicating a whole-body system weakness.

Contrary to previous suggestions of long-term athletic development models, the results of chapter six suggest that the risk of apophysitis injury was not significantly influenced by loading exposure. This finding should be treated with caution and not cited in isolation as a reason not to manage player training load during puberty as the loading exposure measure was determined based on variability of normal training exposure and not complete rest. Nevertheless, it could be that genetic susceptibility and growth is sufficient to predispose youth footballers to apophysitis injury regardless of exposure to football activity. Interestingly, the influence of the *COL1A1* (rs1800012) T allele may be significantly affected by dietary calcium intake (Brown et al., 2001). Brown et al. (2001) found that T allele carriers with low calcium intake lost significantly more bone than GG homozygotes; however, T allele carriers of high calcium intake actually gained significantly more bone than GG individuals (Brown et al., 2001). Bone collagen synthesis stimulated by nutritional stimulation and can occur rapidly (Babraj et al., 2005). Therefore, increased dietary calcium intake may protect *COL1A1* (rs1800012) T allele carriers against apophysitis injury, which could be implemented as a preventative application in elite male youth football development to support long-term player development. Furthermore, this may mean that players do not need to miss training and musculoskeletal loading, which could be beneficial for long-term bone health.

The overall risk of injury for each player, within each session, is low and actively managed via applied interventions and training design. Even when considering all non-contact injuries, which represent the broadest injury classification analysis investigated, the overall risk of injury per player per session was less than 1% (Table 26). While small variations in this baseline risk can exponentially increase when considered upon the squad and long-term development scale, this resulted in an overall low number of injuries in which to explore the genetic influence of injury susceptibility. This was especially evident for tendon injuries with only 27 injuries recorded from 38,378 potentially injurious sessions, despite including data from 95 participants over more than five years, equating to a tendon injury frequency of 0.5 ± 1.3 injuries per 1000 player hours (Table 28). Therefore, although significant genetic associations could be observed when considering only the frequency of injury over the duration of the injury surveillance period (Table 22), when attempting to take this to the level of applicability with injury based on loading exposure over each session no significant influence is observed. It may be argued that chapter five was underpowered to identify significant genetic associations because of the rarity of certain injuries. This may be evident from a scientific research perspective; however, in attempting to scrutinise the current applicability of genetic data to inform the daily practice of elite male football development, this may also highlight the current inability to purposefully implement these data. It may be that other limitations within the study design restricted the ability of the models to sensitively identify the circumstances in which different TGS for tissue-specific injury mediate injury susceptibility. Nevertheless, despite promising preliminary findings, the current application of genetic information might only be able to infer a generic predisposition to injury. This may still be valuable information for the applied practitioner to steer injury prevention training and nutritional strategies but delimiting this inference is important. That withstanding, the rarity of injuries makes the genetic association observed between the *COL1A1* (rs1800012) T allele and apophysitis injury worth further exploration and future validation attempts should be considered .

7.4 Potential applications to support development

7.4.1 COL1A1 (rs1800012), calcium intake and apophysitis injury

The results of chapters three to six indicate potential avenues of further investigation for applications of genetic information to support long-term footballer development. The *COL1A1* (rs1800012) SNP was repeatedly implicated with injury risk; however, genetic associations appear to be significantly influenced by sex and growth rate (chapters three, five and seven). These findings improve our understanding of how *COL1A1* (rs1800012) affects injury risk by highlighting the need to account for genetic penetrance, sex, growth rate, and potentially maturation and / or aging when applying individual genetic information. If the genetic association between *COL1A1* (rs1800012) and apophysitis injury can be replicated, then targeted nutritional and / or training interventions could be viable strategies for genetically individualised protection from injury. Brown et al. (2001) found that the *COL1A1* (rs1800012) SNP influenced BMD and a significant interaction was observed depending on calcium intake. Possession of the *COL1A1* (rs1800012) T allele has been linked with increased bone turnover (Keen et al., 1999), which may explain why high calcium intake was associated with reduced, and low calcium intake increased, bone loss for T allele carriers compared to GG homozygotes (Brown et al., 2001). Although these findings were observed in elderly participants

(mean age, 69 years), the findings of Brown et al. (2001) suggest that high calcium intake could mediate and even reverse the increased susceptibility of *COL1A1* (rs1800012) T allele carriers to apophysitis injury during periods of rapid growth. Bone mass accrual peaks during puberty, which appears to influence life-long bone strength (Parfitt, 1994). Therefore, if effective, this intervention would become particularly powerful for *COL1A1* (rs1800012) T allele carriers to reduce injury risk both at that time, and potentially long-term, if bone strength can be improved with increased bone mass accrual during puberty with continued exposure to appropriate musculoskeletal loading (Parfitt, 1994; Torres-Costoso et al., 2020). This, in turn, could be of particular importance for *GDF5* (rs143383) A and *COL1A2* (rs412777) C allele carriers, who appear to have an increased risk of fracture injury, as shown in chapter five, and by others (Suuriniemi et al., 2003; Zhao et al., 2016).

Interestingly, young physically active *COL1A1* (rs1800012) T allele carriers have also been shown to experience fewer ACL ruptures in adulthood (Ficek et al., 2013; Khoschnau et al., 2008; Posthumus et al., 2009a). Therefore, although this was not observed in the present study, as the number of ACL ruptures was so low, senior TT and TG genotype footballers may be more resilient to severe injury overall, especially if supported with nutritional interventions earlier in their development. The protective effect of the *COL1A1* (rs1800012) T allele for ACL rupture is unclear (Posthumus et al., 2009a). Previous work on fracture risk in the elderly suggested that the T allele may reduce bone tissue strength because of increased type 1 collagen pro-collagen fibres formed exclusively of COL1A1 (instead of 2xCOL1A1 and 1xCOL1A2), which may be weaker (Mann et al., 2001). However, the divergent influence of the *COL1A1* (rs1800012) T allele on injury risk depending on sex, age and growth suggest this is not the case, otherwise all would be expected to be at greater risk of injury. Keen et al. (1999) suggest that their results indicate that the *COL1A1* (rs1800012) SNP may have a regulatory effect on total body type I collagen turnover. Indeed, *COL1A1* (rs1800012) is located in a transcriptional regulatory region, which could therefore affect collagen synthesis (Keen et al., 1999). Consequently, *COL1A1* (rs1800012) could influence the regulation of type I collagen, which could interact with other factors affecting type I collagen regulation to explain how the influence of the T allele varies by injury, sex, age and growth rate. Therefore, understanding how the *COL1A1* (rs1800012) SNP interacts with other physiological and environmental factors to predispose individuals to injury risk could provide greater insights for applied interventions to reduce injury risk.

7.4.2 Growth, maturation, COL1A1 (rs1800012), loading exposure and apophysitis injury

The results of chapter three highlight the current inability to sensitively identify at-risk individuals during pubertal development, using current growth and maturation guidelines, in practically meaningful timeframes in elite male footballers. Nevertheless, generalisable findings consistently show increased apophysitis injury incidence around pubertal growth periods (Light et al., 2021; Materne et al., 2020; Read et al., 2018). As apophysitis injuries are generally considered insidious loading exposure-based injuries, this has led many to suggest that training load should be reduced during periods of pubertal development (Arnold et al., 2017; DiFiori et al., 2014; Faigenbaum et al., 2009). However, appropriate musculoskeletal loading during puberty can augment bone strength and support long-term injury resilience (Parfitt, 1994; Torres-Costoso et al., 2020).

Therefore, despite the best intentions, applying generalisable findings to all individuals may be detrimental to some. Furthermore, the results of chapter six indicate that loading exposure had no significant influence on apophysitis injury risk when considered with *COL1A1* (rs1800012) genotype and growth rate category in the current group of elite male footballers. This suggests that there was no difference in apophysitis injury risk when players experienced normal training variability, substantial de-load or overload in total distance, despite significant variations observed for growth rate category and *COL1A1* (rs1800012) genotype. This supports the argument that blanket strategies to modify training load during pubertal growth may not be appropriate for all individuals and that even reduced exposure to football activity during periods of susceptibility can result in injury occurrence.

This issue is exacerbated when considering the limitations of somatic maturation estimates, which are frequently used to categorise players in intervention groups in football and youth development models (Parr et al., 2020; Ryan et al., 2018). This should not devalue somatic estimates of maturation, but maturation estimates alone are not sufficient to individualise a training programme. Indeed, because of the prolonged periods during which adolescence may experience pubertal maturation (Abbassi, 1998; Parr et al., 2020), these alterations are arguably not truly individualised. Nevertheless, some individuals do appear to suffer repeated, prolonged, and painful apophysitis injuries (Light et al., 2021; Materne et al., 2020). These individuals deserve additional support and should be protected, but not at the expense of others' developmental opportunities. Therefore, a tension exists whereby the needs of individuals diverge substantially during pubertal development, yet practitioners lack effective tools to identify and differentiate these individuals. If genetic information can be used to improve the sensitivity of identification of the at-risk individuals during pubertal development, then all players may benefit from more appropriate training interventions and greater individualisation to support-long term development. In this way, genetically informed interventions, such as calcium supplementation discussed previously, could support the long-term development opportunities of both the at-risk and not-at-risk individuals with greater specificity and individualisation of a player's programme. The results of this research indicate that the *COL1A1* (rs1800012) SNP is a viable candidate for further exploration towards this goal. Nevertheless, further replication of the *COL1A1* (rs1800012) T allele association with apophysitis injury risk is required to assert this effect with confidence. However, if replication is observed then future research could explore how dietary calcium intake affects apophysitis injury risk between *COL1A1* (rs1800012) genotypes in elite male youth footballers.

7.4.3 Apophysitis injury record

The contributory influence of genetic variation on apophysitis injury is supported by observations that up to 25% of players who suffer apophysitis injuries experience them bilaterally (Le Gall et al., 2006) and many experienced multiple occurrences in the present research. However, this finding was not explicitly analysed and only noted from the injury surveillance data. Therefore, even if the genetic association between apophysitis injury and the *COL1A1* (rs1800012) SNP is not replicated in future research, practitioners may look to examine the risk of re-injury for those who have previously experienced apophysitis injury. Genetic testing arguably remains a "nice to have" rather than a vital piece of information at this time. Therefore, the associated cost of genotyping

players may not be accessible for all. Nevertheless, those involved in long-term footballer development may benefit from ensuring accurate records of apophysitis injury history are maintained as it represents a low-cost strategy, which may highlight individuals at greater risk of apophysitis injury. Therefore, future apophysitis injury risk models or research, inclusive of genetics or not, may benefit from exploring the influence of previous apophysitis injury on subsequent injury risk further.

7.4.4 Genetic associations with exercise induced muscle damage and fracture

The *COL1A2* (rs412777), *MMP3* (rs679620), *VDR* (rs2228570), *COL5A1* (rs12722) and *GDF5* (rs143383) SNPs were each associated with the interindividual variability of injury risk in elite male youth football players. Consistent with previous findings, the *COL1A2* (rs412777) and *GDF5* (rs143383) variants appeared to significantly influence the risk of fracture injury. Additionally, the *MMP3* (rs679620) and *VDR* (rs2228570) SNPs significantly influenced the incidence of all non-contact injuries, with *COL5A1* (rs12722) specifically related to non-contact muscle injuries. However, when these SNPs were combined with other variants previously associated with injury, tissue-specific TGSs showed no significant influence on injury incidence, when also accounting for load exposure and maturation in chapter six. Therefore, a greater understanding, and inclusion of more genetic factors which influence tissue-specific injury is required to differentiate the heritable difference in individual susceptibility. Consequently, future research should continue to try and identify novel genetic variants which contribute to the observed heritable variability in tissue-specific injury risk. Once tissue-specific injury susceptibility models can be established, more bespoke protective interventions could be developed. Including environmental and physiological variables into risk models such as this would also provide further insight into applied interventions, which may vary depending on age, maturation, growth, loading exposure and recovery.

7.4.4.1 Exercise-induced muscle damage and loading exposure

In chapter five, *MMP3* (rs679620) T allele carriers and *VDR* (rs2228570) AG genotype individuals were associated with a significantly greater risk of non-contact injury than *MMP3* (rs679620) CC, and *VDR* (rs2228570) AA and GG individuals. Additionally, *COL5A1* (rs12722) TC heterozygotes appeared to have a significantly reduced risk of non-contact muscle injury specifically. Genetic associations with injury typically assume a linear, additive model of injury risk with an increasing number of risk alleles possessed, considered to increase injury susceptibility (Massidda et al., 2015b; Miyamoto-Mikami et al., 2019). Therefore, overdominant (XY *vs.* XX + YY) genetic association models are often not analysed, in favour of codominant (XX *vs.* XY *vs.* YY), dominant (XX + XY *vs.* YY) and recessive (XX *vs.* XY + YY) models (Massidda et al., 2015b; Miyamoto-Mikami et al., 2019). This strategy may be chosen to reduce the number of tests conducted if heterozygotes are not expected to display significant variations from the combined homozygote groups. Indeed, interpretations from observations that *COL5A1* (rs12722) and *VDR* (rs2228570) heterozygotes are associated with significant differences in injury risk from chapter five remain speculative. However, as our understanding of the complex interaction of genetic and environmental factors remains limited, and inheritance patterns unknown, evaluation of different genetic comparison models in warranted

(Lee, 2015). Indeed, it is possible that heterozygotes display unique injury risk profiles when compared to homozygotes if a balance, or dilution, of protective effects, exerted by each allele, are expressed.

The potential for heterozygotes to possess significantly increased, or decreased, musculoskeletal injury risk may be exemplified by previous observations of the *COL5A1* (rs12722) SNP. The *COL5A1* (rs12722) T allele has been associated with both enhanced endurance performance (Brown et al., 2011a; Posthumus et al., 2011b), and increased tendon (Altinisik et al., 2015; Mokone et al., 2006) and ligament injury (O'Connell et al., 2015; Posthumus et al., 2009b). Type V collagen plays an important role in the regulation of fibrilogenesis in non-cartilaginous tissue (Collins & Posthumus, 2011). Therefore, heterozygotes may be protected because of a trade-off between reduced fatiguability and increased tissue loading capacity, both of which are considered protective against non-contact muscle injury (Gabbett, 2016; Kalkhoven, 2021). Indeed, the results of chapter five suggest that *COL5A1* (rs12722) C allele carriers are less than half as likely to suffer a tendon injury than T homozygotes and the musculotendinous junction is particularly susceptible to injury (Jakobsen & Krogsgaard, 2021), which may also explain why heterozygotes are at reduced risk of non-contact muscle injury. Nevertheless, no significant association has previously been observed between the *COL5A1* (rs12722) and muscle injuries in football (Larruskain et al., 2018), using an over-dominant model, nor mixed athlete groups (Miyamoto-Mikami et al., 2019). However, Larruskain et al. (2018) included only non-contact hamstring muscle injuries. Therefore, future research may look to explore the influence of the *COL5A1* (rs12722) SNP on all lower limb non-contact muscle injuries using an additive model to further explore the influence of this variant.

The increased susceptibility to non-contact injury observed for VDR (rs2228570) heterozygotes in chapter five is harder to explain. Vitamin D receptors play an important role in transcriptional regulation, which mediates the effects of vitamin D in the body (Wang et al., 2005). Vitamin D has been shown to influence muscle repair, muscle function, immunity, cardiac function and bone homeostasis (Dahlquist et al., 2015; Owens et al., 2018). Therefore, VDR (rs2228570) GA heterozygotes may be at greater risk of non-contact injury due to the loss of some protective or compensatory effect expressed in homozygotes. Furthermore, although a similar study in elite male footballers found no significant difference in muscle injury incidence between the three VDR (rs2228570) genotypes, only the codominant genetic model was assessed (Massidda et al., 2015b). Nevertheless, injury incidence was higher in heterozygotes than both homozygotes; however, as no overdominant model was used, this cannot be considered statistically significant (Massidda et al., 2015b). Therefore, future research should look to examine if the VDR (rs2228570) SNP is associated with non-contact injury using more genetic association comparison models.

A recent study by Baumert et al. (2022) identified a TGS with seven SNPs associated with exercise induced muscle damage and recovery, including the *COL5A1* (rs12722), *MMP3* (rs679620) and *VDR* (rs2228570) SNPs. Participants were segregated into three groups based on how many preferential (considered protective against muscle damage) alleles they possessed, with the *COL5A1* (rs12722) C allele, *MMP3* (rs679620) TT genotype, and *VDR* (rs2228570) G allele considered preferential (Baumert et al., 2022). They found that the preferential TGS group was stronger, with reduced muscle soreness, and greater range of motion than the non-preferential group following exercise-induced muscle damage (Baumert et al., 2022). These findings support the

potential influence of the *COL5A1* (rs12722), *MMP3* (rs679620) and *VDR* (rs2228570) SNPs to differentiate individuals at greater risk of non-contact injury. Nevertheless, when the *COL5A1* (rs12722), *MMP3* (rs679620) and *VDR* (rs2228570) SNPs were included in the non-contact and non-contact muscle injury risk TGS (Table 25) of chapter six, no significant association was observed with injury risk when accounting for maturation and loading exposure. This may be because the genotype risk score allocation in the present study was incorrect - *COL5A1* (rs12722) and *VDR* (rs2228570) heterozygotes were given a risk score of zero (low) and two (high), respectively, for non-contact muscle and non-contact injury. However, these values were chosen based on the findings of chapter five and previous research. Alternatively, the complexity of tissue-specific injury susceptibility in elite male footballer was unable to be sensitively differentiated, with only three SNPs and more genetic variants required to understand the heritable differences in injury susceptibility.

Non-contact injuries are considered to be more preventable and attributable to training and competitive loading exposure football (Bowen et al., 2017, 2020). Exercise-induced muscle damage will likely contribute to the susceptibility of non-contact muscle injury incidence in football and inclusion of the *COL2A1* (rs2070739), *IGF2-AS* (rs4244808), *TRIM63* (rs2275950), and *TTN-AS1* (rs3731749) variants, used by Baumert et al. (2022), into the TGS could improve the model. Muscle damage is proportional to musculoskeletal loading (Simmons et al., 2021) and repeated exposure to muscle damage with insufficient recovery can lead to injury (Bowen et al., 2017, 2020; Kalkhoven et al., 2021). Therefore, if a TGS can differentiate players susceptibility to exercise-induced muscle damage then more targeted recovery or training intervention strategies could be implemented to support players with the aim of reducing risk of injury. While a greater understanding of the genetic factors which contribute to tissue-specific injury risk is still needed, integration of both genetic and environmental / exposure factors will provide greater actionable applications of genetic information to support long-term individual player development. Substantial work exists on strategies for improved recovery (Brink et al., 2010; Halson, 2014) and appropriate progression and exposure to training and matches in football (Bowen et al., 2017, 2020). Furthermore, continued monitoring and screening processes of the observable strengths and weaknesses of individual players is still recommended (Buckthorpe et al., 2019; Hughes et al., 2017, 2018). However, differentiating players based on their individual biological predisposition could, in turn, further refine this personalised process of support. Therefore, sophisticated integration of personal, environmental and genetic information into a complex understanding of the individual could enhance our ability to support player development. Training load, growth and maturation represent some of the most notable factors for consideration but integration of numerous factors to create a better understanding of a player is possible and likely valuable.

The variation in training load was not associated with significant differences in injury incidence in chapter six. However, this was not expected and different to previous findings in elite male football (Bowen et al., 2017, 2020). This may be because loading exposure classification in chapter six was based upon the z-score of an individual's TD_{ACDiff} and HID_{ACDiff} in each training session relative to the average team TD_{ACDiff} and HID_{ACDiff} within each season. Loading exposure was then classified as substantial overload with a z-score of ≥1, substantial de-load when ≤-1 and within normal variability when between -1 and 1. This was expected to represent a meaningful change in training exposure, which would increase the risk of injury based on previous research in football (Bowen et al., 2017,

2020). However, this also represents a novel attempt to account for the numerous issues which have been discussed around the use of acute-to-chronic workload ratios (Impellizzeri et al., 2019, 2020; Lolli et al., 2019). Specifically, the absolute acute-to-chronic load difference was used to account for the issue of mathematical coupling and the use of ratio data (Impellizzeri et al., 2020), but arguably also provides a more practically interpretable value of training exposure. For example, the absolute acute-to-chronic difference in high-intensity distance for an under-18 player was 244 ± 815 m (Table 28). Although this would vary each season based on the cohort and training, a +1SD increase in the absolute acute-to-chronic high-intensity distance was generally over 1000 m. This is arguably easier to understand and provides a clear difference in exposure that may be expected to exert a meaningful influence on the ability of the individual to withstand that exposure given their recent history. For example, instead of a 1.5 increase in acute-to-chronic workload ratio of high-intensity distance, which could be 20 m or 200 m, an absolute increase in acute-to-chronic workload of 20 m or 200 m can be interpreted relative to match or training demands. Nevertheless, it may be that even +1SD was too low to accurately reflect training load exposures, which would significantly increase injury risk. Alternatively, injury incidence was not frequent enough in the current sample to detect a meaningful change at exposures of higher magnitudes. However, the likelihood of injury would be expected to increase the greater the increase in acute to chronic workload (Bowen et al., 2017, 2020). Nevertheless, this also may reflect the potential lack of predictive power to detect injury with loading exposure alone.

A further noteworthy consideration is that load monitoring only accounted for whole system physical exposure experienced by the players. Although this is common practise and has been able to identify significant relationships with injury in football before (Bowen et al., 2017, 2020), this does not accurately account for tissue-specific loading (Kalkhoven et al., 2021). Considerable challenges exist in understanding how training load measures translate into tissue-level mechanical exposure in an applied context (Kalkhoven et al., 2021). Nevertheless, the forces experienced at different locations, and tissues, around the musculoskeletal system from elite football training and competition are likely to vary substantially to influence tissue-specific damage and injury risk (Kalkhoven et al., 2021). Furthermore, the loading exposure measures included total exposures of distance covered, which do not include other sport-specific actions (e.g., velocity of ball, number of ball kicks, changes of direction, jumping, landing, tackling, etc.) that put a strain on the musculoskeletal system. Therefore, a greater understanding of the mechanical loads experienced at the tissue-specific level could further clarify the influence of tissue-specific genetic risk on injury in along with growth and maturation factors.

7.4.4.2 Fracture risk

The influence of the *COL1A2* (rs412777) and *GDF5* (rs143383) SNPs on fracture risk could guide nutritional monitoring and interventions towards those at the greatest need. Vitamin D supplementation (Anderson et al., 2017) and increased strength (Clark et al., 2011) have been shown to provide a protective effect against fracture injury. Therefore, targeted interventions could be suggested for individuals at increased risk of fracture and the *GDF5* (rs143383) T and *COL1A2* (rs412777) C alleles appear to confer a greater risk of fracture. As mentioned, it may also be

important to ensure musculoskeletal loading during pubertal growth is not unnecessarily removed for those who are not at risk of injury. This is because appropriate and progressive loading of the musculoskeletal system during puberty can stimulate increased bone mass accrual, and strength, for life-long protection from fracture (Torres-Costoso et al., 2020). This may be especially important for *GDF5* (rs143383) T and *COL1A2* (rs412777) C allele carriers who are at increased risk of fracture injury but may be offloaded from football during puberty to reduce apophysitis injury risk. Alternatively, teaching players effective strategies to avoid contact and / or to fall safely, when possible, to absorb potentially injurious forces may also be viable interventions targeted towards fracture risk allele carriers. Although limited research has explored the influence of effective falling strategies on fracture risk, many fractures in the upper limbs occur in football due to a fall on an outstretched hand (Hedström et al., 2010). Therefore, it is reasonable to assume that learning to roll and dissipate forces from a fall would reduce fracture incidence (Wormhoudt et al., 2017).

7.4.5 Ethical implications

The use of genetic information to inform physical programming can be divisive (Pickering & Kiely, 2019; Wackerhage et al., 2009; Webborn et al., 2015). Potentially the most controversial use of genetic information is for talent identification (Varley et al., 2017) and there is general consensus against this in youth athletes (Williams et al., 2016). Furthermore, the complex nature of football performance (Reilly et al., 2000), combined with our limited knowledge of the genetic determinants of physical performance (Gibson, 2016), mean that genetic information will have limited utility for talent identification in football above non-genetic physical, technical, and perceptual-cognitive performance testing. Arguably a more promising and cost-effective use of genetic information is to support the individualisation of elite player development to provide all with a greater opportunity to achieve their potential (Pickering & Kiely, 2019). However, this still presents several ethical issues worthy of consideration. Genetic information does not change with age and not all genetic associations are known at the time of testing. Therefore, players could be tested for a genetic variant associated with musculoskeletal injury, which is later shown to influence more severe diseases (Vlahovich et al., 2016; Williams et al., 2016). This knowledge could be distressing for players, and close relatives due to the heritability of genetics, who may or may not wish to know this information (Williams et al., 2016). Therefore, it is suggested that mandatory genetic counselling occurs, prior to testing, to explain this possibility and agree how it should be managed, alongside consultation when results are reported to contextualise the findings (Williams et al., 2016). Further consideration is required on who should be able to request genetic tests (players, parents, coaches, club doctors), however, freely given informed consent for testing from the player is essential without discrimination or risk of penalty imposed on the players decision (Vlahovich et al., 2016; Williams et al., 2016). Furthermore, the privacy of genetic information requires specific consideration (Vlahovich et al., 2016). Elite athletes often sign medical confidentiality waivers to allow coaching staff to understand how medical conditions may impact on training and match participation (Vlahovich et al., 2016). However, confidentiality clauses are also often included into these agreements for particularly sensitive medical information and clear agreement, supported by written informed consent, is required on how and to whom genetic information will be stored and accessible (Vlahovich et al., 2016). Indeed, if football clubs decide to offer genetic testing to their players, then specific

agreements should be established or these results should be kept confidential (Vlahovich et al., 2016). Therefore, although not insurmountable, the use of genetic information in football poses numerous ethical and legal considerations that require additional support from experts who can interpret the results, which has further implications for who clubs need to employ or partner with.

7.5 Reflections on the programme of research presented within the thesis

The aims of the thesis sought to explore how inter-individual genetic variants, physiological development and loading exposure might inform applied interventions to reduce injury risk in elite male youth football. Implementation of novel interventions in elite sport can be challenging as deviation from tried and tested practises may be resisted (Burden et al., 2022). This resistance to change may be easier to overcome by presenting the benefits of a novel intervention with confidence based on supporting scientific evidence of its effectiveness. However, a thorough understanding of the contextual challenges in which a novel intervention will be implemented will also likely improve its success (Burden et al., 2022). The aims of the thesis were explored from a research-practitioner perspective and sought to draw upon both the scientific and contextual information to develop impactful applied innovations (Burden et al., 2022). In this way, the research attempted to discover if, and how, genetic information may be utilised to genuinely improve applied practise in elite male youth football. In addition to, maintaining critical scepticism on the value of any interventions, which may result in discovery of no applied interventions that can be utilised with confidence using our current knowledge.

Considering both the research-practitioner perspective and the aims of the project, this thesis attempts to synthesise the findings of the research and discuss its potential, with further investigation, to support long-term player development with greater understanding in the future. Overall, despite some promising observations the conclusions of the research are not considered to suggest a consistently strong and clearly identifiable effect to assert confidence in directing an intervention in an elite male youth football context at this time. The observed influence of the *COL1A1* (rs1800012) SNP with apophysitis injury during periods of rapid growth in elite male youth footballers appears to be the most promising initial finding of the present research. However, despite previous evidence indicating a plausible mechanism for the variability in apophysitis injury risk associated with *COL1A1* (rs1800012), this remains a novel finding, which requires replication in another independent group of elite male youth footballers to confirm the result and be confident in its effect.

7.6 Limitations of research presented within the thesis

The results of this thesis contribute to the existent literature by progressing the understanding of genetic factors associated with injury in elite male football when considering the influence of physiological and environmental factors. However, these findings should be considered alongside the context and limitations of the research. The specific limitations of each study are discussed in

each chapter (sections 3.5, 4.4, 5.4, and 6.4). Therefore, this section will focus on reflecting on the overall programme of research, alongside some considerations and recommendations for researchers to bear in mind when building on the research presented in this thesis.

The present programme of research utilised a convenience sample of elite male youth footballers at one club, which limited the number of players available for participation in the studies. At the start of the research, *a priori* sample size calculations using G*Power software (Faul et al., 2007) suggested that 128 participants would be needed to achieve 80% statistical power using a two-tailed test of difference with medium effect size (Cohen's D = 0.5), significance set at 0.05, and minor allele frequency equal to 0.45. Therefore, we sought to recruit approximately 150 of the elite male footballers from the Fulham Football Club academy and first team squads between the ages of 8-40 years. In total, 112 players, and parents of those aged below 18 years, were included in the genetic association study of chapter five from the convenience sample. Of these, 95 were also included in the TGS, loading and maturation or growth analysis of chapter six. Therefore, the results of the statistical analyses in chapters five and six may be considered to be underpowered. Nevertheless, significant genetic associations were still observed, even for the infrequent *COL1A1* (rs1800012) T allele and apophysitis injuries. Consequently, although the results should be considered with some scepticism, it is likely that the genetic associations observed in chapter five do influence injury risk. However, the lack of power may explain the null findings observed for tissue-specific TGS and injury incidence in chapter six.

Convenience sampling of players from one club will reduce the generalisability of the project findings. However, from the research-practitioner perspective this strategy provided multiple benefits to support innovation development including: deeper understanding of the applied context; improved longitudinal monitoring of players; increased ability to scrutinise historical records; and reduced variability that may be expected to result from inclusion of players from multiple teams with different testing, training and monitoring philosophies. Nevertheless, these factors also: limit the ability of the findings to transfer to different contexts, even within elite male youth football development; reduce the potential to differentiate the influence of genetic variants depending on variable such as sex and ethnicity; and lower the power to observe significant effects, which may be valid but small.

One advantage of embedding the research within one football club was the access to longitudinal data. This allowed the long-term influence of genetic, environmental and physiological factors on injury risk to be examined. The growth, maturation, injury and loading data analysed throughout the thesis was retrospectively captured from club databases. Although this data is purposefully collected prospectively and used to determine workload exposure and physical development to inform daily applied decision making, using retrospective data in research has its disadvantages (Sedgwick, 2014). Causation can be inferred from an association between a risk factor and outcome in observational studies in specific circumstances (Hill, 1965). However, the retrospective observational nature of the studies included in this thesis mean that a direct causal effect is unable to be established. Nevertheless, the retrospective observational data of this thesis is less likely to be affected by selection bias as all players at the club were offered participation in the study regardless of injury history (Sedgwick, 2014). Furthermore, practitioners assessing injury, collecting loading and measuring growth and maturation were blinded to the genetic risk factors of each participant when the data were collected, and genotyping services completed by individuals

blinded to the environmental and physiological factors. Recall bias is also worth considering with retrospective data and the main investigator did collate all information together towards the latter stages of the project. However, genetic information was the last data to be included in chapter five after all the injury history data had been added. Although caution is required when interpreting the results of retrospective, observational genetic association studies to infer causality, practically meaningful outcomes can still be highlighted. Causality could provide greater insight into mitigating factors and applied strategies to reduce injury risk but in reality if an association between a genetic variant and tissue-specific injury risk is clear and consistent enough then bespoke and targeted applied interventions can still be designed to support the long-term development of individual players. Acknowledging the differences between association and causality is crucial but the application of the observations may not be so different.

It is not possible to control for all of the factors which could influence injury incidence in elite male youth football. Indeed, the observational nature of the research also meant that the club was continually and actively seeking strategies to reduce injury incidence, which changed over time. The ethical and commercial implications of attempting to control different factors is also questionable if a new strategy to reduce injury incidence was to become available during the experimental period. Nevertheless, although not all factors can be controlled for, many can be accounted for. Indeed, understanding how nutritional factors differed between injured and non-injured, or genotype groups, could have allowed further inferences on the applicability of genetic factors on injury incidence to be made. Additionally, only workload exposure and injuries experienced as a registered player, during Fulham Football Club activities were included in the analyses. Nevertheless, a large portion of the participants also played in school sports teams, during which time they could become injured as a result of fatigue accumulated from the physical demands of involvement within an elite football academy. Vice-versa players could suffer an injury during club activities that was due to school sport exposure. Although this issue predominantly affects only the under-9 to under-16 players it is believed to be noteworthy of consideration, especially when interpreting the results of chapter six.

7.7 Conclusions

Untangling the complexity of genetic predisposition is challenging and the current understanding of the genetic factors that influence injury risk remain limited (Gibson, 2016). Nevertheless, it is clear that players are biologically unique, presenting with different growth, maturation and aging profiles, as well as experiencing variable loading exposures (Hautala et al., 2006; Hubal et al., 2005; Pickering & Kiely, 2019). This raises the need for an individualised approach in sport and exercise; however, this is not a novel concept and individuality is frequently discussed (Abt & Lovell, 2009; Halson, 2014). Indeed, heritable factors appear to contribute substantially to the variability in physical performance and injury risk between individuals (Andrew et al., 2004; Bouchard et al., 1998; De Moor et al., 2007; Magnusson et al., 2020). Therefore, the interindividual variability of genetic predisposition should be acknowledged and understanding this could allow for more targeted and bespoke training interventions to be designed to support the long-term development of every individual (Pickering & Kiely, 2019). However, an individual's genetic predisposition interacts with other complex and dynamically changing biological processes to influence player adaptation,

performance development and injury risk (Pickering & Kiely, 2019). Therefore, the practical applicability of genetic information could be enhanced if considered in conjunction with other physiological and environmental factors (Pickering & Kiely, 2019). The concept of genetic penetrance may be of particular value, as the advantageous or deleterious effects of a particular genetic factor may only become expressed under specific circumstances. In an elite male youth football programme, these physiological and environmental factors could be actively managed or mitigated with training, or at least accounted for in daily practice. Therefore, understanding how each individual's inherent biological predisposition may interact with physiological and environmental factors could allow coaches to be better informed on interventions to support player development (Pickering & Kiely, 2019).

The findings of the present thesis contribute to the wider literature in several ways. The most exciting observation with potentially meaningful practical application is the novel observation of an association between the *COL1A1* (rs1800012) T allele and apophysitis injury risk in rapidly growing elite male adolescent football players. The results of chapter six indicate that those with the *COL1A1* (rs1800012) TT or TG genotype had a four-fold greater risk of apophysitis than G allele homozygotes ($p = 0.03$). Growth rate >0.6 $cm.m^{-1}$ was also associated with a more than three times greater risk of apophysitis injury in chapter six, compared with those at lower growth ($p = 0.02$). However, no significant association ($p > 0.06$) was observed between apophysitis injury and growth rate, even >1.0 $cm.m^{-1}$, in chapter four when not accounting for *COL1A1* (rs1800012) genotype or loading exposure in the model. Despite others observing a significant association with rapid growth and bone and growth plate injuries >0.7 $cm.m^{-1}$ (Wik et al., 2020b). When combined, T allele carriers experiencing rapid growth were almost twelve times more likely to suffer an apophysitis injury than GG homozygotes at low growth ($p = 0.006$). Current long-term athletic development models indicate that training load should be reduced during pubertal growth to reduce the risk of apophysitis injury (Arnold et al., 2017; DiFiori et al., 2014; Faigenbaum et al., 2009; Ryan et al., 2018). However, this period of physiological development also represents the peak in life-time bone mass accrual which influences life-time bone strength (Parfitt, 1994; Khan et al., 2000). Furthermore, bone mass accrual during this time can be augmented with appropriate musculoskeletal loading (Parfitt, 1994; Torres-Costoso et al., 2020). Therefore, blanket training modifications during puberty may not be necessary for all and potentially detrimental to long-term fracture resilience / bone strength.

The results of chapter four also challenge the practical applicability of using measures of stature growth and maturation alone to identify players at increased risk of injury in a practically meaningful measurement interval (5-7 weeks). Increased age and maturation were associated with increases in all injuries, non-contact injuries and non-contact muscle injuries with spikes observed around academy football phase transition age groups (under-13 and under-18). This provides high-level implications for consideration around the progression and management of training and competition exposures around these transitions. However, maturational phases occur over prolonged periods of time (Parr et al., 2020) and training modifications might not be necessary throughout this time. This issue is compounded by limitations in the current methods of maturation estimation, which are frequently used to guide training modifications during adolescence in elite male youth football (Malina et al., 2005b; Parr et al., 2020; Ryan et al., 2018). Stature growth is frequently cited as a risk factor for injury (DiFiori et al., 2014; Faigenbaum et al., 2009; Kemper et al., 2015; Wik

et al., 2020b). However, when stature measurements are collected more frequently than annual measurements no association was observed in elite male youth footballers when considering stature growth alone. This challenges the current common narrative in the literature that stature growth is a clear risk factor for injury, and it is possible that other confounding variables associated with pubertal maturation are more attributable to the changes in injury incidence. Therefore, the results of this thesis suggest that more complex and sensitive models are required to confidently identify players at risk of injury in an elite youth football environment.

Injuries which occur around important developmental milestones can have greater implications than just the time lost to injury itself (Jones et al., 2019; Larruskain et al., 2021). Academy players are typically given two-year registrations, after which their participation in the programme is evaluated to ensure players with the greatest credible chance of becoming professional are retained, and those who are not are supported to leave the programme. Therefore, the Under-12, Under-14 and Under-16 age groups represent major milestones in the progressive journey of elite male youth football players development, as their future within the elite player performance pathway is evaluated. Nevertheless, these age groups also roughly align with periods of maturational development which may temporarily disrupt their performance and / or increase injury susceptibility (Read et al., 2018b; Wang et al., 2010). Players who become injured around these times (or throughout these periods) can lose significant developmental time, or temporarily appear fragile and unable to withstand the demands of elite football participation. These players may become lost to the academy development system, despite being potentially very talented. Indeed, sustaining an injury influences the development of elite male youth football players (Jones et al., 2019; Larruskain et al., 2021). Sport-specific skills appear to be independent of maturation status, despite clear physical advantages (Coelho-E-Silva et al., 2010; Malina et al., 2007b; Matthys et al., 2012). Adult physical performance is predominantly determined by genetics and training (Davids & Baker, 2007). Therefore, injury during periods of maturation around registration deadlines could mask the identity of genuine future potential as players experience temporary disruption due to injury, which may not affect future potential but disrupts performance development.

Although the association between *COL1A1* (rs1800012) and apophysitis injury risk requires further validation in an independent sample of elite male youth footballers, if this observation can be replicated the *COL1A1* (rs1800012) SNP could be used as a marker to identify players in need of additional attention and bespoke interventions to reduce apophysitis injury during adolescence in elite male youth football. Previous research has shown the influence of the *COL1A1* (rs1800012) genotype on bone mass to be significantly influenced by calcium intake (Keen et al., 1999). Therefore, future research should examine how variations in nutrition and diet influence the effect of *COL1A1* (rs1800012) genotype on apophysitis injury more closely. If calcium intake deficiency is observed through screening processes, then this may become a viable candidate to explore as a genetically informed intervention strategy. This individualisation would hopefully support both the players in need of additional management (those at risk of apophysitis injury) and those who can continue to be exposed to full training and competition (those at low risk of apophysitis injury). Indeed, although unexpected and requiring further examination, loading exposure was not associated with a significant difference in apophysitis injury when also accounting for *COL1A1* (rs1800012) genotype and growth rate in chapter six. As with current individual screening processes, genetic predisposition

would likely only represent one step in the decision-making processes in applied practice, which should incorporate multi-factorial considerations for that individual. Nevertheless, genetic information could provide greater understanding of an individual, which can be acted upon with greater evidence of an effect at an individual level. Further replication of the *COL1A1* (rs1800012) T allele association with apophysitis injury risk is required to assert confidence of this effect. Nevertheless, this observation represents a promising avenue of further exploration and a novel contribution of the thesis.

The *COL1A1* (rs1800012) SNP has been associated with the risk of other musculoskeletal injuries (Ficek et al., 2013; Gibbon et al., 2020; Khoschnau et al., 2008). However, the influence of the *COL1A1* (rs1800012) SNP appears to differ significantly based on sex and age as shown in chapters three and six. The findings of chapter three highlight how sex and age should be considered when interpreting the potential influence of the *COL1A1* (rs1800012) T allele with fracture risk. Sex-specific analysis indicated a protective effect of the *COL1A1* (rs1800012) T allele in females despite previous associations with increased risk of osteoporotic fracture in the elderly. This suggests that the genetic penetrance of the T allele is influenced by sex / age and is not ubiquitously detrimental to bone strength as has been suggested previously (Mann et al., 2001). The *COL1A1* (rs1800012) T allele was thought to influence fracture injury risk as a result of an increased proportion in type I pro-collagen formed exclusively from three COL1A1 sub-units, instead of the more abundant COL1A1, COL1A2 combination form (Mann et al., 2001). Type I collagen formed from an increased proportion of three COL1A1 pro-collagen was hypothesised to result in weaker type I collagen tissue (Mann et al., 2001). However, the findings of this thesis indicate that the influence of the *COL1A1* (rs1800012) T allele is significantly influenced by physiological and environmental factors and others have shown a protective effect for ACL injury. The T allele appears to be particularly protective against injury during childhood and in the elderly (Blades et al., 2010; Mann et al., 2001), during which notable changes in BMD occur, as bone turnover is influenced (Mann et al., 2001; Parfitt, 1994). As the *COL1A1* (rs1800012) T allele is both observed to be associated with protection from ligament injury, and increased risk of bone fracture and apophysitis injury it is arguably unlikely that the mechanism of influence results from a direct change in tissue structural integrity. However, the high sensitivity to calcium intake during puberty associated with the *COL1A1* (rs1800012) SNP may suggest that this variant influences the dynamic regulation of type I collagen in some way, which may explain how the T allele can be both protective and deleterious at different ages in different tissue when physiological processes may interact with this to mediate injury susceptibility. This observation is completely hypothetical but the observations of chapters three, five and six in the present thesis challenge the previous hypothesis that the *COL1A1* (rs1800012) T allele is ubiquitously determinantal to tissue strength (Mann et al., 2001).

The results of chapter five provide further evidence to support previously observed genetic associations between *COL1A2* (rs412777), *MMP3* (rs679620), *VDR* (rs2228570), *COL5A1* (rs12722), *GDF5* (rs143383) and *COL1A1* (rs1800012) with interindividual variability in injury risk. Specifically, the *COL1A2* (rs412777) and *GDF5* (rs143383) SNPs appear to play a significant role in the risk of fracture injury. In the initial findings of chapter three, no significant overall effect was observed for fracture risk in physical active participants for any SNP - including *COL1A2* (rs412777) and *GDF5* (rs143383) - using any genetic association comparison model. Only two studies were

included for *COL1A2* (rs412777), which showed strong significant and directly contradictory results. Therefore, when the results of chapter five were included in the meta-analysis of chapter three no overall effect was observed. Nevertheless, the *COL1A2* (rs412777) SNP appears to be associated with fracture risk in some way but more consistent replication attempts are required to confirm this association and its direction. However, when the results of chapter five were included with those of chapter three, a significant overall effect was observed for all genetic comparison models which suggested that the *GDF5* (rs143383) T allele increased the risk of fracture injury in young, healthy physically active males. Only one other study has explored this observation and future research should look to validate these findings (Zhao et al., 2016). Nevertheless, the *COL1A2* (rs412777) and *GDF5* (rs143383) SNPs appear to be associated with fracture risk in elite male youth footballers in the present research. Understanding the genotype of these SNPs in elite male youth footballers could inform targeted nutritional or training interventions aimed at reducing fracture injury. Further research is required to explore the applicability of these interventions to protect players from harm, but the results presented in this thesis directly contribute to our understanding of these genetic variants with fracture risk in physically active participants.

Prognostic tools to identify individuals at risk of injury are commonplace in elite football to support training interventions and management strategies to mitigate injury risk (Hughes et al., 2020). However, the ability to sensitively identify individuals at significantly increased risk of injury with enough certainty to validate a practical intervention remains questionable. Injury aetiology is a multifactorial emergent occurrence and understanding the contributory factors to try and reduce injury incidence is important and theoretically possible but further research is required (Hughes et al., 2018; Kalkhoven et al., 2021). Several prognostic factors have been presented which may provide a feasible insight into the injury susceptibility of an individual including age and previous injury (Hughes et al., 2017). However, the prognostic value of screening tools remains limited, and more information is needed to identify the at-risk individual (Hughes et al., 2020). Therefore, greater understanding of the inherent genetic variations affecting injury susceptibility could further support the individualised training decisions to support long-term development of every individual. The *MMP3* (rs679620) and *VDR* (rs2228570) SNPs appear to be associated with the incidence of all non-contact injuries, and *COL5A1* (rs12722) specifically associated with non-contact muscle injuries. Non-contact injuries are more preventable and attributable to variations in training load and / or maturation than contact injuries (Gabbett, 2016; Kalkhoven et al., 2021; Read et al., 2018). Furthermore, all three of these SNPs were recently implicated in variations in exercise induced muscle damage (Baumert et al., 2022). Therefore, the relationship between loading exposure and the *MMP3* (rs679620), *VDR* (rs2228570) and *COL5A1* (rs12722) SNPs may provide more granular information, which could support the practical application of genetically individualised training programmes to reduce the risk of injury.

Nevertheless, the most accurate evaluation of an individual's current physical performance, strength and movement competency remains direct assessment of the athlete at that time (Wackerhage & Schoenfeld, 2021). The results of this thesis should not be considered as a replacement for direct measurement and monitoring of physical performance qualities, which remain the most effective evaluation of the tissue-specific load capacity of an individual at that moment (Wackerhage & Schoenfeld, 2021). Indeed, the overall conclusion of the research is that despite

some promising observations, no consistent, strong and clearly identifiable genetic effect capable of directing an applied intervention with confidence in an elite male youth football was observed at this time. Nevertheless, rejecting the potential utility of genetic information also represents an unhelpful extreme. Continuing to develop a greater understanding of the genetic influence on injury risk could strengthen the power of predictive models to guide practical implications for coaches to support the long-term development of every individual. Understanding the influence of genetic predisposition represents a fundamental component of the athlete's inherent susceptibility that should be acknowledged within the decision-making process of training exposure to minimise injury risk. Integrating the influence of physiological and environmental factors could further increase this risk stratification as injury is a dynamic and complex phenomenon. Nevertheless, identifying which genetic variants contribute to the heritable variability in injury susceptibility is challenging and healthy scepticism of initial findings is justifiable. Eventually, the lines between genetic and environmental factors may become blurred, as epigenetic modifications mediate the genetic influence as a result of environmental experiences and exposures (Ehlert et al., 2013; Pickering & Kiely, 2019). Epigenetic modifications can be heritable but may also be modified over time to have important implications for the adaptive response to training, long-term development, and injury risk (Ehlert et al., 2013; Pickering & Kiely, 2019). Therefore, future research focused on understanding the complex interaction between nature and nurture, acknowledging individual differences, which may change over time, could support the long-term development for every individual.

CHAPTER 8: References

Abbassi, V. (1998). Growth and normal puberty. *Pediatrics, 102*(2 III), 507–511. https://doi.org/10.1542/peds.102.2.S1.507

Abrams, S. A. (2003). Normal acquisition and loss of bone mass. *Hormone Research, 60* Suppl 3(SUPPL. 3), 71–76. https://doi.org/10.1159/000074505

Abt, G., & Lovell, R. (2009). The use of individualized speed and intensity thresholds for determining the distance run at high-intensity in professional soccer. *Journal of Sports Sciences, 27*(9), 893–898. https://doi.org/10.1080/02640410902998239

Achar, S., & Yamanaka, J. (2019). Apophysitis and Osteochondrosis: Common Causes of Pain in Growing Bones. *American Family Physician, 99*(10), 610–618. http://www.ncbi.nlm.nih.gov/pubmed/31083875

Ahmetov, I., Druzhevskaya, A. M., Astratenkova, I. V., Popov, D. V., Vinogradova, O. L., & Rogozkin, V. A. (2010). The ACTN3 R577X polymorphism in Russian endurance athletes. *British Journal of Sports Medicine, 44*(9), 649–652. https://doi.org/10.1136/bjsm.2008.051540

Ahmetov, I., Egorova, E. S., Gabdrakhmanova, L. J., & Fedotovskaya, O. N. (2016). *Genes and Athletic Performance: An Update. In Medicine and Sport Science* (Vol. 61, pp. 41–54). https://doi.org/10.1159/000445240

Ahmetov, I., Gavrilov, D. N., Astratenkova, I. V., Druzhevskaya, A. M., Malinin, A. V., Romanova, E. E., & Rogozkin, V. A. (2013). The association of ACE, ACTN3 and PPARA gene variants with strength phenotypes in middle school-age children. *The Journal of Physiological Sciences, 63*(1), 79–85. https://doi.org/10.1007/s12576-012-0233-8

Ahmetov, I., Mozhayskaya, I., Flavell, D., Astratenkova, I., Komkova, A., Lyubaeva, E., Tarakin, P., Shenkman, B., Vdovina, A., Netreba, A., Popov, D., Vinogradova, O., Montgomery, H., & Rogozkin, V. (2006). PPARα gene variation and physical performance in Russian athletes. *European Journal of Applied Physiology, 97*(1), 103–108. https://doi.org/10.1007/s00421-006-0154-4

Ahmetov, I., Williams, A. G., Popov, D., Lyubaeva, E., Hakimullina, A., Fedotovskaya, O., Mozhayskaya, I., Vinogradova, O., Astratenkova, I., Montgomery, H., & Rogozkin, V. (2009). The combined impact of metabolic gene polymorphisms on elite endurance athlete status and related phenotypes. *Human Genetics, 126*(6), 751–761. https://doi.org/10.1007/s00439-009-0728-4

Alajlouni, D., Bliuc, D., Tran, T., Eisman, J. A., Nguyen, T. V, & Center, J. R. (2020). Decline in Muscle Strength and Performance Predicts Fracture Risk in Elderly Women and Men. *The Journal of Clinical Endocrinology & Metabolism, 105*(9), e3363–e3373. https://doi.org/10.1210/clinem/dgaa414

Alfred, T., Ben-Shlomo, Y., Cooper, R., Hardy, R., Cooper, C., Deary, I. J., Gunnell, D., Harris, S. E., Kumari, M., Martin, R. M., Moran, C. N., Pitsiladis, Y. P., Ring, S. M., Sayer, A. A., Smith, G. D., Starr, J. M., Kuh, D., & Day, I. N. M. (2011). ACTN3 genotype, athletic status, and life course physical capability: meta-analysis of the published literature and findings from nine studies. *Human Mutation, 32*(9), 1008–1018. https://doi.org/10.1002/humu.21526

Almeida, J. A., Boullosa, D. A., Pardono, E., Lima, R. M., Morais, P. K., Denadai, B. S., Souza, V. C., Nóbrega, O. T., Campbell, C. S. G., & Simões, H. G. (2012). The influence of ACE genotype on cardiovascular fitness of moderately active young men. *Arquivos Brasileiros de Cardiologia, 98*(4), 315–320. https://doi.org/10.1590/s0066-782x2012005000029

Alonso, P., Gratacòs, M., Segalàs, C., Escaramís, G., Real, E., Bayés, M., Labad, J., Pertusa, A., Vallejo, J., Estivill, X., & Menchón, J. M. (2011). Variants in estrogen receptor alpha gene are associated with phenotypical expression of obsessive-compulsive disorder. *Psychoneuroendocrinology, 36*(4), 473–483. https://doi.org/10.1016/j.psyneuen.2010.07.022

Altinisik, J., Meric, G., Erduran, M., Ates, O., Ulusal, A. E., & Akseki, D. (2015). The BstUI and DpnII Variants of the COL5A1 Gene Are Associated With Tennis Elbow. *The American Journal of Sports Medicine, 43*(7), 1784–1789. https://doi.org/10.1177/0363546515578661

Alves, C. R., Fernandes, T., Lemos, J. R., Magalhães, F. de C., Trombetta, I. C., Alves, G. B., Mota, G. de F. A. da, Dias, R. G., Pereira, A. C., Krieger, J. E., Negrão, C. E., & Oliveira, E. M. (2018). Aerobic exercise training

differentially affects ACE C- and N-domain activities in humans: Interactions with ACE I/D polymorphism and association with vascular reactivity. *JRAAS - Journal of the Renin-Angiotensin-Aldosterone System, 19*(2). https://doi.org/10.1177/1470320318761725

Alves, G. B., Oliveira, E. M., Alves, C. R., Rached, H. R. S., Mota, G. F. A., Pereira, A. C., Rondon, M. U., Hashimoto, N. Y., Azevedo, L. F., Krieger, J. E., & Negrão, C. E. (2009). Influence of angiotensinogen and angiotensin-converting enzyme polymorphisms on cardiac hypertrophy and improvement on maximal aerobic capacity caused by exercise training. *European Journal of Cardiovascular Prevention & Rehabilitation, 16*(4), 487–492. https://doi.org/10.1097/HJR.0b013e32832c5a8a

American Evaluation Association. (2014). What is evaluation? chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=https%3A%2F%2Fwww.eval.org%2FPortals%2F0%2FWhat%2520is%2520evaluation%2520Document.pdf&clen=576728&chunk=true

Amir, O., Amir, R., Yamin, C., Attias, E., Eynon, N., Sagiv, M., Sagiv, M., & Meckel, Y. (2007). The ACE deletion allele is associated with Israeli elite endurance athletes. *Experimental Physiology, 92*(5), 881–886. https://doi.org/10.1113/expphysiol.2007.038711

Anderson, L. N., Heong, S. W., Chen, Y., Thorpe, K. E., Adeli, K., Howard, A., Sochett, E., Birken, C. S., Parkin, P. C., Maguire, J. L., & TARGet Kids Collaboration. (2017). Vitamin D and Fracture Risk in Early Childhood: A Case-Control Study. *American Journal of Epidemiology, 185*(12), 1255–1262. https://doi.org/10.1093/aje/kww204

Andrew, T., Antioniades, L., Scurrah, K. J., MacGregor, A. J., & Spector, T. D. (2004). Risk of Wrist Fracture in Women Is Heritable and Is Influenced by Genes That Are Largely Independent of Those Influencing BMD. *Journal of Bone and Mineral Research, 20*(1), 67–74. https://doi.org/10.1359/JBMR.041015

Arai, H., Miyamoto, K.-I., Taketani, Y., Yamamoto, H., Iemori, Y., Morita, K., Tonai, T., Nishisho, T., Mori, S., & Takeda, E. (1997). A Vitamin D Receptor Gene Polymorphism in the Translation Initiation Codon: Effect on Protein Activity and Relation to Bone Mineral Density in Japanese Women. *Journal of Bone and Mineral Research, 12*(6), 915–921. https://doi.org/10.1359/jbmr.1997.12.6.915

Arendt, E. A., Agel, J., & Dick, R. (1999). Anterior cruciate ligament injury patterns among collegiate men and women. *Journal of Athletic Training, 34*(2), 86–92. http://www.ncbi.nlm.nih.gov/pubmed/16558564

Arnold, A., Thigpen, C. A., Beattie, P. F., Kissenberth, M. J., & Shanley, E. (2017). Overuse Physeal Injuries in Youth Athletes. *Sports Health: A Multidisciplinary Approach, 9*(2), 139–147. https://doi.org/10.1177/1941738117690847

Arruda, A. F. S., Carling, C., Zanetti, V., Aoki, M. S., Coutts, A. J., & Moreira, A. (2015). Effects of a Very Congested Match Schedule on Body-Load Impacts, Accelerations, and Running Measures in Youth Soccer Players. *International Journal of Sports Physiology and Performance, 10*(2), 248–252. https://doi.org/10.1123/ijspp.2014-0148

Ash, G. I., Scott, R. A., Deason, M., Dawson, T. A., Wolde, B., Bekele, Z., Teka, S., & Pitsiladis, Y. P. (2011). No association between ACE gene variation and endurance athlete status in Ethiopians. *Medicine and Science in Sports and Exercise, 43*(4), 590–597. https://doi.org/10.1249/MSS.0b013e3181f70bd6

Atan, S. A., Foskett, A., & Ali, A. (2016). Motion analysis of match play in New Zealand U13 to U15 age-group soccer players. *Journal of Strength and Conditioning Research, 30*(9), 2416–2423. https://doi.org/10.1519/JSC.0000000000001336

Attia, J. (2003). Meta-analyses of molecular association studies: Methodologic lessons for genetic epidemiology. *Journal of Clinical Epidemiology, 56*(4), 297–303. https://doi.org/10.1016/S0895-4356(03)00011-8

Babraj, J. A., Smith, K., Cuthbertson, D. J. R., Rickhuss, P., Dorling, J. S., & Rennie, M. J. (2005). Human Bone Collagen Synthesis Is a Rapid, Nutritionally Modulated Process. *Journal of Bone and Mineral Research, 20*(6), 930–937. https://doi.org/10.1359/JBMR.050201

Bacon, C. S., & Mauger, A. R. (2017). Prediction of Overuse Injuries in Professional U18-u21 Footballers Using Metrics of Training Distance and Intensity. *Journal of Strength and Conditioning Research, 31*(11), 3067–3076. https://doi.org/10.1519/jsc.0000000000001744

Bahr, R., Clarsen, B., Derman, W., Dvorak, J., Emery, C. A., Finch, C. F., Hägglund, M., Junge, A., Kemp, S., Khan, K. M., Marshall, S. W., Meeuwisse, W., Mountjoy, M., Orchard, J. W., Pluim, B., Quarrie, K. L., Reider,

B., Schwellnus, M. P., Soligard, T., … Chamari, K. (2020). International Olympic Committee consensus statement: methods for recording and reporting of epidemiological data on injury and illness in sport 2020 (including STROBE Extension for Sport Injury and Illness Surveillance (STROBE-SIIS)). *British Journal of Sports Medicine, 54*(7), 372–389. https://doi.org/10.1136/bjsports-2019-101969

Bangsbo, J., Mohr, M., & Krustrup, P. (2006). Physical and metabolic demands of training and match-play in the elite football player. *Nutrition and Football: The FIFA/FMARC Consensus on Sports Nutrition*, *24*(July), 1–18. https://doi.org/10.4324/9780203967430

Battié, M. C., Levalahti, E., Videman, T., Burton, K., & Kaprio, J. (2008). Heritability of lumbar flexibility and the role of disc degeneration and body weight. *Journal of Applied Physiology (Bethesda, Md. : 1985), 104*(2), 379–385. https://doi.org/10.1152/japplphysiol.01009.2007

Baumert, P., Cocks, M., Strauss, J. A., Shepherd, S. O., Drust, B., Lake, M. J., Stewart, C. E., & Erskine, R. M. (2022). Polygenic mechanisms underpinning the response to exercise-induced muscle damage in humans: In vivo and in vitro evidence. *Journal of Cellular Physiology, December 2021*, 1–15. https://doi.org/10.1002/jcp.30723

Baumert, P., Hall, E. C., & Erskine, R. M. (2019). The genetic association with exercise-induced muscle damage and muscle injury risk. In *Sports, Exercise, and Nutritional Genomics: Current Status and Future Directions*. Elsevier Inc. https://doi.org/10.1016/B978-0-12-816193-7.00017-8

Baumert, P., Lake, M. J., Stewart, C. E., Drust, B., Erskine, R. M., Lake, M. J., Stewart, C. E., & Erskine, R. M. (2016). Genetic variation and exercise-induced muscle damage: implications for athletic performance, injury and ageing. *European Journal of Applied Physiology, 116*(9), 1595–1625. https://doi.org/10.1007/s00421-016-3411-1

Baumert, P., Stewart, C. E., Lake, M. J., Drust, B., & Erskine, R. M. (2018). Variations of collagen-encoding genes are associated with exercise-induced muscle damage. *Physiological Genomics, 50*(9), 691–693. https://doi.org/10.1152/physiolgenomics.00145.2017

Belli, T., Crisp, A. H., & Verlengia, R. (2017). Greater muscle damage in athletes with ACTN3 R577X (RS1815739) gene polymorphism after an ultra-endurance race: a pilot study. *Biology of Sport, 34*(2), 105–110. https://doi.org/10.5114/biolsport.2017.64583

Ben-Zaken, S., Meckel, Y., Nemet, D., & Eliakim, A. (2015). Genetic score of power-speed and endurance track and field athletes. *Scandinavian Journal of Medicine and Science in Sports, 25*(2), 166–174. https://doi.org/10.1111/sms.12141

Ben-Zaken, S., Eliakim, A., Nemet, D., & Meckel, Y. (2019). Genetic Variability Among Power Athletes: The Stronger vs. the Faster. *Journal of Strength and Conditioning Research, 33*(6), 1505–1511. https://doi.org/10.1519/JSC.0000000000001356

Bennell, K., Matheson, G., Meeuwisse, W., & Brukner, P. (1999). Risk factors for stress fractures. *Sports Medicine, 28*(2), 1–32. papers2://publication/uuid/350BAA03-79CA-435F-8DCD-7BA3598B16E7

Bertuzzi, R., Pasqua, L. A., Bueno, S., Lima-Silva, A. E., Matsuda, M., Marquezini, M., & Saldiva, P. H. (2014). Is the COL5A1 rs12722 gene polymorphism associated with running economy? *PLoS ONE, 9*(9), e106581. https://doi.org/10.1371/journal.pone.0106581

Beunen, G. P., Malina, R. M., Lefevre, J., Claessens, A. L., Renson, R., & Simons, J. (1997). Prediction of adult stature and noninvasive assessment of biological maturation. *Medicine and Science in Sports and Exercise, 29*(2), 225–230. https://doi.org/10.1097/00005768-199702000-00010

Birk, D. E., Fitch, J. M., Babiarz, J. P., Doane, K. J., & Linsenmayer, T. F. (1990). Collagen fibrillogenesis in vitro: interaction of types I and V collagen regulates fibril diameter. *Journal of Cell Science, 95* (Pt 4), 649–657. https://doi.org/2384532

Birkedal-Hansen, H., Moore, W. G. I., Bodden, M. K., Windsor, L. J., Birkedal-Hansen, B., DeCarlo, A., & Engler, J. A. (1993). Matrix Metalloproteinases: A Review. *Critical Reviews in Oral Biology & Medicine, 4*(2), 197–250. https://doi.org/10.1177/10454411930040020401

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

Blades, H. Z., Arundel, P., Carlino, W. A., Dalton, A., Crook, J. S., Freeman, J. V., & Bishop, N. J. (2010). Collagen gene polymorphisms influence fracture risk and bone mass acquisition during childhood and adolescent growth. *Bone, 47*(5), 989–994. https://doi.org/10.1016/j.bone.2010.08.014

Bouchard, C., An, P., Rice, T. K., Skinner, J. S., Wilmore, J. H., Gagnon, J., Pérusse, L., Leon, A. S., Rao, D. C., Pé russe, L., Leon, A. S., Rao, D. C., An, P., Pé russe, L., & Rao Familial aggrega-, D. C. (1999). Familial aggregation of VO2max response to exercise training: results from the HERITAGE Family Study. *Journal of Applied Physiology, 87*(3), 1003–1008.

Bouchard, C., Daw, E. W., Rice, T., Pérusse, L., Gagnon, J., Province, M. A., Leon, A. S., Rao, D. C., Skinner, J. S., & Wilmore, J. H. (1998). Familial resemblance for VO2max in the sedentary state: the HERITAGE family study. *Medicine and Science in Sports and Exercise, 30*(2), 252–258. http://www.ncbi.nlm.nih.gov/pubmed/9502354

Bouchard, C., & Rankinen, T. (2001). Individual differences in response to regular physical activity. *Medicine and Science in Sports and Exercise, 33*(6 Suppl), S446-51; discussion S452-3. https://doi.org/10.1097/00005768-200106001-00013

Bouchard, C., Sarzynski, M. A., Rice, T. K., Kraus, W. E., Church, T. S., Sung, Y. J., Rao, D. C., & Rankinen, T. (2011). Genomic predictors of the maximal O2 uptake response to standardized exercise training programs. *Journal of Applied Physiology, 110*(5), 1160–1170. https://doi.org/10.1152/japplphysiol.00973.2010

Bowen, L., Gross, A. S., Gimpel, M., Bruce-Low, S., & Li, F.-X. (2020). Spikes in acute:chronic workload ratio (ACWR) associated with a 5–7 times greater injury rate in English Premier League football players: a comprehensive 3-year study. *British Journal of Sports Medicine, 54*(12), 731–738. https://doi.org/10.1136/bjsports-2018-099422

Bowen, L., Gross, A. S., Gimpel, M., & Li, F.-X. (2017). Accumulated workloads and the acute:chronic workload ratio relate to injury risk in elite youth football players. *British Journal of Sports Medicine, 51*(5), 452–459. https://doi.org/10.1136/bjsports-2015-095820

Bradley, P. S., Carling, C., Archer, D., Roberts, J., Dodds, A., Di Mascio, M., Paul, D., Gomez Diaz, A., Peart, D., & Krustrup, P. (2011). The effect of playing formation on high-intensity running and technical profiles in English FA Premier League soccer matches. *Journal of Sports Sciences, 29*(8), 821–830. https://doi.org/10.1080/02640414.2011.561868

Bradley, P. S., Sheldon, W., Wooster, B., Olsen, P., Boanas, P., & Krustrup, P. (2009). High-intensity running in English FA Premier League soccer matches. *Journal of Sports Sciences, 27*(2), 159–168. https://doi.org/10.1080/02640410802512775

Brink, M. S., Visscher, C., Arends, S., Zwerver, J., Post, W. J., & Lemmink, K. A. P. M. (2010). Monitoring stress and recovery: new insights for the prevention of injuries and illnesses in elite youth soccer players. *British Journal of Sports Medicine, 44*(11), 809–815. https://doi.org/10.1136/bjsm.2009.069476

Briški, N., Vrgoč, G., Knjaz, D., Janković, S., Ivković, A., Pećina, M., & Lauc, G. (2021). Association of the matrix metalloproteinase 3 (MMP3) single nucleotide polymorphisms with tendinopathies: case-control study in high-level athletes. *International Orthopaedics, 45*(5), 1163–1168. https://doi.org/10.1007/s00264-020-04684-w

Brito, J., Malina, R. M., Seabra, A., Massada, J. L., Soares, J. M., Krustrup, P., & Rebelo, A. (2012). Injuries in Portuguese Youth Soccer Players During Training and Match Play. *Journal of Athletic Training, 47*(2), 191–197. https://doi.org/10.4085/1062-6050-47.2.191

Broos, S., Malisoux, L., Theisen, D., Van Thienen, R., Francaux, M., Thomis, M. A., & Deldicque, L. (2019). The stiffness response of type IIa fibres after eccentric exercise-induced muscle damage is dependent on ACTN3 r577X polymorphism. *European Journal of Sport Science, 19*(4), 480–489. https://doi.org/10.1080/17461391.2018.1529200

Broos, S., Malisoux, L., Theisen, D., van Thienen, R., Ramaekers, M., Jamart, C., Deldicque, L., Thomis, M. A., & Francaux, M. (2016). Evidence for ACTN3 as a Speed Gene in Isolated Human Muscle Fibers. *PLOS ONE, 11*(3), e0150594. https://doi.org/10.1371/journal.pone.0150594

Broos, S., Van Leemputte, M., Deldicque, L., & Thomis, M. A. (2015). History-dependent force, angular velocity and muscular endurance in ACTN3 genotypes. *European Journal of Applied Physiology, 115*(8), 1637–1643. https://doi.org/10.1007/s00421-015-3144-6

Brown, J. C., Miller, C.-J., Posthumus, M., Schwellnus, M. P., & Collins, M. (2011a). The COL5A1 gene, ultra-marathon running performance, and range of motion. *International Journal of Sports Physiology and Performance, 6*(4), 485–496. https://doi.org/10.2307/2383716

Brown, J. C., Miller, C.-J., Schwellnus, M. P., & Collins, M. (2011b). Range of motion measurements diverge with increasing age for COL5A1 genotypes. *Scandinavian Journal of Medicine & Science in Sports, 21*(6), e266-72. https://doi.org/10.1111/j.1600-0838.2010.01271.x

Brown, K. L., Seale, K. B., El Khoury, L. Y., Posthumus, M., Ribbans, W. J., Raleigh, S. M., Collins, M., & September, A. V. (2016). Polymorphisms within the COL5A1 gene and regulators of the extracellular matrix modify the risk of Achilles tendon pathology in a British case-control study. *Journal of Sports Sciences, 35*(15), 1–9. https://doi.org/10.1080/02640414.2016.1221524

Brown, M. A., Haughton, M. A., Grant, S. F. A., Gunnell, A. S., Henderson, N. K., & Eisman, J. A. (2001). Genetic Control of Bone Density and Turnover: Role of the Collagen 1α1, Estrogen Receptor, and Vitamin D Receptor Genes. *Journal of Bone and Mineral Research, 16*(4), 758–764. https://doi.org/10.1359/jbmr.2001.16.4.758

Buchheit, M. (2017). Applying the acute:chronic workload ratio in elite football: worth the effort? *British Journal of Sports Medicine, 51*(18), 1325–1327. https://doi.org/10.1136/bjsports-2016-097017

Buchheit, M., Allen, A., Poon, T. K., Modonutti, M., Gregson, W., & Di Salvo, V. (2014). Integrating different tracking systems in football: multiple camera semi-automatic system, local position measurement and GPS technologies. *Journal of Sports Sciences, 32*(20), 1844–1857. https://doi.org/10.1080/02640414.2014.942687

Buchheit, M., Mendez-Villanueva, A., Simpson, B. M., & Bourdon, P. C. (2010a). Match running performance and fitness in youth soccer. *International Journal of Sports Medicine, 31*(11), 818–825. https://doi.org/10.1055/s-0030-1262838

Buchheit, M., Mendez-Villanueva, A., Simpson, B. M., & Bourdon, P. C. (2010b). Repeated-sprint sequences during youth soccer matches. *International Journal of Sports Medicine, 31*(10), 709–716. https://doi.org/10.1055/s-0030-1261897

Buckthorpe, M., Wright, S., Bruce-Low, S., Nanni, G., Sturdy, T., Gross, A. S., Bowen, L., Styles, B., Della Villa, S., Davison, M., & Gimpel, M. (2019). Recommendations for hamstring injury prevention in elite football: translating research into practice. *British Journal of Sports Medicine, 53*(7), 449–456. https://doi.org/10.1136/bjsports-2018-099616

Bullen, E. C., Longaker, M. T., Updike, D. L., Benton, R., Ladin, D., Hou, Z., & Howard, E. W. (1995). Tissue inhibitor of metalloproteinases-1 is decreased and activated gelatinases are increased in chronic wounds. *The Journal of Investigative Dermatology, 104*(2), 236–240. https://doi.org/10.1111/1523-1747.ep12612786

Burden, R. J., Biswas, A., & McCarron, J. P. (2022). Elite sport, innovation and the application gap. *Experimental Physiology*, *107*(1), 3–5. https://doi.org/10.1113/EP090194

Bush, M., Barnes, C., Archer, D. T., Hogg, B., & Bradley, P. S. (2015). Evolution of match performance parameters for various playing positions in the English Premier League. *Human Movement Science, 39*, 1–11. https://doi.org/10.1016/j.humov.2014.10.003

Buxens, A., Ruiz, J. R., Arteta, D., Artieda, M., Santiago, C., González-Freire, M., Martínez, A., Tejedor, D., Lao, J. I., Gómez-Gallego, F., & Lucia, A. (2011). Can we predict top-level sports performance in power vs endurance events? A genetic approach. *Scandinavian Journal of Medicine & Science in Sports, 21*(4), 570–579. https://doi.org/10.1111/j.1600-0838.2009.01079.x

Buxton, P., Edwards, C., Archer, C. W., & Francis-West, P. (2001). Growth/Differentiation Factor-5 (GDF-5) and Skeletal Development. *The Journal of Bone and Joint Surgery-American Volume, 83*(SUPPL. 1), S1-23–S1–30. https://doi.org/10.2106/00004623-200100001-00004

Caine, D., Maffulli, N., & Caine, C. (2008). Epidemiology of Injury in Child and Adolescent Sports: Injury Rates, Risk Factors, and Prevention. *Clinics in Sports Medicine, 27*(1), 19–50. https://doi.org/10.1016/j.csm.2007.10.008

Cam, S., Colakoglu, M., Colakoglu, S., Sekuri, C., & Berdeli, A. (2007). ACE I/D gene polymorphism and aerobic endurance development in response to training in a non-elite female cohort. *Journal of Sports Medicine and Physical Fitness, 47*(2), 234–238. https://doi.org/10.1111/j.1600-0838.2006.00590.x

Cam, S., Colakoglu, M., Sekuri, C., Colakoglu, S., Sahan, Ç., & Berdeli, A. (2005). Association between the ACE I/D gene polymorphism and physical performance in a homogeneous non-elite cohort. *Canadian Journal of Applied Physiology, 30*(1), 74–86. https://doi.org/10.1139/h05-106

Carling, C., McCall, A., Le Gall, F., & Dupont, G. (2016). The impact of short periods of match congestion on injury risk and patterns in an elite football club. *British Journal of Sports Medicine, 50*(12), 764–768. https://doi.org/10.1136/bjsports-2015-095501

Cauci, S., Migliozzi, F., Trombetta, C. S., Venuto, I., Saccheri, P., Travan, L., & Chiriacò, G. (2017). Low back pain and FokI (rs2228570) polymorphism of vitamin D receptor in athletes. *BMC Sports Science, Medicine & Rehabilitation, 9*(1), 4. https://doi.org/10.1186/s13102-017-0069-x

Cerit, M., Colakoglu, M., Erdogan, M., Berdeli, A., & Cam, F. S. (2006). Relationship between ace genotype and short duration aerobic performance development. *European Journal of Applied Physiology, 98*(5), 461–465. https://doi.org/10.1007/s00421-006-0286-6

Chakravarti, A. (1999). Population genetics--making sense out of sequence. *Nature Genetics, 21*(1 Suppl), 56–60. https://doi.org/10.1038/4482

Charbonneau, D. E., Hanson, E. D., Ludlow, A. T., Delmonico, M. J., Hurley, B. F., & Roth, S. M. (2008). ACE genotype and the muscle hypertrophic and strength responses to strength training. *Medicine and Science in Sports and Exercise, 40*(4), 677–683. https://doi.org/10.1249/MSS.0b013e318161eab9

Chatzipapas, C., Boikos, S., Drosos, G. I., Kazakos, K., Tripsianis, G., Serbis, A., Stergiopoulos, S., Tilkeridis, C., Verettas, D.-A., & Stratakis, C. A. (2009). Polymorphisms of the vitamin D receptor gene and stress fractures. *Hormone and Metabolic Research = Hormon- Und Stoffwechselforschung = Hormones et Metabolisme, 41*(8), 635–640. https://doi.org/10.1055/s-0029-1216375

Chen, B., Li, B., Qi, Y.-J., Tie, K., & Chen, L.-B. (2015). Association study between growth differentiation factor 5 polymorphism and non-contact anterior cruciate ligament rupture in Chinese Han population. *International Journal of Clinical and Experimental Medicine, 8*(12), 22484–22490. http://www.ncbi.nlm.nih.gov/pubmed/26885231%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4730017

Chin, E. R., Olson, E. N., Richardson, J. A., Yang, Q., Humphries, C., Shelton, J. M., Wu, H., Zhu, W., Bassel-Duby, R., & Williams, R. S. (1998). A calcineurin-dependent transcriptional pathway controls skeletal muscle fiber type. *Genes & Development, 12*(16), 2499–2509. https://doi.org/10.1101/gad.12.16.2499

Cieszczyk, P., Krupecki, K., Maciejewska, A., & Sawczuk, M. (2009). The Angiotensin Converting Enzyme Gene I/D Polymorphism in Polish Rowers. *International Journal of Sports Medicine, 30*(08), 624–627. https://doi.org/10.1055/s-0029-1202825

Cięszczyk, P., Leońska-Duniec, A., Maciejewska-Skrendo, A., Sawczuk, M., Leźnicka, K., Contrò, V., Trybek, G., & Lulińska-Kuklik, E. (2016). Variation in the Ace Gene in Elite Polish Football Players. *Human Movement, 17*(4), 237–241. https://doi.org/10.1515/humo-2016-0032

Circi, E., Atalay, Y., & Beyzadeoglu, T. (2017). Treatment of Osgood-Schlatter disease: review of the literature. *Musculoskeletal Surgery*, 1–6. https://doi.org/10.1007/s12306-017-0479-7

Clark, E. M., Tobias, J. H., Murray, L., & Boreham, C. (2011). Children with low muscle strength are at an increased risk of fracture with exposure to exercise. *Journal of Musculoskeletal & Neuronal Interactions, 11*(2), 196–202. http://www.ncbi.nlm.nih.gov/pubmed/21625056

Clarkson, P. M., Hoffman, E. P., Zambraski, E., Gordish-Dressman, H., Kearns, A., Hubal, M., Harmon, B., & Devaney, J. M. (2005). ACTN3 and MLCK genotype associations with exertional muscle damage. *Journal of Applied Physiology (Bethesda, Md. : 1985), 99*(2), 564–569. https://doi.org/10.1152/japplphysiol.00130.2005

Clos, E., Pruna, R., Lundblad, M., Artells, R., & Esquirol Caussa, J. (2019). ACTN3 single nucleotide polymorphism is associated with non-contact musculoskeletal soft-tissue injury incidence in elite professional football players. *Knee Surgery, Sports Traumatology, Arthroscopy : Official Journal of the ESSKA, 27*(12), 4055–4061. https://doi.org/10.1007/s00167-019-05381-x

Coelho-E-Silva, M. J., Moreira Carvalho, H., Gonçalves, C. E., Figueiredo, A. J., Elferink-Gemser, M. T., Philippaerts, R. M., & Malina, R. M. (2010). Growth, maturation, functional capacities and sport-specific skills in 12-13 year-old- basketball players. *The Journal of Sports Medicine and Physical Fitness, 50*(2), 174–181. http://www.ncbi.nlm.nih.gov/pubmed/20585295

Coelho, D. B., Pimenta, E., Rosse, I. C., Veneroso, C., Pussieldi, G., Becker, L. K., Carvalho, M.-R., & Silami-Garcia, E. (2016). Angiotensin-converting enzyme (ACE-I/D) polymorphism frequency in Brazilian soccer players. *Applied Physiology, Nutrition, and Metabolism, 41*(6), 692–694. https://doi.org/10.1139/apnm-2015-0514

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd Editio). Routledge.

Colakoglu, M., Cam, F. S., Kayitken, B., Cetinoz, F., Colakoglu, S., Turkmen, M., & Sayin, M. (2005). ACE Genotype May Have an Effect on Single versus Multiple Set Preferences in Strength Training. *European Journal of Applied Physiology, 95*(1), 20–26. https://doi.org/10.1007/s00421-005-1335-2

Collins, M., Mokone, G. G., September, A. V., Van Der Merwe, L., & Schwellnus, M. P. (2009). The COL5A1 genotype is associated with range of motion measurements. *Scandinavian Journal of Medicine and Science in Sports, 19*(6), 803–810. https://doi.org/10.1111/j.1600-0838.2009.00915.x

Collins, M., & Raleigh, S. M. (2009). Genetic risk factors for musculoskeletal soft tissue injuries. *Medicine and Sport Science, 54*(February), 136–149. https://doi.org/10.1159/000235701

Collins, M., Xenophontos, S. L., Cariolou, M. A., Mokone, G. G., Hudson, D. E., Anastasiades, L., & Noakes, T. D. (2004). The ACE gene and endurance performance during the South African Ironman Triathlons. *Medicine and Science in Sports and Exercise, 36*(8), 1314–1320. https://doi.org/10.1249/01.MSS.0000135779.41475.42

Colombini, A., Brayda-Bruno, M., Lombardi, G., Croiset, S. J., Vrech, V., Maione, V., Banfi, G., & Cauci, S. (2014). FokI polymorphism in the vitamin D receptor gene (VDR) and its association with lumbar spine pathologies in the Italian population: A case-control study. *PLoS ONE, 9*(5). https://doi.org/10.1371/journal.pone.0097027

Colombini, A., De Luca, P., Cangelosi, D., Perucca Orfei, C., Ragni, E., Viganò, M., Malacarne, M., Castagnetta, M., Brayda-Bruno, M., Coviello, D., & de Girolamo, L. (2021). High-Throughput Gene and Protein Analysis Revealed the Response of Disc Cells to Vitamin D, Depending on the VDR FokI Variants. *International Journal of Molecular Sciences, 22*(17). https://doi.org/10.3390/ijms22179603

Conn, J. M., Annest, J. L., & Gilchrist, J. (2003). Sports and recreation related injury episodes in the US population, 1997-99. *Injury Prevention, 9*(2), 117–123. https://doi.org/10.1136/ip.9.2.117

Cook, J., & Docking, S. (2015). "Rehabilitation will increase the 'capacity' of your …insert musculoskeletal tissue here…." Defining 'tissue capacity': a core concept for clinicians. *British Journal of Sports Medicine, 49*(23), 1484–1485. https://doi.org/10.1136/bjsports-2015-094849

Cooper, C., Dennison, E. M., Leufkens, H. G. M., Bishop, N., & van Staa, T. P. (2004). Epidemiology of childhood fractures in Britain: a study using the general practice research database. *Journal of Bone and Mineral Research : The Official Journal of the American Society for Bone and Mineral Research, 19*(12), 1976–1981. https://doi.org/10.1359/JBMR.040902

Cooper, D. N., Krawczak, M., Polychronakos, C., Tyler-Smith, C., & Kehrer-Sawatzki, H. (2013). Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease. *Human Genetics, 132*(10), 1077–1130. https://doi.org/10.1007/s00439-013-1331-2

Cosman, F., Ruffing, J., Zion, M., Uhorchak, J., Ralston, S., Tendy, S., McGuigan, F. E. A., Lindsay, R., & Nieves, J. (2013). Determinants of stress fracture risk in United States Military Academy cadets. *Bone, 55*(2), 359–366. https://doi.org/10.1016/j.bone.2013.04.011

Costa, A. M., Silva, A. J., Garrido, N. D., Louro, H., de Oliveira, R. J., & Breitenfeld, L. (2009). Association between ACE D allele and elite short distance swimming. *European Journal of Applied Physiology, 106*(6), 785–790. https://doi.org/10.1007/s00421-009-1080-z

Cunanan, A. J., DeWeese, B. H., Wagle, J. P., Carroll, K. M., Sausaman, R., Hornsby, W. G., Haff, G. G., Triplett, N. T., Pierce, K. C., & Stone, M. H. (2018). The General Adaptation Syndrome: A Foundation for the Concept of Periodization. *Sports Medicine, 48*(4), 787–797. https://doi.org/10.1007/s40279-017-0855-3

Curtis, E. M., van der Velde, R., Moon, R. J., van den Bergh, J. P. W., Geusens, P., de Vries, F., van Staa, T. P., Cooper, C., & Harvey, N. C. (2016). Epidemiology of fractures in the United Kingdom 1988–2012: Variation with age, sex, geography, ethnicity and socioeconomic status. *Bone, 87*(0), 19–26. https://doi.org/10.1016/j.bone.2016.03.006

Cutler, G. (1997). The Role of Estrogen in Bone Growth and Maturation During Childhood and Adolescence. *The Journal of Steroid Biochemistry and Molecular Biology, 61*(3–6), 141–144. https://doi.org/10.1016/S0960-0760(96)00219-1

Dahlquist, D. T., Dieter, B. P., & Koehle, M. S. (2015). Plausible ergogenic effects of vitamin D on athletic performance and recovery. *Journal of the International Society of Sports Nutrition, 12*(1), 33. https://doi.org/10.1186/s12970-015-0093-8

Dalen-Lorentsen, T., Bjørneboe, J., Clarsen, B., Vagle, M., Fagerland, M. W., & Andersen, T. E. (2021). Does load management using the acute:chronic workload ratio prevent health problems? A cluster randomised trial of 482 elite youth footballers of both sexes. *British Journal of Sports Medicine, 55*(2), 108–114. https://doi.org/10.1136/bjsports-2020-103003

Danser, A. H. J., Schalekamp, M. A. D. H., Bax, W. A., Van Den Brink, A. M., Saxena, P. R., Riegger, G. A. J., & Schunkert, H. (1995). Angiotensin-converting enzyme in the human heart: Effect of the deletion/insertion polymorphism. *Circulation, 92*(6), 1387–1388. https://doi.org/10.1161/01.CIR.92.6.1387

Davids, K., & Baker, J. (2007). Genes, Environment and Sport Performance. *Sports Medicine, 37*(11), 961–980. https://doi.org/10.2165/00007256-200737110-00004

de Loës, M. (1995). Epidemiology of sports injuries in the Swiss organization "Youth and Sports" 1987-1989. Injuries, exposure and risks of main diagnoses. *International Journal of Sports Medicine, 16*(2), 134–138. http://www.ncbi.nlm.nih.gov/pubmed/7751078

De Moor, M. H. M., Spector, T. D., Cherkas, L. F., Falchi, M., Hottenga, J. J., Boomsma, D. I., & De Geus, E. J. C. (2007). Genome-wide linkage scan for athlete status in 700 British female DZ twin pairs. *Twin Research and Human Genetics, 10*(6), 812–820. https://doi.org/10.1375/twin.10.6.812

Del Coso, J., Hiam, D., Houweling, P., Pérez, L. M., Eynon, N., & Lucía, A. (2019). More than a 'speed gene': ACTN3 R577X genotype, trainability, muscle damage, and the risk for injuries. *European Journal of Applied Physiology, 119*(1), 49–60. https://doi.org/10.1007/s00421-018-4010-0

Del Coso, J., Salinero, J. J., Lara, B., Gallo-Salazar, C., Areces, F., Puente, C., & Herrero, D. (2016). ACTN3 X-allele carriers had greater levels of muscle damage during a half-ironman. *European Journal of Applied Physiology, 117*(1), 151–158. https://doi.org/10.1007/s00421-016-3507-7

Del Coso, J., Valero, M., Salinero, J. J., Lara, B., Díaz, G., Gallo-Salazar, C., Ruiz-Vicente, D., Areces, F., Puente, C., Carril, J. C., & Cacabelos, R. (2017a). ACTN3 genotype influences exercise-induced muscle damage during a marathon competition. *European Journal of Applied Physiology, 117*(3), 409–416. https://doi.org/10.1007/s00421-017-3542-z

Del Coso, J., Valero, M., Salinero, J. J., Lara, B., Gallo-Salazar, C., & Areces, F. (2017b). Optimum polygenic profile to resist exertional rhabdomyolysis during a marathon. *PLOS ONE, 12*(3), e0172965. https://doi.org/10.1371/journal.pone.0172965

Delmonico, M. J., Kostek, M. C., Doldo, N. A., Hand, B. D., Walsh, S., Conway, J. M., Carignan, C. R., Roth, S. M., & Hurley, B. F. (2007). Alpha-actinin-3 (ACTN3) R577X polymorphism influences knee extensor peak power response to strength training in older men and women. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 62*(2), 206–212. https://doi.org/10.1093/gerona/62.2.206

DeLuca, H. (2005). Overview of general physiologic features and functions of vitamin D1-4. *The American Journal of Clinical Nutrition, 80*, 1689S-96S. https://doi.org/10.1093/ajcn/80.6.1689S

Demontiero, O., Vidal, C., & Duque, G. (2012). Aging and bone loss: new insights for the clinician. *Therapeutic Advances in Musculoskeletal Disease, 4*(2), 61–76. https://doi.org/10.1177/1759720X11430858

Deprez, D., Fransen, J., Boone, J., Lenoir, M., Philippaerts, R., & Vaeyens, R. (2015). Characteristics of high-level youth soccer players: variation by playing position. *Journal of Sports Sciences, 33*(3), 243–254. https://doi.org/10.1080/02640414.2014.934707

Di Cagno, A., Sapere, N., Piazza, M., Aquino, G., Iuliano, E., Intrieri, M., & Calcagno, G. (2013). ACE and AGTR1 Polymorphisms in Elite Rhythmic Gymnastics. *Genetic Testing and Molecular Biomarkers, 17*(2), 99–103. https://doi.org/10.1089/gtmb.2012.0209

Di Mauro, M., Izzicupo, P., Santarelli, F., Falone, S., Pennelli, A., Amicarelli, F., Calafiore, A. M., di Baldassarre, A., & Gallina, S. (2010). ACE and AGTR1 polymorphisms and left ventricular hypertrophy in

endurance athletes. *Medicine and Science in Sports and Exercise, 42*(5), 915–921. https://doi.org/10.1249/MSS.0b013e3181c29e79

Diet, F., Graf, C., Mahnke, N., Wassmer, G., Predel, H. G., Palma-Hohmann, I., Rost, R., & Bohm, M. (2001). ACE and angiotensinogen gene genotypes and left ventricular mass in athletes. *European Journal of Clinical Investigation, 31*(10), 836–842. https://doi.org/10.1046/j.1365-2362.2001.00886.x

DiFiori, J. P., Benjamin, H. J., Brenner, J. S., Gregory, A., Jayanthi, N., Landry, G. L., & Luke, A. (2014). Overuse injuries and burnout in youth sports: a position statement from the American Medical Society for Sports Medicine. *British Journal of Sports Medicine, 48*(4), 287–288. https://doi.org/10.1136/bjsports-2013-093299

DiGirolamo, D. J., Clemens, T. L., & Kousteni, S. (2012). The skeleton as an endocrine organ. *In Nature Reviews Rheumatology* (Vol. 8, Issue 11, pp. 674–683). https://doi.org/10.1038/nrrheum.2012.157

Diogenes, M. E. L., Bezerra, F. F., Cabello, G. M. K., Cabello, P. H., Mendonça, L. M. C., Oliveira Júnior, A. V, & Donangelo, C. M. (2010). Vitamin D receptor gene FokI polymorphisms influence bone mass in adolescent football (soccer) players. *European Journal of Applied Physiology, 108*(1), 31–38. https://doi.org/10.1007/s00421-009-1191-6

Dionísio, T. J., Thiengo, C. R., Brozoski, D. T., Dionísio, E. J., Talamoni, G. A., Silva, R. B., Garlet, G. P., Santos, C. F., & Amaral, S. L. (2017). The influence of genetic polymorphisms on performance and cardiac and hemodynamic parameters among Brazilian soccer players. *Applied Physiology, Nutrition, and Metabolism, 42*(6), 596–604. https://doi.org/10.1139/apnm-2016-0608

Dionne, F. T., Turcotte, L., Thibault, M. C., Boulay, M. R., Skinner, J. S., & Bouchard, C. (1991). Mitochondrial DNA sequence polymorphism, VO2max, and response to endurance training. *Medicine and Science in Sports and Exercise, 23*(2), 177–185. https://doi.org/10.1249/00005768-199102000-00006

Djarova, T., Bardarev, D., Boyanov, D., Kaneva, R., & Atanasov, P. (2013). Performance enhancing genetic variants, oxygen uptake, heart rate, blood pressure and body mass index of elite high altitude mountaineers. *Acta Physiologica Hungarica, 100*(3), 289–301. https://doi.org/10.1556/APhysiol.100.2013.3.5

Dzau, V. J. (1993). Tissue Renin-Angiotensin System in Myocardial Hypertrophy and Failure. *Archives of Internal Medicine, 153*(8), 937. https://doi.org/10.1001/archinte.1993.00410080011002

Efstathiadou, Z., Tsatsoulis, A., & Ioannidis, J. P. A. (2001). Association of Collagen Iα 1 Sp1 Polymorphism with the Risk of Prevalent Fractures: A Meta-Analysis. *Journal of Bone and Mineral Research, 16*(9), 1586–1592. https://doi.org/10.1359/jbmr.2001.16.9.1586

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ (Clinical Research Ed.), 315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Ehlert, T., Simon, P., & Moser, D. A. (2013). Epigenetics in Sports. *Sports Medicine, 43*(2), 93–110. https://doi.org/10.1007/s40279-012-0012-y

El Khoury, L. Y., Ribbans, W. J., & Raleigh, S. M. (2016). MMP3 and TIMP2 gene variants as predisposing factors for Achilles tendon pathologies: Attempted replication study in a British case-control cohort. *Meta Gene, 9*, 52–55. https://doi.org/10.1016/j.mgene.2016.03.007

Elengard, T., Karlsson, J., & Silbernagel, K. G. (2010). Aspects of treatment for posterior heel pain in young athletes. *Open Access Journal of Sports Medicine, 1*, 223–232. https://doi.org/10.2147/OAJSM.S15413

Emery, C. A., Meeuwisse, W. H., & Hartmann, S. E. (2005). Evaluation of Risk Factors for Injury in Adolescent Soccer. *The American Journal of Sports Medicine, 33*(12), 1882–1891. https://doi.org/10.1177/0363546505279576

Emery, C. A., Roy, T.-O., Whittaker, J. L., Nettel-Aguirre, A., & van Mechelen, W. (2015). Neuromuscular training injury prevention strategies in youth sport: a systematic review and meta-analysis. *British Journal of Sports Medicine, 49*(13), 865–870. https://doi.org/10.1136/bjsports-2015-094639

Engelman, C. D., Fingerlin, T. E., Langefeld, C. D., Hicks, P. J., Rich, S. S., Wagenknecht, L. E., Bowden, D. W., & Norris, J. M. (2008). Genetic and Environmental Determinants of 25-Hydroxyvitamin D and 1,25-Dihydroxyvitamin D Levels in Hispanic and African Americans. *The Journal of Clinical Endocrinology & Metabolism, 93*(9), 3381–3388. https://doi.org/10.1210/jc.2007-2702

Erduran, M., Altinisik, J., Meric, G., Ates, O., Ulusal, A. E., & Akseki, D. (2014). Is Sp1 site polymorphism within COL1A1 gene associated with tennis elbow? *Gene, 537*(2), 308–311. https://doi.org/10.1016/j.gene.2013.12.014

Ernst & Young LLP. (2022). Premier League Economic and social impact. The Premier League. https://www.premierleague.com/news/2435089#:~:text=Between%20the%202012%2F13%20and,EFL%20to%20develop%20homegrown%20talent.

Erskine, R. M., Williams, A. G., Jones, D. A., Stewart, C. E., & Degens, H. (2014). The individual and combined influence of ACE and ACTN3 genotypes on muscle phenotypes before and after strength training. *Scandinavian Journal of Medicine & Science in Sports, 24*(4), 642–648. https://doi.org/10.1111/sms.12055

Eynon, N., Alves, A. J., Yamin, C., Sagiv, M., Duarte, J. A., Oliveira, J., Ayalon, M., Goldhammer, E., Sagiv, M., & Meckel, Y. (2009). Is There an ACE ID – ACTN3 R577X Polymorphisms Interaction that Influences Sprint Performance? *International Journal of Sports Medicine, 30*(12), 888–891. https://doi.org/10.1055/s-0029-1238291

Faigenbaum, A. D., Kraemer, W. J., Blimkie, C. J. R., Jeffreys, I., Micheli, L. J., Nitka, M., & Rowland, T. W. (2009). Youth Resistance Training: Updated Position Statement Paper From the National Strength and Conditioning Association. *Journal of Strength and Conditioning Research, 23*, S60–S79. https://doi.org/10.1519/JSC.0b013e31819df407

Falconer, D. S., & MacKay, T. F. S. (2011). Introduction to qunatitative genetics (4th ed.).

Falk, B., & Dotan, R. (2006). Child-Adult Differences in the Recovery from High-Intensity Exercise. *Exercise and Sport Sciences Reviews, 34*(3), 107–112. https://doi.org/10.1249/00003677-200607000-00004

Fanchini, M., Rampinini, E., Riggio, M., Coutts, A. J., Pecci, C., & McCall, A. (2018). Despite association, the acute:chronic work load ratio does not predict non-contact injury in elite footballers. *Science and Medicine in Football, 2*(2), 1–7. https://doi.org/10.1080/24733938.2018.1429014

Fanchini, M., Steendahl, I. B., Impellizzeri, F. M., Pruna, R., Dupont, G., Coutts, A. J., Meyer, T., & McCall, A. (2020). Exercise-Based Strategies to Prevent Muscle Injury in Elite Footballers: A Systematic Review and Best Evidence Synthesis. *Sports Medicine, 0123456789*. https://doi.org/10.1007/s40279-020-01282-z

Fatini, C., Guazzelli, R., Manetti, P., Battaglini, B., Gensini, F., Vono, R., Toncelli, L., Zilli, P., Capalbo, A., Abbate, R., Gensini, G. F., & Galanti, G. (2000). RAS genes influence exercise-induced left ventricular hypertrophy: an elite athletes study. *Medicine and Science in Sports and Exercise, 32*(11), 1868–1872. https://doi.org/10.1097/00005768-200011000-00008

Faude, O., Rößler, R., & Junge, A. (2013). Football injuries in children and adolescent players: Are there clues for prevention? *Sports Medicine, 43*(9), 819–837. https://doi.org/10.1007/s40279-013-0061-x

Faude, O., Rössler, R., Petushek, E. J., Roth, R., Zahner, L., & Donath, L. (2017). Neuromuscular Adaptations to Multimodal Injury Prevention Programs in Youth Sports: A Systematic Review with Meta-Analysis of Randomized Controlled Trials. *Frontiers in Physiology, 8*(OCT), 1–15. https://doi.org/10.3389/fphys.2017.00791

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Faulkner, R. A., Davison, K. S., Bailey, D. A., Mirwald, R. L., & Baxter-Jones, A. (2006). Size-Corrected BMD Decreases During Peak Linear Growth: Implications for Fracture Incidence During Adolescence. *Journal of Bone and Mineral Research, 21*(12), 1864–1870. https://doi.org/10.1359/jbmr.060907

Ficek, K., Cięszczyk, P., Kaczmarczyk, M., Maciejewska-Karłowska, A., Sawczuk, M., Cholewinski, J., Leońska-Duniec, A., Stępień-Słodkowska, M., Zarębska, A., Stepto, N. K., Bishop, D. J., & Eynon, N. (2013). Gene variants within the COL1A1 gene are associated with reduced anterior cruciate ligament injury in professional soccer players. *Journal of Science and Medicine in Sport, 16*(5), 396–400. https://doi.org/10.1016/j.jsams.2012.10.004

Folland, J., Leach, B., Little, T., Hawker, K., Myerson, S., Montgomery, H., & Jones, D. (2000). Angiotensin-converting enzyme genotype affects the response of human skeletal muscle to functional overload. *Experimental Physiology, 85*(5), 575–579. http://www.ncbi.nlm.nih.gov/pubmed/11038409

Fonseca, H., Moreira-Gonçalves, D., Coriolano, H.-J. A. J. A., & Duarte, J. A. (2014). Bone Quality: The Determinants of Bone Strength and Fragility. *Sports Medicine, 44*(1), 37–53. https://doi.org/10.1007/s40279-013-0100-7

Foster, B. P., Morse, C. I., Onambele, G. L., & Williams, A. G. (2014). Human COL5A1 rs12722 gene polymorphism and tendon properties in vivo in an asymptomatic population. *European Journal of Applied Physiology, 114*(7), 1393–1402. https://doi.org/10.1007/s00421-014-2868-z

Fridén, J., & Lieber, R. (2001). Eccentric exercise-induced injuries to contractile and cytoskeletal muscle fibre components. *Acta Physiologica Scandinavica, 171*(3), 321–326. https://doi.org/10.1046/j.1365-201X.2001.00834.x

Fuller, C. W. (2006). Consensus statement on injury definitions and data collection procedures in studies of football (soccer) injuries. *British Journal of Sports Medicine, 40*(3), 193–201. https://doi.org/10.1136/bjsm.2005.025270

Fuller, C. W. (2019). Assessing the Return on Investment of Injury Prevention Procedures in Professional Football. *Sports Medicine, 49*(4), 621–629. https://doi.org/10.1007/s40279-019-01083-z

Gabbett, T. J. (2016). The training—injury prevention paradox: should athletes be training smarter and harder? *British Journal of Sports Medicine, 50*(5), 273–280. https://doi.org/10.1136/bjsports-2015-095788

Gabbett, T. J. (2018). Debunking the myths about training load, injury and performance: empirical evidence, hot topics and recommendations for practitioners. *British Journal of Sports Medicine, 0*(0), bjsports-2018-099784. https://doi.org/10.1136/bjsports-2018-099784

García-Hermoso, A., Ramírez-Campillo, R., & Izquierdo, M. (2019). Is Muscular Fitness Associated with Future Health Benefits in Children and Adolescents? A Systematic Review and Meta-Analysis of Longitudinal Studies. *Sports Medicine, 49*(7), 1079–1094. https://doi.org/10.1007/s40279-019-01098-6

Garton, F. C., Seto, J. T., Quinlan, K. G. R., Yang, N., Houweling, P. J., & North, K. N. (2014). α-Actinin-3 deficiency alters muscle adaptation in response to denervation and immobilization. *Human Molecular Genetics, 23*(7), 1879–1893. https://doi.org/10.1093/hmg/ddt580

Gayagay, G., Yu, B., Hambly, B., Boston, T., Hahn, A., Celermajer, D. S., & Trent, R. J. (1998). Elite endurance athletes and the ACE I allele--the role of genes in athletic performance. *Human Genetics, 103*(1), 48–50. https://doi.org/10.1007/s004390050781

Ge, W., Mu, J., & Huang, C. (2014). The GDF5 SNP is associated with meniscus injury and function recovery in male Chinese soldiers. International Journal of Sports Medicine, 35(7), 625–628. https://doi.org/10.1055/s-0033-1355417

Gentil, P., Lima, R. M., Pereira, R. W., Mourot, J., Leite, T. K., & Bottaro, M. (2012). Lack of association of the ACE genotype with the muscle strength response to resistance training. *European Journal of Sport Science, 12*(4), 331–337. https://doi.org/10.1080/17461391.2011.573581

Gentil, P., Pereira, R. W., Leite, T. K. M., & Bottaro, M. (2011). ACTN3 R577X Polymorphism and Neuromuscular Response to Resistance Training. *Journal of Sports Science & Medicine, 10*(2), 393–399. http://search.ebscohost.com/login.aspx?direct=true&db=s3h&AN=67432854&site=ehost-live

Gholve, P., Scher, D. M., Khakharia, S., Widmann, R. F., & Green, D. W. (2007). Osgood Schlatter syndrome. *Current Opinion in Pediatrics, 19*(1), 44–50. https://doi.org/10.1097/MOP.0b013e328013dbea

Ghosh, A. K. (2002). Factors Involved in Regulation of Type I Collagen Gene Expression: Implication in Fibrosis. *Experimental Biology and Medicine, 227*(5), 301–314. https://doi.org/10.1172/JCI113959

Gibbon, A., Hobbs, H., van der Merwe, W., Raleigh, S. M., Cook, J., Handley, C. J., Posthumus, M., Collins, M., & September, A. V. (2016). The MMP3 gene in musculoskeletal soft tissue injury risk profiling: A study in two independent sample groups. *Journal of Sports Sciences, 35*(7), 655–662. https://doi.org/10.1080/02640414.2016.1183806

Gibbon, A., Raleigh, S. M., Ribbans, W. J., Posthumus, M., Collins, M., & September, A. V. (2020). Functional COL1A1 variants are associated with the risk of acute musculoskeletal soft tissue injuries. *Journal of Orthopaedic Research, 38*(10), 2290–2298. https://doi.org/10.1002/jor.24621

Gibson, W. T. (2009). Genetic association studies for complex traits: relevance for the sports medicine practitioner. *British Journal of Sports Medicine, 43*(5), 314–316. https://doi.org/10.1136/bjsm.2008.052191

Gibson, W. T. (2016). Core Concepts in Human Genetics: Understanding the Complex Phenotype of Sport Performance and Susceptibility to Sport Injury. *Medicine and Sport Science, 61*, 1–14. https://doi.org/10.1159/000445237

Ginevičienė, V., Jakaitiene, A., Aksenov, M., Aksenova, A., Druzhevskaya, A., Astratenkova, I., Egorova, E., Gabdrakhmanova, L., Tubelis, L., Kucinskas, V., & Utkus, A. (2016). Association analysis of ACE, ACTN3 and PPARGC1A gene polymorphisms in two cohorts of European strength and power athletes. *Biology of Sport, 33*(3), 199–206. https://doi.org/10.5604/20831862.1201051

Ginevičienė, V., Jakaitiene, A., Tubelis, L., & Kucinskas, V. (2014). Variation in the ACE, PPARGC1A and PPARA genes in Lithuanian football players. *Eur J Sport Sci, 14 Suppl 1*(June), S289-95. https://doi.org/10.1080/17461391.2012.691117

Gómez-Gallego, F., Santiago, C., González-Freire, M., Muniesa, C., Fernández del Valle, M., Pérez, M., Foster, C., & Lucia, A. (2009). Endurance Performance: Genes or Gene Combinations? *International Journal of Sports Medicine, 30*(01), 66–72. https://doi.org/10.1055/s-2008-1038677

Gonzalez, A. M., Hoffman, J. R., Stout, J. R., Fukuda, D. H., & Willoughby, D. S. (2016). Intramuscular Anabolic Signaling and Endocrine Response Following Resistance Exercise: Implications for Muscle Hypertrophy. *Sports Medicine, 46*(5), 671–685. https://doi.org/10.1007/s40279-015-0450-4

Goodlin, G. T., Roos, A. K., Roos, T. R., Hawkins, C., Beache, S., Baur, S., & Kim, S. K. (2015). Applying personal genetic data to injury risk assessment in athletes. *PLoS ONE, 10*(4), 1–13. https://doi.org/10.1371/journal.pone.0122676

Goto, H., Morris, J. G., & Nevill, M. E. (2015). Motion analysis of U11 to U16 elite English Premier League Academy players. *Journal of Sports Sciences, 33*(12), 1248–1258. https://doi.org/10.1080/02640414.2014.999700

Granados, A., Gebremariam, A., & Lee, J. M. (2015). Relationship Between Timing of Peak Height Velocity and Pubertal Staging in Boys and Girls. *Journal of Clinical Research in Pediatric Endocrinology, 7*(3), 235–237. https://doi.org/10.4274/jcrpe.2007

Granger, E., Di Nardo, F., Harrison, A., Patterson, L., Holmes, R., & Verma, A. (2017). A systematic review of the relationship of physical activity and health status in adolescents. *European Journal of Public Health, 27*(suppl_2), 100–106. https://doi.org/10.1093/eurpub/ckw187

Grealy, R., Herruer, J., Smith, C. L. E., Hiller, D., Haseler, L. J., & Griffiths, L. R. (2015). Evaluation of a 7-gene genetic profile for athletic endurance phenotype in ironman championship triathletes. *PLoS ONE, 10*(12), 1–20. https://doi.org/10.1371/journal.pone.0145171

Grealy, R., Smith, C. L. E. E., Chen, T., Hiller, D., Haseler, L. J., & Griffiths, L. R. (2013). The genetics of endurance: Frequency of the ACTN3 R577X variant in Ironman World Championship athletes. *Journal of Science and Medicine in Sport, 16*(4), 365–371. https://doi.org/10.1016/j.jsams.2012.08.013

Guilherme, J. P. L. F., Bertuzzi, R., Lima-Silva, A. E., Pereira, A. da C., & Lancha Junior, A. H. (2018). Analysis of sports-relevant polymorphisms in a large Brazilian cohort of top-level athletes. *Annals of Human Genetics, 82*(5), 254–264. https://doi.org/10.1111/ahg.12248

Guldhammer, C., Rathleff, M. S., Jensen, H. P., & Holden, S. (2019). Long-term Prognosis and Impact of Osgood-Schlatter Disease 4 Years After Diagnosis: A Retrospective Study. *Orthopaedic Journal of Sports Medicine, 7*(10), 232596711987813. https://doi.org/10.1177/2325967119878136

Hagberg, J. M., Ferrell, R. E., McCole, S. D., Wilund, K. R., & Moore, G. E. (1998). VO2 max is associated with ACE genotype in postmenopausal women. *Journal of Applied Physiology (Bethesda, Md. : 1985), 85*(5), 1842–1846. https://doi.org/10.1152/jappl.1998.85.5.1842

Hägglund, M., Waldén, M., Magnusson, H., Kristenson, K., Bengtsson, H., & Ekstrand, J. (2013). Injuries affect team performance negatively in professional football: an 11-year follow-up of the UEFA Champions League injury study. *British Journal of Sports Medicine, 47*(12), 738–742. https://doi.org/10.1136/bjsports-2013-092215

Hakim, A. J., Cherkas, L. F., Spector, T. D., & MacGregor, A. J. (2003). Genetic associations between frozen shoulder and tennis elbow: a female twin study. *Rheumatology (Oxford, England), 42*(6), 739–742. https://doi.org/10.1093/rheumatology/keg159

Hall, E. C. R., Baumert, P., Larruskain, J., Gil, S. M., Lekue, J. A., Rienzi, E., Moreno, S., Tannure, M., Murtagh, C. F., Ade, J. D., Squires, P., Orme, P., Anderson, L., Brownlee, T. E., Whitworth-Turner, C. M., Morton, J. P., Drust, B., Williams, A. G., & Erskine, R. M. (2022). The genetic association with injury risk in male academy soccer players depends on maturity status. *Scandinavian Journal of Medicine & Science in Sports, 32*(2), 338–350. https://doi.org/10.1111/sms.14077

Hall, E. C. R., Larruskain, J., Gil, S. M., Lekue, J. A., Baumert, P., Rienzi, E., Moreno, S., Tannure, M., Murtagh, C. F., Ade, J. D., Squires, P., Orme, P., Anderson, L., Whitworth-Turner, C. M., Morton, J. P., Drust, B., Williams, A. G., & Erskine, R. M. (2020). An injury audit in high-level male youth soccer players from English, Spanish, Uruguayan and Brazilian academies. *Physical Therapy in Sport, 44*, 53–60. https://doi.org/10.1016/j.ptsp.2020.04.033

Halson, S. L. (2014). Monitoring Training Load to Understand Fatigue in Athletes. *Sports Medicine, 44*, 139–147. https://doi.org/10.1007/s40279-014-0253-z

Hannan, M. T., Litman, H. J., Araujo, A. B., McLennan, C. E., McLean, R. R., McKinlay, J. B., Chen, T. C., & Holick, M. F. (2008). Serum 25-Hydroxyvitamin D and Bone Mineral Density in a Racially and Ethnically Diverse Group of Men. *The Journal of Clinical Endocrinology & Metabolism, 93*(1), 40–46. https://doi.org/10.1210/jc.2007-1217

Hansen, L., Bangsbo, J., Twisk, J., & Klausen, K. (1999). Development of muscle strength in relation to training level and testosterone in young male soccer players. *Journal of Applied Physiology, 87*(3), 1141–1147. https://doi.org/10.1152/jappl.1999.87.3.1141

Hanson, E. D., Ludlow, A. T., Sheaff, A. K., Park, J. J., & Roth, S. M. (2010). ACTN3 Genotype Does not Influence Muscle Power. *International Journal of Sports Medicine, 31*(11), 834–838. https://doi.org/10.1055/s-0030-1263116

Hautala, A. J., Kiviniemi, A. M., Mäkikallio, T. H., Kinnunen, H., Nissilä, S., Huikuri, H. V, & Tulppo, M. P. (2006). Individual differences in the responses to endurance and resistance training. *European Journal of Applied Physiology, 96*(5), 535–542. https://doi.org/10.1007/s00421-005-0116-2

Hebestreit, H., Mimura, K., & Bar-Or, O. (1993). Recovery of muscle power after high-intensity short-term exercise: comparing boys and men. *Journal of Applied Physiology, 74*(6), 2875–2880. https://doi.org/10.1152/jappl.1993.74.6.2875

Hedström, E. M., Svensson, O., Bergström, U., & Michno, P. (2010). Epidemiology of fractures in children and adolescents. *Acta Orthopaedica, 81*(1), 148–153. https://doi.org/10.3109/17453671003628780

Heffernan, S. M., Kilduff, L. P., Erskine, R. M., Day, S. H., Stebbings, G. K., Cook, C. J., Raleigh, S. M., Bennett, M. A., Wang, G., Collins, M., Pitsiladis, Y. P., & Williams, A. G. (2017). COL5A1 gene variants previously associated with reduced soft tissue injury risk are associated with elite athlete status in rugby. *BMC Genomics, 18*(S8), 820. https://doi.org/10.1186/s12864-017-4187-3

Helgerud, J., Engen, L., Wisløff, U., & Hoff, J. (2001). Aerobic endurance training improves soccer performance. *Medicine and Science in Sports and Exercise, 33*(13), 1925–1931. https://doi.org/10.1097/00005768-200111000-00019

Helsen, W. F., Starkes, J. L., & Van Winckel, J. (1998). The Influence of Relative Age on Success and Dropout in Male Soccer Players. *American Journal of Human Biology, 10*, 791–798. https://doi.org/10.1002/(SICI)1520-6300(1998)10:6<791::AID-AJHB10>3.0.CO;2-1

Herbert, A. J., Williams, A. G., Hennis, P. J., Erskine, R. M., Sale, C., Day, S. H., & Stebbings, G. K. (2018). The interactions of physical activity, exercise and genetics and their associations with bone mineral density: implications for injury risk in elite athletes. *European Journal of Applied Physiology, 0*(3), 0. https://doi.org/10.1007/s00421-018-4007-8

Hermanussen, M. (1998). The analysis of short-term growth. *In Hormone Research* (Vol. 49, Issue 2, pp. 53–64). https://doi.org/10.1159/000023127

Hernandez, C. J., & van der Meulen, M. C. (2017). Understanding Bone Strength Is Not Enough. Journal of Bone and Mineral Research : *The Official Journal of the American Society for Bone and Mineral Research, 32*(6), 1157–1162. https://doi.org/10.1002/jbmr.3078

Hernández, D., de la Rosa, A., Barragán, A., Barrios, Y., Salido, E., Torres, A., Martín, B., Laynez, I., Duque, A., De Vera, A., Lorenzo, V., & González, A. (2003). The ACE/DD genotype is associated with the extent of

exercise-induced left ventricular growth in endurance athletes. *Journal of the American College of Cardiology, 42*(3), 527–532. https://doi.org/10.1016/S0735-1097(03)00642-9

Hill, A. B. (1965). The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine, 58*(5), 295–300. https://doi.org/10.1177/003591576505800503

Hill, M., Scott, S., Malina, R. M., McGee, D., & Cumming, S. P. (2020). Relative age and maturation selection biases in academy football. *Journal of Sports Sciences, 38*(11–12), 1359–1367. https://doi.org/10.1080/02640414.2019.1649524

Holden, S., & Rathleff, M. S. (2020). Separating the myths from facts: time to take another look at Osgood-Schlatter 'disease.' *British Journal of Sports Medicine, 54*(14), 824–825. https://doi.org/10.1136/bjsports-2019-101888

Horita, N., & Kaneko, T. (2015). Genetic model selection for a case–control study and a meta-analysis. *Meta Gene, 5*, 1–8. https://doi.org/10.1016/j.mgene.2015.04.003

Hu, F. B., Goldberg, J., Hedeker, D., Flay, B. R., & Pentz, M. A. (1998). Comparison of Population-Averaged and Subject-Specific Approaches for Analyzing Repeated Binary Outcomes. *American Journal of Epidemiology, 147*(7), 694–703. https://doi.org/10.1093/oxfordjournals.aje.a009511

Hubal, M. J., Gordish-Dressman, H., Thompson, P. D., Price, T. B., Hoffman, E. P., Angelopoulos, T. J., Gordon, P. M., Moyna, N. M., Pescatello, L. S., Visich, P. S., Zoeller, R. F., Seip, R. L., Clarkson, P. M., Thompson D., P., Price, T. B., Hoffman, E. P., Angelopoulos, T. J., Gordon, P. M., Moyna, N. M., … Clarkson, P. M. (2005). Variability in muscle size and strength gain after unilateral resistance training. *Medicine and Science in Sports and Exercise, 37*(6), 964–972. https://doi.org/10.1249.01.mss.0000170469.90461.5f

Hughes, T., Riley, R. D., Callaghan, M. J., & Sergeant, J. C. (2020). The Value of Preseason Screening for Injury Prediction: The Development and Internal Validation of a Multivariable Prognostic Model to Predict Indirect Muscle Injury Risk in Elite Football (Soccer) Players. *Sports Medicine - Open, 6*(1), 22. https://doi.org/10.1186/s40798-020-00249-8

Hughes, T., Sergeant, J. C., Parkes, M. J., & Callaghan, M. J. (2017). Prognostic factors for specific lower extremity and spinal musculoskeletal injuries identified through medical screening and training load monitoring in professional football (soccer): a systematic review. *BMJ Open Sport & Exercise Medicine, 3*(1), e000263. https://doi.org/10.1136/bmjsem-2017-000263

Hughes, T., Sergeant, J. C., van der Windt, D. A., Riley, R., & Callaghan, M. J. (2018). Periodic Health Examination and Injury Prediction in Professional Football (Soccer): Theoretically, the Prognosis is Good. *Sports Medicine, 48*(11), 2443–2448. https://doi.org/10.1007/s40279-018-0928-y

Hulin, B. T., Gabbett, T. J., Blanch, P., Chapman, P., Bailey, D., & Orchard, J. W. (2014). Spikes in acute workload are associated with increased injury risk in elite cricket fast bowlers. *British Journal of Sports Medicine, 48*(8), 708–712. https://doi.org/10.1136/bjsports-2013-092524

Hulin, B. T., Gabbett, T. J., Lawson, D. W., Caputi, P., & Sampson, J. A. (2016). The acute:chronic workload ratio predicts injury: high chronic workload may decrease injury risk in elite rugby league players. *British Journal of Sports Medicine, 50*(4), 231–236. https://doi.org/10.1136/bjsports-2015-094817

Impellizzeri, F. M., Woocock, S., McCall, A., Ward, P., & Coutts, A. J. (2019). The acute-chronic workload ratio-injury figure and its "sweet spot" are flawed. https://doi.org/10.31236/osf.io/gs8yu

Impellizzeri, F. M., Tenan, M. S., Kempton, T., Novak, A., & Coutts, A. J. (2020). Acute:Chronic Workload Ratio: Conceptual Issues and Fundamental Pitfalls. International Journal of Sports Physiology and Performance, 15(6), 907–913. https://doi.org/10.1123/ijspp.2019-0864

International HapMap Consortium. (2005). A haplotype map of the human genome. *Nature, 437*(7063), 1299–1320. https://doi.org/10.1038/nature04226

Ioannidis, J. P. A., Ntzani, E. E., Trikalinos, T. A., & Contopoulos-Ioannidis, D. G. (2001). Replication validity of genetic association studies. *Nature Genetics, 29*, 306–309. https://doi.org/10.1038/ng749

Ioannidis, J. P. A., Trikalinos, T. A., Ntzani, E. E., & Contopoulos-Ioannidis, D. G. (2003). Genetic associations in large versus small studies: An empirical assessment. *Lancet, 361*(9357), 567–571. https://doi.org/10.1016/S0140-6736(03)12516-0

Iwao-Koizumi, K., Ota, T., Hayashida, M., Yonetani, Y., Nakata, K., Kinoshita, K., Koyanagi, Y., & Murata, S. (2014). The ACTN3 Gene is a Potential Biomarker for the Risk of Non-Contact Sports Injury in Female Athletes. *Journal of Molecular Biomarkers & Diagnosis, s6*. https://doi.org/10.4172/2155-9929.S6-002

Jackson, R. C., & Comber, G. (2020). Hill on a mountaintop: A longitudinal and cross-sectional analysis of the relative age effect in competitive youth football. *Journal of Sports Sciences, 38*(11–12), 1352–1358. https://doi.org/10.1080/02640414.2019.1706830

Jackson, S. E., & Chester, J. D. (2015). Personalised cancer medicine. *International Journal of Cancer, 137*(2), 262–266. https://doi.org/10.1002/ijc.28940

Jakobsen, J. R., & Krogsgaard, M. R. (2021). The Myotendinous Junction—A Vulnerable Companion in Sports. A Narrative Review. *Frontiers in Physiology, 12*(March), 1–13. https://doi.org/10.3389/fphys.2021.635561

Ji, G.-R., Yao, M., Sun, C.-Y., Li, Z.-H., & Han, Z. (2010). BsmI, TaqI, ApaI and FokI polymorphisms in the vitamin D receptor (VDR) gene and risk of fracture in Caucasians: a meta-analysis. *Bone, 47*(3), 681–686. https://doi.org/10.1016/j.bone.2010.06.024

Jin, H., Stewart, T. L., Hof, R. V., Reid, D. M., Aspden, R. M., & Ralston, S. (2009). A Rare Haplotype in the Upstream Regulatory Region of COL1A1 Is Associated With Reduced Bone Quality and Hip Fracture. *Journal of Bone and Mineral Research, 24*(3), 448–454. https://doi.org/10.1359/jbmr.081111

Jo, B.-S., & Choi, S. S. (2015). Introns: The Functional Benefits of Introns in Genomes. *Genomics & Informatics, 13*(4), 112–118. https://doi.org/10.5808/GI.2015.13.4.112

John, C., Rahlf, A. L., Hamacher, D., & Zech, A. (2019). Influence of biological maturity on static and dynamic postural control among male youth soccer players. *Gait & Posture, 68*, 18–22. https://doi.org/10.1016/j.gaitpost.2018.10.036

Johnson, D. M., Williams, S., Bradley, B., Sayer, S., Murray Fisher, J., & Cumming, S. (2020). Growing pains: Maturity associated variation in injury risk in academy football. *European Journal of Sport Science, 20*(4), 544–552. https://doi.org/10.1080/17461391.2019.1633416

Jones, N., Kiely, J., Suraci, B., Collins, D. J., de Lorenzo, D., Pickering, C., Grimaldi, K. A., Lorenzo, D. D., Pickering, C., & Grimaldi, K. A. (2016). A genetic-based algorithm for personalized resistance training. *Biology of Sport, 33*(2), 117–126. https://doi.org/10.5604/20831862.1198210

Jones, S., Almousa, S., Gibb, A., Allamby, N., Mullen, R., Andersen, T. E., & Williams, M. (2019). Injury Incidence, Prevalence and Severity in High-Level Male Youth Football: A Systematic Review. Sports Medicine, 49(12), 1879–1899. https://doi.org/10.1007/s40279-019-01169-8

Juffer, P., Furrer, R., González-Freire, M., Santiago, C., Verde, Z., Serratosa, L., Morate, F., Rubio, J., Martin, M., Ruiz, J. R., Arenas, J., Gómez-Gallego, F., & Lucia, A. (2009). Genotype Distributions in Top-level Soccer Players: A Role for ACE ? *International Journal of Sports Medicine, 30*(05), 387–392. https://doi.org/10.1055/s-0028-1105931

Jurutka, P. W., Remus, L. S., Whitfield, G. K., Thompson, P. D., Hsieh, J.-C., Zitzer, H., Tavakkoli, P., Galligan, M. A., Dang, H. T. L., Haussler, C. A., & Haussler, M. R. (2000). The Polymorphic N Terminus in Human Vitamin D Receptor Isoforms Influences Transcriptional Activity by Modulating Interaction with Transcription Factor IIB. *Molecular Endocrinology, 14*(3), 401–420. https://doi.org/10.1210/mend.14.3.0435

Kadler, K. E., Baldock, C., Bella, J., & Boot-Handford, R. P. (2007). Collagens at a glance. *Journal of Cell Science, 120*(12), 1955–1958. https://doi.org/10.1242/jcs.03453

Kalkhoven, J. T. (2021). Training load and injury : speculation is not knowledge. *The Sport and Exercise Scientist, 68*, 20–21. https://www.bases.org.uk/imgs/8931_bas_bases_tses_summer_2021_online_real_world794.pdf

Kalkhoven, J. T., Watsford, M. L., Coutts, A. J., Edwards, W. B., & Impellizzeri, F. M. (2021). Training Load and Injury: Causal Pathways and Future Directions. *Sports Medicine, 51*(6), 1137–1150. https://doi.org/10.1007/s40279-020-01413-6

Kalkhoven, J. T., Watsford, M. L., & Impellizzeri, F. M. (2020). A conceptual model and detailed framework for stress-related, strain-related, and overuse athletic injury. *Journal of Science and Medicine in Sport, 23*(8), 726–734. https://doi.org/10.1016/j.jsams.2020.02.002

Kasikcioglu, E., Kayserilioglu, A., Ciloglu, F., Akhan, H., Oflaz, H., Yildiz, S., & Peker, I. (2004). Angiotensin-converting enzyme gene polymorphism, left ventricular remodeling, and exercise capacity in strength-trained athletes. *Heart and Vessels, 19*(6), 287–293. https://doi.org/10.1007/s00380-004-0783-7

Kaufman, K., Brodine, S., & Shaffer, R. (2000). Military training-related injuries Surveillance, research, and prevention. *American Journal of Preventive Medicine, 18*(1), 54–63. https://doi.org/10.1016/S0749-3797(00)00114-8

Kaya, D. O., Toprak, U., Baltaci, G., Yosmaoglu, B., & Ozer, H. (2013). Long-term functional and sonographic outcomes in Osgood–Schlatter disease. *Knee Surgery, Sports Traumatology, Arthroscopy, 21*(5), 1131–1139. https://doi.org/10.1007/s00167-012-2116-1

Keen, R. W., Woodford-Richens, K. L., Grant, S. F. A., Ralston, S. H., Lanchbury, J. S., & Spector, T. D. (1999). Association of polymorphism at the type I collagen (COL1A1) locus with reduced bone mineral density, increased fracture risk, and increased collagen turnover. *Arthritis & Rheumatism, 42*(2), 285–290. https://doi.org/10.1002/1529-0131(199902)42:2<285::AID-ANR10>3.0.CO;2-3

Kelly, P. M., & Diméglio, A. (2008). Lower-limb growth: How predictable are predictions? *Journal of Children's Orthopaedics, 2*(6), 407–415. https://doi.org/10.1007/s11832-008-0119-8

Kemp, J. P., Morris, J. A., Medina-Gomez, C., Forgetta, V., Warrington, N. M., Youlten, S. E., Zheng, J., Gregson, C. L., Grundberg, E., Trajanoska, K., Logan, J. G., Pollard, A. S., Sparkes, P. C., Ghirardello, E. J., Allen, R., Leitch, V. D., Butterfield, N. C., Komla-Ebri, D., Adoum, A. T., … Evans, D. M. (2017). Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nature Genetics, 49*(10), 1468–1475. https://doi.org/10.1038/ng.3949

Kemper, G., Van Der Sluis, A., Brink, M., Visscher, C., Frencken, W., & Elferink-Gemser, M. (2015). Anthropometric Injury Risk Factors in Elite-standard Youth Soccer. *International Journal of Sports Medicine, 36*(13), 1112–1117. https://doi.org/10.1055/s-0035-1555778

Kerssemakers, S. P., Fotiadou, A. N., de Jonge, M. C., Karantanas, A. H., & Maas, M. (2009). Sport injuries in the paediatric and adolescent patient: a growing problem. *Pediatric Radiology, 39*(5), 471–484. https://doi.org/10.1007/s00247-009-1191-z

Khamis, H. J., & Roche, A. F. (1994). Predicting Adult Stature Without Using Skeletal Age: Khamis-Roche Method. *Pediatrics, 94*(4), 504–507. https://doi.org/10.1016/S0740-624X(98)90020-X

Khan, K. M., McKay, H. A., Haapasalo, H., Bennell, K. L., Forwood, M. R., Kannus, P., & Wark, J. D. (2000). Does childhood and adolescence provide a unique opportunity for exercise to strengthen the skeleton? *Journal of Science and Medicine in Sport, 3*(2), 150–164. https://doi.org/10.1016/S1440-2440(00)80077-8

Khoschnau, S., Melhus, H., Jacobson, A., Rahme, H., Bengtsson, H., Ribom, E., Grundberg, E. E., Mallmin, H., Michaëlsson, K., Melhus, H., Jacobson, A., Rahme, H., Bengtsson, H., Ribom, E., Grundberg, E. E., Mallmin, H., Michaëlsson, K., Melhus, H., Jacobson, A., … Michaelsson, K. (2008). Type I Collagen a1 Sp1 Polymorphism and the Risk of Cruciate Ligament Ruptures or Shoulder Dislocations. *The American Journal of Sports Medicine, 36*(12), 2432–2436. https://doi.org/10.1177/0363546508320805

Khosla, S., Oursler, M. J., & Monroe, D. G. (2012). Estrogen and the skeleton. *Trends in Endocrinology and Metabolism, 23*(11), 576–581. https://doi.org/10.1016/j.tem.2012.03.008

Kikuchi, N., Min, S., Ueda, D. A. I., Igawa, S., & Nakazato, K. (2012). Higher Frequency of the ACTN3 R Allele + ACE DD Genotype in Japanese Elite Wrestlers. *Journal of Strength and Conditioning Research, 26*(12), 3275–3280. https://doi.org/10.1519/JSC.0b013e318273679d

Kikuchi, N., Miyamoto-Mikami, E., Murakami, H., Nakamura, T., Min, S.-K. K., Mizuno, M., Naito, H., Miyachi, M., Nakazato, K., & Fuku, N. (2016). ACTN3 R577X genotype and athletic performance in a large cohort of Japanese athletes. *European Journal of Sport Science, 16*(6), 694–701. https://doi.org/10.1080/17461391.2015.1071879

Kikuchi, N., Ueda, D., Min, S., Nakazato, K., & Igawa, S. (2013). The ACTN3 XX genotype's underrepresentation in Japanese elite wrestlers. *International Journal of Sports Physiology and Performance, 8*(1), 57–61. https://doi.org/10.1123/ijspp.8.1.57

Kim, H., Song, K.-H., & Kim, C.-H. (2014a). The ACTN3 R577X variant in sprint and strength performance. *Journal of Exercise Nutrition and Biochemistry, 18*(4), 347–353. https://doi.org/10.5717/jenb.2014.18.4.347

Kim, J. H., Jung, E. S., Kim, C.-H., Youn, H., & Kim, H. R. (2014b). Genetic associations of body composition, flexibility and injury risk with ACE, ACTN3 and COL5A1 polymorphisms in Korean ballerinas. *Journal of Exercise Nutrition and Biochemistry, 18*(2), 205–214. https://doi.org/10.5717/jenb.2014.18.2.205

Kim, S. K., Kleimeyer, J. P., Ahmed, M. A., Avins, A. L., Fredericson, M., Dragoo, J. L., & Ioannidis, J. P. a. (2017a). A Genetic Marker Associated with Shoulder Dislocation. *International Journal of Sports Medicine, 38*(7), 508–514. https://doi.org/10.1055/s-0043-106190

Kim, S. K., Roos, T. R., Roos, A. K., Kleimeyer, J. P., Ahmed, M. A., Goodlin, G. T., Fredericson, M., Ioannidis, J. P. A., Avins, A. L., & Dragoo, J. L. (2017b). Genome-wide association screens for achilles tendon and ACL tears and tendinopathy. *PLoS ONE, 12*(3), 1–16. https://doi.org/10.1371/journal.pone.0170422

Kohrt, W. M., Bloomfield, S. A., Little, K. D., Nelson, M. E., & Yingling, V. R. (2004). Physical activity and bone health. *Medicine and Science in Sports and Exercise, 36*(11), 1985–1996. https://doi.org/10.1249/01.MSS.0000142662.21767.58

Korvala, J., Hartikka, H., Pihlajamäki, H., Solovieva, S., Ruohola, J.-P., Sahi, T., Barral, S., Ott, J., Ala-Kokko, L., & Männikkö, M. (2010). Genetic predisposition for femoral neck stress fractures in military conscripts. *BMC Genetics, 11*(1), 95. https://doi.org/10.1186/1471-2156-11-95

Kozlovskaia, M., Vlahovich, N., Ashton, K. J., & Hughes, D. C. (2017). Biomedical Risk Factors of Achilles Tendinopathy in Physically Active People: a Systematic Review. *Sports Medicine - Open, 3*(1), 20. https://doi.org/10.1186/s40798-017-0087-y

Krabbe, S., Transbøl, I., & Christiansen, C. (1982). Bone mineral homeostasis, bone growth, and mineralisation during years of pubertal growth: a unifying concept. *Archives of Disease in Childhood, 57*(5), 359–363. https://doi.org/10.1136/adc.57.5.359

Kraemer, W. J., & Ratamess, N. A. (2004). Fundamentals of Resistance Training: Progression and Exercise Prescription. *In Medicine and Science in Sports and Exercise* (Vol. 36, Issue 4, pp. 674–688). https://doi.org/10.1249/01.MSS.0000121945.36635.61

Kubo, K., Yata, H., & Tsunoda, N. (2013). Effect of gene polymorphisms on the mechanical properties of human tendon structures. *SpringerPlus, 2*(1), 343. https://doi.org/10.1186/2193-1801-2-343

Kumagai, H., Miyamoto-Mikami, E., Hirata, K., Kikuchi, N., Kamiya, N., Hoshikawa, S., Zempo, H., Naito, H., Miyamoto, N., & Fuku, N. (2018). ESR1 rs2234693 Polymorphism Is Associated with Muscle Injury and Muscle Stiffness. *Medicine & Science in Sports & Exercise, 51*(1), 1. https://doi.org/10.1249/MSS.0000000000001750

Kurt, O., Yilmaz-Aydogan, H., Uyar, M., Isbir, T., Seyhan, M. F., & Can, A. (2012). Evaluation of ERα and VDR gene polymorphisms in relation to bone mineral density in Turkish postmenopausal women. *Molecular Biology Reports, 39*(6), 6723–6730. https://doi.org/10.1007/s11033-012-1496-0

Laguette, M.-J., Abrahams, Y., Prince, S., & Collins, M. (2011). Sequence variants within the 3'-UTR of the COL5A1 gene alters mRNA stability: Implications for musculoskeletal soft tissue injuries. *Matrix Biology, 30*(5–6), 338–345. https://doi.org/10.1016/j.matbio.2011.05.001

Lahti, J., Mendiguchia, J., Ahtiainen, J., Anula, L., Kononen, T., Kujala, M., Matinlauri, A., Peltonen, V., Thibault, M., Toivonen, R., Edouard, P., & Morin, J. B. (2020). Multifactorial individualised programme for hamstring muscle injury risk reduction in professional football: protocol for a prospective cohort study. *BMJ Open Sport & Exercise Medicine, 6*(1), e000758. https://doi.org/10.1136/bmjsem-2020-000758

Lampl, M., & Johnson, M. L. (1993). A case study of daily growth during adolescence: a single spurt or changes in the dynamics of saltatory growth? *Annals of Human Biology, 20*(6), 595–603. https://doi.org/10.1080/03014469300003002

Lander, E., & Schork, N. (1994). Genetic dissection of complex traits. *Science, 265*(5181), 2037–2048. https://doi.org/10.1126/science.8091226

Larruskain, J., Celorrio, D., Barrio, I., Odriozola, A., Gil, S. M., Fernandez-Lopez, J. R., Nozal, R., Ortuzar, I., Lekue, J. A., & Aznar, J. M. (2018). Genetic Variants and Hamstring Injury in Soccer: An Association and Validation Study. *Medicine and Science in Sports and Exercise, 50*(2), 361–368. https://doi.org/10.1249/MSS.0000000000001434

Larruskain, J., Lekue, J. A., Martin-Garetxana, I., Barrio, I., McCall, A., & Gil, S. M. (2021). Injuries are negatively associated with player progression in an elite football academy. *Science and Medicine in Football, 00*(00), 1–10. https://doi.org/10.1080/24733938.2021.1943756

Launay, F. (2015). Sports-related overuse injuries in children. *Orthopaedics & Traumatology: Surgery & Research, 101*(1), S139–S147. https://doi.org/10.1016/j.otsr.2014.06.030

Le Gall, F., Carling, C., & Reilly, T. (2007). Biological maturity and injury in elite youth football. *Scandinavian Journal of Medicine and Science in Sports, 17*(5), 564–572. https://doi.org/10.1111/j.1600-0838.2006.00594.x

Le Gall, F., Carling, C., Reilly, T., Vandewalle, H., Church, J., & Rochcongar, P. (2006). Incidence of Injuries in Elite French Youth Soccer Players: A 10-Season Study. *American Journal of Sports Medicine, 34*(6), 928–938. https://doi.org/10.1177/0363546505283271

Lee, Y. H. (2015). Meta-Analysis of Genetic Association Studies. *Annals of Laboratory Medicine, 35*(3), 283–287. https://doi.org/10.3343/alm.2015.35.3.283

Lemes, I. R., Pinto, R. Z., Lage, V. N., Roch, B. A. B., Verhagen, E., Bolling, C., Aquino, C. F., Fonseca, S. T., & Souza, T. R. (2021). Do exercise-based prevention programmes reduce non-contact musculoskeletal injuries in football (soccer)? A systematic review and meta-analysis with 13 355 athletes and more than 1 million exposure hours. *British Journal of Sports Medicine, 55*(20), 1170–1178. https://doi.org/10.1136/bjsports-2020-103683

Levinger, I., Yan, X., Bishop, D., Houweling, P. J., Papadimitriou, I., Munson, F., Byrnes, E., Vicari, D., Brennan-Speranza, T. C., & Eynon, N. (2017). The influence of α-actinin-3 deficiency on bone remodelling markers in young men. *Bone, 98*, 26–30. https://doi.org/10.1016/j.bone.2017.02.010

Levy, S., Sutton, G., Ng, P., Feuk, L., Halpern, A., Walenz, B., Axelrod, N., Huang, J., Kirkness, E., Denisov, G., Lin, Y., MacDonald, J., Pang, A. W. C., Shago, M., Stockwell, T., Tsiamouri, A., Bafna, V., Bansal, V., Kravitz, S., … Venter, C. (2007). The diploid genome sequence of an individual human. *PLoS Biology, 5*(10), e254. https://doi.org/10.1371/journal.pbio.0050254

Lewontin, R. C. (1964). The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. *Genetics, 49*(1), 49–67. http://www.ncbi.nlm.nih.gov/pubmed/17248194%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1210557

Li, Y. C., Wang, L. Q., Yi, L. Y., Liu, J. H., Hu, Y., Lu, Y. F., & Wang, M. (2017). ACTN3 R577X genotype and performance of elite middle-long distance swimmers in China. *Biology of Sport, 1*(1), 39–43. https://doi.org/10.5114/biolsport.2017.63731

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ, 339*(jul21 1), b2700–b2700. https://doi.org/10.1136/bmj.b2700

Light, N., Johnson, A., Williams, S., Smith, N., Hale, B., & Thorborg, K. (2021). Injuries in youth football and the relationship to player maturation: An analysis of time-loss injuries during four seasons in an English elite male football academy. *Scandinavian Journal of Medicine & Science in Sports, August* 2020, sms.13933. https://doi.org/10.1111/sms.13933

Lim, T., Santiago, C., Pareja-Galeano, H., Iturriaga, T., Sosa-Pedreschi, A., Fuku, N., Pérez-Ruiz, M., & Yvert, T. (2021). Genetic variations associated with non-contact muscle injuries in sport: A systematic review. *Scandinavian Journal of Medicine & Science in Sports, May*, sms.14020. https://doi.org/10.1111/sms.14020

Lima, R., Leite, T., Pereira, R., Rabelo, H., Roth, S., & Oliveira, R. (2011). ACE and ACTN3 Genotypes in Older Women: Muscular Phenotypes. *International Journal of Sports Medicine, 32*(01), 66–72. https://doi.org/10.1055/s-0030-1267229

Little, J., Higgins, J. P. T., Ioannidis, J. P. A., Moher, D., Gagnon, F., von Elm, E., Khoury, M. J., Cohen, B., Davey-Smith, G., Grimshaw, J., Scheet, P., Gwinn, M., Williamson, R. E., Zou, G. Y., Hutchings, K., Johnson, C. Y., Tait, V., Wiens, M., Golding, J., … Birkett, N. (2009). STrengthening the REporting of Genetic Association Studies (STREGA)-an extension of the STROBE statement. *Genetic Epidemiology, 33*(7), 581–598. https://doi.org/10.1002/gepi.20410

Liu, J., Cai, W., Zhang, H., He, C., & Deng, L. (2013). Rs143383 in the growth differentiation factor 5 (GDF5) gene significantly associated with osteoarthritis (OA)-a comprehensive meta-analysis. *International Journal of Medical Sciences, 10*(3), 312–319. https://doi.org/10.7150/ijms.5455

Lloyd, R. S., Cronin, J. B., Faigenbaum, A. D., Haff, G. G., Howard, R., Kraemer, W. J., Micheli, L. J., Myer, G. D., & Oliver, J. L. (2016). National Strength and Conditioning Association Position Statement on Long-Term Athletic Development. *Journal of Strength and Conditioning Research, 30*(6), 1491–1509. https://doi.org/10.1519/JSC.0000000000001387

Lloyd, R. S., Oliver, J. L., Faigenbaum, A. D., Howard, R., De Ste Croix, M. B. A., Williams, C. A., Best, T. M., Alvar, B. A., Micheli, L. J., Thomas, D. P., Hatfield, D. L., Cronin, J. B., & Myer, G. D. (2015). Long-Term Athletic Development- Part 1. *Journal of Strength and Conditioning Research, 29*(5), 1439–1450. https://doi.org/10.1519/JSC.0000000000000756

Löffek, S., Schilling, O., & Franzke, C.-W. (2011). Series "matrix metalloproteinases in lung health and disease": Biological role of matrix metalloproteinases: a critical balance. *The European Respiratory Journal, 38*(1), 191–208. https://doi.org/10.1183/09031936.00146510

Lolli, L., Batterham, A. M., Hawkins, R., Kelly, D. M., Strudwick, A. J., Thorpe, R., Gregson, W., & Atkinson, G. (2019). Mathematical coupling causes spurious correlation within the conventional acute-to-chronic workload ratio calculations. *British Journal of Sports Medicine, 53*(15), 921–922. https://doi.org/10.1136/bjsports-2017-098110

Lubans, D., Richards, J., Hillman, C., Faulkner, G., Beauchamp, M., Nilsson, M., Kelly, P., Smith, J., Raine, L., & Biddle, S. (2016). Physical Activity for Cognitive and Mental Health in Youth: A Systematic Review of Mechanisms. *Pediatrics, 138*(3), e20161642–e20161642. https://doi.org/10.1542/peds.2016-1642

Lucia, A., Gomez-Gallego, F., Santiago, C., Bandres, F., Earnest, C., Rabadan, M., Alonso, J. M., Hoyos, J., Cordova, A., Villa, G., Foster, C., Gómez-Gallego, F., Santiago, C., Bandrés, F., Earnest, C., Rabadán, M., Alonso, J. M., Hoyos, J., Córdova, A., … Foster, C. (2006). ACTN3 Genotype in Professional Endurance Cyclists. *International Journal of Sports Medicine, 27*(11), 880–884. https://doi.org/10.1055/s-2006-923862

Lulińska-Kuklik, E., Rahim, M., Domańska-Senderowska, D., Ficek, K., Michałowska-Sawczyn, M., Moska, W., Kaczmarczyk, M., Brzeziański, M., Brzeziańska-Lasota, E., Cięszczyk, P., & September, A. V. (2018). Interactions Between COL5A1 Gene and Risk of the Anterior Cruciate Ligament Rupture. *Journal of Human Kinetics, 62*(1), 65–71. https://doi.org/10.1515/hukin-2017-0177

Lv, Z.-T., Gao, S.-T., Cheng, P., Liang, S., Yu, S.-Y., Yang, Q., & Chen, A.-M. (2017). Association between polymorphism rs12722 in COL5A1 and musculoskeletal soft tissue injuries: a systematic review and meta-analysis. *Oncotarget, 9*(20), 15365–15374. https://doi.org/10.18632/oncotarget.23805

Ma, F., Yang, Y., Li, X., Zhou, F., Gao, C., Li, M., & Gao, L. (2013). The Association of Sport Performance with ACE and ACTN3 Genetic Polymorphisms: A Systematic Review and Meta-Analysis. *PLoS ONE, 8*(1), e54685. https://doi.org/10.1371/journal.pone.0054685

MacArthur, D. G., & North, K. N. (2004). A gene for speed? The evolution and function of alpha-actinin-3. BioEssays : News and Reviews in Molecular, *Cellular and Developmental Biology, 26*(7), 786–795. https://doi.org/10.1002/bies.20061

MacArthur, D. G., & North, K. N. (2007). ACTN3: A Genetic Influence on Muscle Function and Athletic Performance. *Exerc. Sport Sci. Rev, 35*(1), 30–34. https://doi.org/10.1097/JES.0b013e31802d8874\r00003677-200701000-00007

MacArthur, D. G., Seto, J. T., Raftery, J. M., Quinlan, K. G., Huttley, G. A., Hook, J. W., Lemckert, F. A., Kee, A. J., Edwards, M. R., Berman, Y., Hardeman, E. C., Gunning, P. W., Easteal, S., Yang, N., & North, K. N. (2007). Loss of ACTN3 gene function alters mouse muscle metabolism and shows evidence of positive selection in humans. *Nature Genetics, 39*(10), 1261–1265. https://doi.org/10.1038/ng2122

Mägi, A., Unt, E., Prans, E., Raus, L., Eha, J., Veraksitš, A., Kingo, K., & Kõks, S. (2016). The Association Analysis between ACE and ACTN3 Genes Polymorphisms and Endurance Capacity in Young Cross-Country Skiers: Longitudinal Study. *Journal of Sports Science & Medicine, 15*(2), 287–294. http://search.ebscohost.com/login.aspx?direct=true&db=s3h&AN=115590509&site=ehost-live

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

Magra, M., & Maffulli, N. (2008). Genetic aspects of tendinopathy. *Journal of Science and Medicine in Sport, 11*(3), 243–247. https://doi.org/10.1016/j.jsams.2007.04.007

Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). *Growth, Maturation and Physical Activity*. Human Kinetics.

Malina, R. M., Choh, A. C., Czerwinski, S. A., & Chumlea, W. C. (2016). Validation of Maturity Offset in the Fels Longitudinal Study. *Pediatric Exercise Science, 28*(3), 439–455. https://doi.org/10.1123/pes.2015-0090

Malina, R. M., Coelho E Silva, M. J., Figueiredo, A. J., Carling, C., & Beunen, G. P. (2012). Interrelationships among invasive and non-invasive indicators of biological maturation in adolescent male soccer players. *Journal of Sports Sciences, 30*(15), 1705–1717. https://doi.org/10.1080/02640414.2011.639382

Malina, R. M., Cumming, S. P., Kontos, A. P., Eisenmann, J. C., Ribeiro, B., & Aroso, J. (2005a). Maturity-associated variation in sport-specific skills of youth soccer players aged 13 – 15 years. *Journal of Sports Sciences, 23*(5), 515–522. https://doi.org/10.1080/02640410410001729928

Malina, R. M., Cumming, S. P., Morano, P. J., Barron, M. J., & Miller, S. J. (2005b). Maturity status of youth football players: A noninvasive estimate. *Medicine and Science in Sports and Exercise, 37*(6), 1044–1052. https://doi.org/10.1249/01.mss.0000171622.45134.cb

Malina, R. M., Dompier, T. P., Powell, J. W., Barron, M. J., & Moore, M. T. (2007). Validation of a Noninvasive Maturity Estimate Relative to Skeletal Age in Youth Football Players. *Clinical Journal of Sport Medicine, 17*(5), 362–368. https://doi.org/10.1097/JSM.0b013e31815400f4

Malina, R. M., Eisenmann, J. C., Cumming, S. P., Ribeiro, B., & Aroso, J. (2004). Maturity-associated variation in the growth and functional capacities of youth football (soccer) players 13-15 years. *European Journal of Applied Physiology, 91*(5–6), 555–562. https://doi.org/10.1007/s00421-003-0995-z

Malina, R. M., Ribeiro, B., Aroso, J., & Cumming, S. P. (2007). Characteristics of youth soccer players aged 13-15 years classified by skill level. *British Journal of Sports Medicine, 41*(5), 290–295; discussion 295. https://doi.org/10.1136/bjsm.2006.031294

Malina, R. M., Rogol, A. D., Cumming, S. P., Coelho e Silva, M. J., & Figueiredo, A. J. (2015). Biological maturation of youth athletes: assessment and implications. *British Journal of Sports Medicine, 49*(13), 852–859. https://doi.org/10.1136/bjsports-2015-094623

Malone, S., Hughes, B., Doran, D. A., Collins, K., & Gabbett, T. J. (2019). Can the workload–injury relationship be moderated by improved strength, speed and repeated-sprint qualities? *Journal of Science and Medicine in Sport, 22*(1), 29–34. https://doi.org/10.1016/j.jsams.2018.01.010

Mann, V., Hobson, E. E., Li, B., Stewart, T. L., Grant, S. F. A., Robins, S. P., Aspden, R. M., & Ralston, S. H. (2001). A COL1A1 Sp1 binding site polymorphism predisposes to osteoporotic fracture by affecting bone density and quality. *Journal of Clinical Investigation, 107*(7), 899–907. https://doi.org/10.1172/JCI10347

Mann, V., & Ralston, S. H. (2003). Meta-analysis of COL1A1 Sp1 polymorphism in relation to bone mineral density and osteoporotic fracture. Bone, 32(6), 711–717. https://doi.org/10.1016/S8756-3282(03)00087-5

Marshall, W. A., & Tanner, J. M. (1970). Variations in the Pattern of Pubertal Changes in Boys. *Archives of Disease in Childhood, 45*(239), 13–23. https://doi.org/10.1136/adc.45.239.13

Massidda, M., Bachis, V., Corrias, L., Piras, F., Scorcu, M., & Calò, C. M. (2015a). Influence of the COL5A1 rs12722 on musculoskeletal injuries in professional soccer players. *The Journal of Sports Medicine and Physical Fitness, 55*(11), 1348–1353. http://www.ncbi.nlm.nih.gov/pubmed/25583227

Massidda, M., Corrias, L., Bachis, V., Cugia, P., Piras, F., Scorcu, M., & Calò, C. M. (2015b). Vitamin D receptor gene polymorphisms and musculoskeletal injuries in professional football players. *Experimental and Therapeutic Medicine, 9*(5), 1974–1978. https://doi.org/10.3892/etm.2015.2364

Massidda, M., Corrias, L., Ibba, G., Scorcu, M., Vona, G., & Calò, C. M. (2012). Genetic markers and explosive leg-muscle strength in elite Italian soccer players. *The Journal of Sports Medicine and Physical Fitness, 52*(3), 328–334. http://www.ncbi.nlm.nih.gov/pubmed/22648472

Massidda, M., Miyamoto-Mikami, E., Kumagai, H., Ikeda, H., Shimasaki, Y., Yoshimura, M., Cugia, P., Piras, F., Scorcu, M., Kikuchi, N., Calò, C. M., & Fuku, N. (2020). Association between the ACE I/D polymorphism and muscle injuries in Italian and Japanese elite football players. *Journal of Sports Sciences, 38*(21), 2423–2429. https://doi.org/10.1080/02640414.2020.1787683

Massidda, M., Voisin, S., Culigioni, C., Piras, F., Cugia, P., Yan, X., Eynon, N., & Calò, C. M. (2019). ACTN3 R577X Polymorphism Is Associated With the Incidence and Severity of Injuries in Professional Football Players. *Clinical Journal of Sport Medicine, 29*(1), 57–61. https://doi.org/10.1097/JSM.0000000000000487

Materne, O., Chamari, K., Farooq, A., Weir, A., Hölmich, P., Bahr, R., Greig, M., & McNaughton, L. R. (2020). Injury incidence and burden in a youth elite football academy: a four-season prospective study of 551 players aged from under 9 to under 19 years. *British Journal of Sports Medicine, bjsports-2020-102859*. https://doi.org/10.1136/bjsports-2020-102859

Materne, O., Chamari, K., Farooq, A., Weir, A., Hölmich, P., Bahr, R., Greig, M., & McNaughton, L. R. (2021). Association of Skeletal Maturity and Injury Risk in Elite Youth Soccer Players: A 4-Season Prospective Study With Survival Analysis. *Orthopaedic Journal of Sports Medicine, 9*(3), 232596712199911. https://doi.org/10.1177/2325967121999113

Matthys, S. P., Vaeyens, R., Coelho-E-Silva, M. J., Lenoir, M., & Philippaerts, R. (2012). The Contribution of Growth and Maturation in the Functional Capacity and Skill Performance of Male Adolescent Handball Players. *International Journal of Sports Medicine, 33*(07), 543–549. https://doi.org/10.1055/s-0031-1298000

McAuley, A. B. T., Hughes, D. C., Tsaprouni, L. G., Varley, I., Suraci, B., Roos, T. R., Herbert, A. J., & Kelly, A. L. (2020). Genetic association research in football: a systematic review. *European Journal of Sport Science, 0*(0), 1–52. https://doi.org/10.1080/17461391.2020.1776401

McCabe, K., & Collins, C. (2018). Can Genetics Predict Sports Injury? The Association of the Genes GDF5, AMPD1, COL5A1 and IGF2 on Soccer Player Injury Occurrence. Sports, 6(1), 21. https://doi.org/10.3390/sports6010021

McCall, A., Dupont, G., & Ekstrand, J. (2018). Internal workload and non-contact injury: a one-season study of five teams from the UEFA Elite Club Injury Study. *British Journal of Sports Medicine, 52*(23), 1517–1522. https://doi.org/10.1136/bjsports-2017-098473

McCauley, T., Mastana, S. S., Hossack, J., MacDonald, M., & Folland, J. P. (2009). Human angiotensin-converting enzyme I/D and α-actinin 3 R577X genotypes and muscle functional and contractile properties. *Experimental Physiology, 94*(1), 81–89. https://doi.org/10.1113/expphysiol.2008.043075

McNamee, M. J., Müller, A., Van Hilvoorde, I., & Holm, S. (2009). Genetic testing and sports medicine ethics. *Sports Medicine, 39*(5), 339–346. https://doi.org/10.2165/00007256-200939050-00001

Meardon, S. A., Willson, J. D., Gries, S. R., Kernozek, T. W., & Derrick, T. R. (2015). Bone stress in runners with tibial stress fracture. *Clinical Biomechanics, 30*(9), 895–902. https://doi.org/10.1016/j.clinbiomech.2015.07.012

Meeusen, R., Duclos, M., Gleeson, M., Rietjens, G., Steinacker, J., & Urhausen, A. (2006). Prevention, diagnosis and treatment of the Overtraining Syndrome. *European Journal of Sport Science, 6*(1), 1–14. https://doi.org/10.1080/17461390600617717

Meeuwisse, W. H. (1994). Assessing Causation in Sport Injury Multifact model. *Clinical Journal of Sport Medicine, 4*(3), 166–170.

Meeuwisse, W. H., Tyreman, H., Hagel, B., & Emery, C. (2007). A Dynamic Model of Etiology in Sport Injury: The Recursive Nature of Risk and Causation. *Clinical Journal of Sport Medicine, 17*(3), 215–219. https://doi.org/10.1097/JSM.0b013e3180592a48

Mendez-Villanueva, A., Buchheit, M., Kuitunen, S., Douglas, A., Peltola, E., & Bourdon, P. (2011). Age-related differences in acceleration, maximum running speed, and repeated-sprint performance in young soccer players. *Journal of Sports Sciences, 29*(5), 477–484. https://doi.org/10.1080/02640414.2010.536248

Michaëlsson, K., Melhus, H., Ferm, H., Ahlbom, A., & Pedersen, N. L. (2005). Genetic Liability to Fractures in the Elderly. *Archives of Internal Medicine, 165*(16), 1825. https://doi.org/10.1001/archinte.165.16.1825

Micheli, M. L., Gulisano, M., Morucci, G., Punzi, T., Ruggiero, M., Ceroti, M., Marella, M., Castellini, E., & Pacini, S. (2011). Angiotensin-Converting Enzyme/Vitamin D Receptor Gene Polymorphisms and Bioelectrical Impedance Analysis in Predicting Athletic Performances of Italian Young Soccer Players. *Journal of Strength and Conditioning Research, 25*(8), 2084–2091. https://doi.org/10.1519/JSC.0b013e31820238aa

Mikic, B. (2004). Multiple Effects of GDF-5 Deficiency on Skeletal Tissues: Implications for Therapeutic Bioengineering. *Annals of Biomedical Engineering, 32*(3), 466–476. https://doi.org/10.1023/B:ABME.0000017549.57126.51

Minelli, C., Thompson, J. R., Abrams, K. R., Thakkinstian, A., & Attia, J. (2009). The Quality of Meta-Analyses of Genetic Association Studies: A Review With Recommendations. *American Journal of Epidemiology, 170*(11), 1333–1343. https://doi.org/10.1093/aje/kwp350

Mitchell, J. A., Chesi, A., Elci, O., McCormack, S. E., Kalkwarf, H. J., Lappe, J. M., Gilsanz, V., Oberfield, S. E., Shepherd, J. A., Kelly, A., Zemel, B. S., & Grant, S. F. A. A. (2015). Genetics of Bone Mass in Childhood and Adolescence: Effects of Sex and Maturation Interactions. *Journal of Bone and Mineral Research : The Official Journal of the American Society for Bone and Mineral Research, 30*(9), 1676–1683. https://doi.org/10.1002/jbmr.2508

Miyamoto-Mikami, E., Miyamoto, N., Kumagai, H., Hirata, K., Kikuchi, N., Zempo, H., Kimura, N., Kamiya, N., Kanehisa, H., Naito, H., & Fuku, N. (2019). COL5A1 rs12722 polymorphism is not associated with passive muscle stiffness and sports-related muscle injury in Japanese athletes. *BMC Medical Genetics, 20*(1), 192. https://doi.org/10.1186/s12881-019-0928-2

Miyamoto, Y., Mabuchi, A., Shi, D., Kubo, T., Takatori, Y., Saito, S., Fujioka, M., Sudo, A., Uchida, A., Yamamoto, S., Ozaki, K., Takigawa, M., Tanaka, T., Nakamura, Y., Jiang, Q., & Ikegawa, S. (2007). A functional polymorphism in the 5′ UTR of GDF5 is associated with susceptibility to osteoarthritis. *Nature Genetics, 39*(4), 529–533. https://doi.org/10.1038/2005

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine, 6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Mohr, M., Krustrup, P., & Bangsbo, J. (2003). Match performance of high-standard soccer players with special reference to development of fatigue. *Journal of Sports Sciences, 21*(7), 519–528. https://doi.org/10.1080/0264041031000071182

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

Monasterio, X., Gil, S. M., Bidaurrazaga-Letona, I., Lekue, J. A., Santisteban, J. M., Diaz-Beitia, G., Lee, D.-J., Zumeta-Olaskoaga, L., Martin-Garetxana, I., Bikandi, E., & Larruskain, J. (2021a). The burden of injuries according to maturity status and timing: a two-decade study with 110 growth curves in an elite football academy. *European Journal of Sport Science, 0*, 1–21. https://doi.org/10.1080/17461391.2021.2006316

Monasterio, X., Gil, S. M., Bidaurrazaga-Letona, I., Lekue, J. A., Santisteban, J. M., Diaz-Beitia, G., Martin-Garetxana, I., Bikandi, E., & Larruskain, J. (2021b). Injuries according to the percentage of adult height in an elite soccer academy. *Journal of Science and Medicine in Sport, 24*(3), 218–223. https://doi.org/10.1016/j.jsams.2020.08.004

Montagna, R., Canonico, R., Alfano, L., Bucci, E., Boffo, S., Staiano, L., Fulco, B., D'Andrea, E., Nicola, A., Maiorano, P., D'Angelo, C., Chirico, A., Nicola, A., & Giordano, A. (2019). Genomic analysis reveals association of specific SNPs with athletic performance and susceptibility to injuries in professional soccer players. *Journal of Cellular Physiology, June*, jcp.29118. https://doi.org/10.1002/jcp.29118

Montgomery, H., Clarkson, P., Barnard, M., Bell, J., Brynes, A., Hajnal, J., Hemingway, H., Mercer, D., Jarman, P., Marshall, R., Prasad, K., Rayson, M., Saeed, N., Thomas, L., Jubb, M., World, M., Dollery, C., Talmud, P., & Humphries, S. (1999). Angiotensin-converting-enzyme gene insertion/deletion polymorphism and response to physical training. *The Lancet, 353*(9152), 541–545. https://doi.org/10.1016/S0140-6736(98)07131-1

Montgomery, H., Clarkson, P., Dollery, C. M., Prasad, K., Losi, M. A., Hemingway, H., Statters, D., Jubb, M., Girvain, M., Varnava, A., World, M., Deanfield, J., Talmud, P., McEwan, J. R., McKenna, W. J., & Humphries, S. (1997). Association of angiotensin-converting enzyme gene I/D polymorphism with change in left ventricular mass in response to physical training. *Circulation, 96*(3), 741–747. https://doi.org/10.1161/01.cir.96.3.741

Montgomery, H., Marshall, R., Hemingway, H., Myerson, S., Clarkson, P., Dollery, C., Hayward, M., Holliman, D. E., Jubb, M., World, M., Thomas, E. L., Brynes, A. E., Saeed, N., Barnard, M., Bell, J. D., Prasad, K., Rayson, M., Talmud, P. J., & Humphries, S. E. (1998). Human gene for physical performance. *Nature, 393*(6682), 221–222. https://doi.org/10.1038/30374

Moran, C. N., & Pitsiladis, Y. P. (2017). Tour de France Champions born or made: where do we take the genetics of performance? *Journal of Sports Sciences, 35*(14), 1411–1419. https://doi.org/10.1080/02640414.2016.1215494

Moran, D. S., McClung, J. P., Kohen, T., & Lieberman, H. R. (2013). Vitamin D and Physical Performance. *Sports Medicine, 43*(7), 601–611. https://doi.org/10.1007/s40279-013-0036-y

Morucci, G., Punzi, T., Innocenti, G., Gulisano, M., Ceroti, M., & Pacini, S. (2014). New frontiers in sport training: genetics and artistic gymnastics. *Journal of Strength and Conditioning Research, 28*(2), 459–466. https://doi.org/10.1519/JSC.0b013e31829aad65

Munafò, M. R., & Flint, J. (2004). Meta-analysis of genetic association studies. *Trends in Genetics, 20*(9), 439–444. https://doi.org/10.1016/j.tig.2004.06.014

Musch, J., & Grondin, S. (2001). Unequal competition as an impediment to personal development: A review of the relative age effect in sport. *Developmental Review, 21*, 147–167. https://doi.org/10.1006

Myerson, S., Hemingway, H., Budget, R., Martin, J., Humphries, S., & Montgomery, H. (1999). Human angiotensin I-converting enzyme gene and endurance performance. *Journal of Applied Physiology, 87*(4), 1313–1316. https://doi.org/10.1152/jappl.1999.87.4.1313

Myllyharju, J., & Kivirikko, K. I. (2001). Collagens and collagen-related diseases. *Annals of Medicine, 33*(1), 7–21. https://doi.org/10.3109/07853890109002055

Nagase, H., Visse, R., & Murphy, G. (2006). Structure and function of matrix metalloproteinases and TIMPs. *Cardiovascular Research, 69*(3), 562–573. https://doi.org/10.1016/j.cardiores.2005.12.002

Nakamura, O., Ishii, T., Ando, Y., Amagai, H., Oto, M., Imafuji, T., & Tokuyama, K. (2002a). Potential role of vitamin D receptor gene polymorphism in determining bone phenotype in young male athletes. *Journal of Applied Physiology, 93*(6), 1973–1979. https://doi.org/10.1152/japplphysiol.00663.2002

Nakamura, O., Ishii, T., Mankyu, H., Tsubakimoto, S., Nomura, T., & Tokuyama, K. (2002b). Contribution of vitamin D receptor genotypes to bone mineral density in young male athletes with different impact loading. *European Journal of Sport Science, 2*(2), 1–8. https://doi.org/10.1080/17461390200072203

Nazarov, I. B., Woods, D., Montgomery, H., Shneider, O., Kazakov, V., Tomilin, N., & Rogozkin, V. (2001). The angiotensin converting enzyme I/D polymorphism in Russian athletes. *European Journal of Human Genetics, 9*(10), 797–801. https://doi.org/10.1038/sj.ejhg.5200711

Nehme, A., Zouein, F. A., Zayeri, Z. D., & Zibara, K. (2019). An Update on the Tissue Renin Angiotensin System and Its Role in Physiology and Pathology. *Journal of Cardiovascular Development and Disease, 6*(2), 14. https://doi.org/10.3390/jcdd6020014

Nie, G., Wen, X., Liang, X., Zhao, H., Li, Y., & Lu, J. (2019). Additional evidence supports association of common genetic variants in MMP3 and TIMP2 with increased risk of chronic Achilles tendinopathy susceptibility. *Journal of Science and Medicine in Sport, 22*(10), 1074–1078. https://doi.org/10.1016/j.jsams.2019.05.021

Nielsen, R. O., Bertelsen, M. L., Møller, M., Hulme, A., Windt, J., Verhagen, E., Mansournia, M. A., Casals, M., & Parner, E. T. (2018). Training load and structure-specific load: applications for sport injury causality and data analyses. *British Journal of Sports Medicine, 52*(16), 1016–1017. https://doi.org/10.1136/bjsports-2017-097838

Norman, B., Esbjörnsson, M., Rundqvist, H. H., Österlund, T., von Walden, F., & Tesch, P. A. (2009). Strength, power, fiber types, and mRNA expression in trained men and women with different ACTN3 R577X genotypes. *Journal of Applied Physiology, 106*(3), 959–965. https://doi.org/10.1152/japplphysiol.91435.2008

Norman, B., Esbjörnsson, M., Rundqvist, H., Österlund, T., Glenmark, B., & Jansson, E. (2014). ACTN3 genotype and modulation of skeletal muscle response to exercise in human subjects. *Journal of Applied Physiology, 116*(9), 1197–1203. https://doi.org/10.1152/japplphysiol.00557.2013

North, K., Yang, N., Wattanasirichaigoon, D., Mills, M., Easteal, S., & Beggs, A. H. (1999). A common nonsense mutation results in α-actinin-3 deficiency in the general population. *Nature, 21*(4), 353–354. https://doi.org/10.1038/7675

Norton, K. I. (2018). Standards for Anthropometry Assessment. *In Kinanthropometry and Exercise Physiology* (pp. 68–137). Routledge. https://doi.org/10.4324/9781315385662-4

O'Connell, K., Knight, H., Ficek, K., Leońska-Duniec, A., Maciejewska-Karłowska, A., Sawczuk, M., Stępień-Słodkowska, M., O'Cuinneagain, D., van der Merwe, W., Posthumus, M., Cięszczyk, P., & Collins, M. (2015). Interactions between collagen gene variants and risk of anterior cruciate ligament rupture. *European Journal of Sport Science, 15*(4), 341–350. https://doi.org/10.1080/17461391.2014.936324

O'Connell, K., Posthumus, M., Schwellnus, M. P., & Collins, M. (2013). Collagen Genes and Exercise-Associated Muscle Cramping. *Clinical Journal of Sport Medicine, 23*(1), 64–69. https://doi.org/10.1097/JSM.0b013e3182686aa7

Ogden, J. A., & Southwick, W. O. (1976). Osgood-Schlatter's disease and tibial tuberosity development. *Clinical Orthopaedics and Related Research, NA*;(116), 180–189. https://doi.org/10.1097/00003086-197605000-00032

Orysiak, J., Mazur-Różycka, J., Busko, K., Gajewski, J., Szczepanska, B., & Malczewska-Lenczowska, J. (2018). Individual and Combined Influence of ACE and ACTN3 Genes on Muscle Phenotypes in Polish Athletes. *Journal of Strength and Conditioning Research, 32*(10), 2776–2782. https://doi.org/10.1519/JSC.0000000000001839

Orysiak, J., Zmijewski, P., Klusiewicz, A., Kaliszewski, P., Malczewska-Lenczowska, J., Gajewski, J., & Pokrywka, A. (2013). The association between ace gene variation and aerobic capacity in winter endurance disciplines. *Biology of Sport, 30*(4), 249–253. https://doi.org/10.5604/20831862.1077549

Owens, D. J., Allison, R., & Close, G. L. (2018). Vitamin D and the Athlete: Current Perspectives and New Challenges. *Sports Medicine, 48*(S1), 3–16. https://doi.org/1-.1007/s40279-017-0841-9

Pabalan, N., Tabangay, L., Jarjanazi, H., Vieira, L. A., dos Santos, A. A., Barbosa, C. P., Rodrigues, L. M. R., & Bianco, B. (2017). Association Between the FokI and ApaI Polymorphisms in the Vitamin D Receptor Gene and Intervertebral Disc Degeneration: A Systematic Review and Meta-Analysis. *Genetic Testing and Molecular Biomarkers, 21*(1), 24–32. https://doi.org/10.1089/gtmb.2016.0054

Pabalan, N., Tharabenjasin, P., Phababpha, S., & Jarjanazi, H. (2018). Association of COL5A1 gene polymorphisms and risk of tendon-ligament injuries among Caucasians: a meta-analysis. *Sports Medicine - Open, 4*(1), 46. https://doi.org/10.1186/s40798-018-0161-0

Papadimitriou, I. D., Lockey, S. J., Voisin, S., Herbert, A. J., Garton, F., Houweling, P. J., Cieszczyk, P., Maciejewska-Skrendo, A., Sawczuk, M., Massidda, M., Calò, C. M., Astratenkova, I. V., Kouvatsi, A., Druzhevskaya, A. M., Jacques, M., Ahmetov, I., Stebbings, G. K., Heffernan, S. M., Day, S. H., … Eynon, N. (2018). No association between ACTN3 R577X and ACE I/D polymorphisms and endurance running times in 698 Caucasian athletes. *BMC Genomics, 19*(1), 13. https://doi.org/10.1186/s12864-017-4412-0

Papadimitriou, I. D., Lucia, A., Pitsiladis, Y. P., Pushkarev, V. P., Dyatlov, D. A., Orekhov, E. F., Artioli, G. G., Guilherme, J. P. L. F., Lancha, A. H., Ginevičienė, V., Cieszczyk, P., Maciejewska-Karlowska, A., Sawczuk, M., Muniesa, C. A., Kouvatsi, A., Massidda, M., Calò, C. M., Garton, F., Houweling, P. J., … Eynon, N. (2016). ACTN3 R577X and ACE I/D gene variants influence performance in elite sprinters: a multi-cohort study. *BMC Genomics, 17*(1), 285. https://doi.org/10.1186/s12864-016-2462-3

Parfitt, A. M. (1994). The two faces of growth: Benefits and risks to bone integrity. *Osteoporosis International, 4*(6), 382–398. https://doi.org/10.1007/BF01622201

Parr, J., Winwood, K., Hodson-Tole, E., Deconinck, F. J. A., Parry, L., Hill, J. P., Malina, R. M., & Cumming, S. P. (2020). Predicting the timing of the peak of the pubertal growth spurt in elite male youth soccer players: evaluation of methods. *Annals of Human Biology, 47*(4), 400–408. https://doi.org/10.1080/03014460.2020.1782989

Pasiakos, S. M., Lieberman, H. R., & McLellan, T. M. (2014). Effects of Protein Supplements on Muscle Damage, Soreness and Recovery of Muscle Function and Physical Performance: A Systematic Review. *Sports Medicine, 44*(5), 655–670. https://doi.org/10.1007/s40279-013-0137-7

Pekala, P. A., Henry, B. M., Taterra, D., Piwowar, M., Vikse, J., Tubbs, R. S., & Tomaszewski, K. A. (2019). FokI as a genetic factor of intervertebral disc degeneration: A PRISMA-compliant systematic review of overlapping meta-analyses. *Journal of Clinical Neuroscience, 60*, 36–43. https://doi.org/10.1016/j.jocn.2018.09.028

Pereira, A., Costa, A. M., Izquierdo, M., Silva, A. J., Bastos, E., & Marques, M. C. (2013a). ACE I/D and ACTN3 R/X polymorphisms as potential factors in modulating exercise-related phenotypes in older women in response to a muscle power training stimuli. *AGE, 35*(5), 1949–1959. https://doi.org/10.1007/s11357-012-9461-3

Pereira, A., Costa, A. M., Leitão, J. C., Monteiro, A. M., Izquierdo, M., Silva, A. J., Bastos, E., & Marques, M. C. (2013b). The influence of ACE ID and ACTN3 R577X polymorphisms on lower-extremity function in older women in response to high-speed power training. *BMC Geriatrics, 13*(1), 131. https://doi.org/10.1186/1471-2318-13-131

Pérez-Lezaun, A., Calafell, F., Mateu, E., Comas, D., Bosch, E., & Bertranpetit, J. (1997). Allele frequencies for 20 microsatellites in a worldwide population survey. *Human Heredity, 47*(4), 189–196. https://doi.org/10.1159/000154412

Pertea, M., & Salzberg, S. L. (2010). Between a chicken and a grape: estimating the number of human genes. Genome *Biology, 11*(5), 206. https://doi.org/10.1186/gb-2010-11-5-206

Pescatello, L. S., Kostek, M. A., Gordish-Dressman, H., Thompson, P. D., Seip, R. L., Price, T. B., Angelopoulos, T. J., Clarkson, P. M., Gordon, P. M., Moyna, N. M., Visich, P. S., Zoeller, R. F., Devaney, J. M., & Hoffman, E. P. (2006). ACE ID Genotype and the Muscle Strength and Size Response to Unilateral Resistance Training. M*edicine & Science in Sports & Exercise, 38*(6), 1074–1081. https://doi.org/10.1249/01.mss.0000222835.28273.80

Pfirrmann, D., Herbst, M., Ingelfinger, P., Simon, P., & Tug, S. (2016). Analysis of injury incidences in male professional adult and elite youth soccer players: A systematic review. *Journal of Athletic Training, 51*(5), 410–424. https://doi.org/10.4085/1062-6050-51.6.03

Philippaerts, R. M., Vaeyens, R., Janssens, M., Van Renterghem, B., Matthys, D., Craen, R., Bourgois, J., Vrijens, J., Beunen, G., & Malina, R. M. (2006). The relationship between peak height velocity and physical performance in youth soccer players. *Journal of Sports Sciences, 24*(3), 221–230. https://doi.org/10.1080/02640410500189371

Pickering, C., & Kiely, J. (2017a). Exercise genetics: seeking clarity from noise. *BMJ Open Sport & Exercise Medicine, 3*(1), e000309. https://doi.org/10.1136/bmjsem-2017-000309

Pickering, C., & Kiely, J. (2017b). ACTN3: More than Just a Gene for Speed. *Frontiers in Physiology, 8*(DEC), 1–9. https://doi.org/10.3389/fphys.2017.01080

Pickering, C., & Kiely, J. (2018). ACTN3, Morbidity, and Healthy Aging. *Frontiers in Genetics, 9*(JAN), 1–8. https://doi.org/10.3389/fgene.2018.00015

Pickering, C., & Kiely, J. (2019). The Development of a Personalised Training Framework: Implementation of Emerging Technologies for Performance. *Journal of Functional Morphology and Kinesiology, 4*(2), 25. https://doi.org/10.3390/jfmk4020025

Pickering, C., Kiely, J., Suraci, B., & Collins, D. (2018). The magnitude of Yo-Yo test improvements following an aerobic training intervention are associated with total genotype score. *PLoS ONE, 13*(11), 1–11. https://doi.org/10.1371/journal.pone.0207597

Pimenta, E., Coelho, D. B., Cruz, I. R., Morandi, R. F., Veneroso, C. E., de Azambuja Pussieldi, G., Carvalho, M. R. S., Silami-Garcia, E., & De Paz Fernández, J. A. (2012). The ACTN3 genotype in soccer players in response to acute eccentric training. *European Journal of Applied Physiology, 112*(4), 1495–1503. https://doi.org/10.1007/s00421-011-2109-7

Posthumus, M., Collins, M., Cook, J., Handley, C. J., Ribbans, W. J., Smith, R. K. W. W., Schwellnus, M. P., & Raleigh, S. M. (2010). Components of the transforming growth factor-β family and the pathogenesis of

human achilles tendon pathology-a genetic association study. *Rheumatology, 49*(11), 2090–2097. https://doi.org/10.1093/rheumatology/keq072

Posthumus, M., Raleigh, S. M., Ribbans, W. J., Schwellnus, M. P., & Collins, M. (2010). A functional variant within the MMP3 gene does not associate with human range of motion. *Journal of Science and Medicine in Sport, 13*(6), 630–632. https://doi.org/10.1016/j.jsams.2010.01.006

Posthumus, M., Saunders, C. J., September, A. V, & Collins, M. (2011). The polygenic profiles in participants with achilles tendinopathy and controls. *British Journal of Sports Medicine, 45*(4), 314. https://doi.org/10.1136/bjsm.2011.084038

Posthumus, M., Schwellnus, M. P., & Collins, M. (2011b). The COL5A1 gene: A novel marker of endurance running performance. *Medicine and Science in Sports and Exercise, 43*(4), 584–589. https://doi.org/10.1249/MSS.0b013e3181f34f4d

Posthumus, M., September, A. V., Keegan, M., O'Cuinneagain, D., Van der Merwe, W., Schwellnus, M. P., & Collins, M. (2009a). Genetic risk factors for anterior cruciate ligament ruptures: COL1A1 gene variant. *British Journal of Sports Medicine, 43*(5), 352–356. https://doi.org/10.1136/bjsm.2008.056150

Posthumus, M., September, A. V., O'Cuinneagain, D., van der Merwe, W., Schwellnus, M. P., & Collins, M. (2009b). The COL5A1 Gene Is Associated With Increased Risk of Anterior Cruciate Ligament Ruptures in Female Participants. *The American Journal of Sports Medicine, 37*(11), 2234–2240. https://doi.org/10.1177/0363546509338266

Posthumus, M., September, A. V, Schwellnus, M. P., & Collins, M. (2009c). Investigation of the Sp1-binding site polymorphism within the COL1A1 gene in participants with Achilles tendon injuries and controls. *Journal of Science and Medicine in Sport, 12*(1), 184–189. https://doi.org/10.1016/j.jsams.2007.12.006

Premier League. (2012). *Elite Player Performance Plan*. London, Premier League.

Price, R. J., Hawkins, R. D., Hulse, M. A., & Hodson, A. (2004). The Football Association medical research programme: an audit of injuries in academy youth football. *British Journal of Sports Medicine, 38*(4), 466–471. https://doi.org/10.1136/bjsm.2003.005165

Pruna, R., Artells, R., Lundblad, M., & Maffulli, N. (2016). Genetic biomarkers in non-contact muscle injuries in elite soccer players. *Knee Surgery, Sports Traumatology, Arthroscopy : Official Journal of the ESSKA, 25*(10), 3311–3318. https://doi.org/10.1007/s00167-016-4081-6

Pruna, R., Artells, R., Ribas, J., Montoro, B., Cos, F., Muñoz, C., Rodas, G., & Maffulli, N. (2013). Single nucleotide polymorphisms associated with non-contact soft tissue injuries in elite professional soccer players: influence on degree of injury and recovery time. *BMC Musculoskeletal Disorders, 14*(1), 221. https://doi.org/10.1186/1471-2474-14-221

Pruna, R., Ribas, J., Montoro, J. B., & Artells, R. (2015). The impact of single nucleotide polymorphisms on patterns of non-contact musculoskeletal soft tissue injuries in a football player population according to ethnicity. *Medicina Clínica, 144*(3), 105–110. https://doi.org/10.1016/j.medcli.2013.09.026

Puthucheary, Z. (2011). The ACE Gene and Human Performance: 12 Years On. *Sports Medicine, 41*(6), 433–448. http://search.ebscohost.com/login.aspx?direct=true&db=s3h&AN=63800128&site=ehost-live

Qasim, A., Mayhew, A. J., Ehtesham, S., Alyass, A., Volckmar, A. -L., Herpertz, S., Hinney, A., Hebebrand, J., & Meyre, D. (2019). Gain-of-function variants in the melanocortin 4 receptor gene confer susceptibility to binge eating disorder in subjects with obesity: a systematic review and meta-analysis. *Obesity Reviews, 20*(1), 13–21. https://doi.org/10.1111/obr.12761

Qi, B., Liu, J. Q., & Liu, G. L. (2016). Genetic association between ACTN3 polymorphism and risk of non-acute ankle sprain. *Genetics and Molecular Research, 15*(4), 1–7. https://doi.org/10.4238/gmr15048962

Quadrilatero, J., Alway, S. E., & Dupont-Versteegden, E. E. (2011). Skeletal muscle apoptotic response to physical activity: potential mechanisms for protection. *Applied Physiology, Nutrition, and Metabolism, 36*(5), 608–617. https://doi.org/10.1139/h11-064

Quatman-Yates, C. C., Quatman, C. E., Meszaros, A. J., Paterno, M. V., & Hewett, T. E. (2012). A systematic review of sensorimotor function during adolescence: a developmental stage of increased motor awkwardness? British Journal of *Sports Medicine, 46*(9), 649–655. https://doi.org/10.1136/bjsm.2010.079616

Rabon-Stith, K. M., Hagberg, J. M., Phares, D. A., Kostek, M. C., Delmonico, M. J., Roth, S. M., Ferrell, R. E., Conway, J. M., Ryan, A. S., & Hurley, B. F. (2005). Vitamin D receptor FokI genotype influences bone mineral density response to strength training, but not aerobic training. *Experimental Physiology, 90*(4), 653–661. https://doi.org/10.1113/expphysiol.2005.030197

Rae, K., & Orchard, J. (2007). The Orchard Sports Injury Classification System (OSICS) Version 10. Clinical Journal of *Sport Medicine, 17*(3), 201–204. https://doi.org/10.1097/JSM.0b013e318059b536

Rahim, M., Gibbon, A., Collins, M., & September, A. V. (2019). Genetics of musculoskeletal soft tissue injuries: Current status, challenges, and future directions. In *Sports, Exercise, and Nutritional Genomics: Current Status and Future Directions*. Elsevier Inc. https://doi.org/10.1016/B978-0-12-816193-7.00015-4

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

Raisz, L. G., & Seeman, E. (2001). Causes of age-related bone loss and bone fragility: an alternative view. *Journal of Bone and Mineral Research : The Official Journal of the American Society for Bone and Mineral Research, 16*(11), 1948–1952. https://doi.org/10.1359/jbmr.2001.16.11.1948

Raleigh, S. M., Posthumus, M., O'Cuinneagain, D., Van Der Merwe, W., & Collins, M. (2013). The GDF5 gene and anterior cruciate ligament rupture. *International Journal of Sports Medicine, 34*(4), 364–367. https://doi.org/10.1055/s-0032-1316361

Raleigh, S. M., van der Merwe, L., Ribbans, W. J., Smith, R. K. W., Schwellnus, M. P., & Collins, M. (2009). Variants within the MMP3 gene are associated with Achilles tendinopathy: possible interaction with the COL5A1 gene. *British Journal of Sports Medicine, 43*(7), 514–520. https://doi.org/10.1136/bjsm.2008.053892

Rankinen, T., Pérusse, L., Gagnon, J., Chagnon, Y. C., Leon, A. S., Skinner, J. S., Wilmore, J. H., Rao, D. C., & Bouchard, C. (2000a). Angiotensin-converting enzyme ID polymorphism and fitness phenotype in the HERITAGE Family Study. *Journal of Applied Physiology, 88*(3), 1029–1035. https://doi.org/10.1152/jappl.2000.88.3.1029

Rankinen, T., Wolfarth, B., Simoneau, J.-A., Maier-Lenz, D., Rauramaa, R., Rivera, M. A., Boulay, M. R., Chagnon, Y. C., Pérusse, L., Keul, J., & Bouchard, C. (2000b). No association between the angiotensin-converting enzyme ID polymorphism and elite endurance athlete status. *Journal of Applied Physiology, 88*(5), 1571–1575. https://doi.org/10.1152/jappl.2000.88.5.1571

Rathleff, M. S., Winiarski, L., Krommes, K., Graven-Nielsen, T., Hölmich, P., Olesen, J. L., Holden, S., & Thorborg, K. (2020). Pain, Sports Participation, and Physical Function in Adolescents With Patellofemoral Pain and Osgood-Schlatter Disease: A Matched Cross-sectional Study. *Journal of Orthopaedic & Sports Physical Therapy, 50*(3), 149–157. https://doi.org/10.2519/jospt.2020.8770

Read, P. J., Oliver, J. L., De Ste Croix, M. B. A., Myer, G. D., & Lloyd, R. S. (2018a). A prospective investigation to evaluate risk factors for lower extremity injury risk in male youth soccer players. *Scandinavian Journal of Medicine & Science in Sports, 28*(3), 1244–1251. https://doi.org/10.1111/sms.13013

Read, P. J., Oliver, J. L., De Ste Croix, M. B. A., Myer, G. D., & Lloyd, R. S. (2018b). An audit of injuries in six english professional soccer academies. *Journal of Sports Sciences, 36*(13), 1542–1548. https://doi.org/10.1080/02640414.2017.1402535

Reilly, T. (2006). Assessments of Young Soccer Players: A Holistic Approach. *Perceptual and Motor Skills, 103*(1), 229–230. https://doi.org/10.2466/pms.103.1.229-230

Reilly, T., Bangsbo, J., & Franks, A. (2000). Anthropometric and physiological predispositions for elite soccer. *Journal of Sports Sciences, 18*(9), 669–683. https://doi.org/10.1080/02640410050120050

Renshaw, A., & Goodwin, P. C. (2016). Injury incidence in a Premier League youth soccer academy using the consensus statement: a prospective cohort study. *BMJ Open Sport & Exercise Medicine, 2*(1), e000132. https://doi.org/10.1136/bmjsem-2016-000132

Renstrom, P., Ljungqvist, A., Arendt, E., Beynnon, B., Fukubayashi, T., Garrett, W., Georgoulis, T., Hewett, T. E., Johnson, R., Krosshaug, T., Mandelbaum, B., Micheli, L., Myklebust, G., Roos, E., Roos, H., Schamasch, P., Shultz, S., Werner, S., Wojtys, E., & Engebretsen, L. (2008). Non-contact ACL injuries in female athletes: an International Olympic Committee current concepts statement. *British Journal of Sports Medicine, 42*(6), 394–412. https://doi.org/10.1136/bjsm.2008.048934

Ribeiro-Oliveira, A., Nogueira, A. I., Pereira, R. M., Boas, W. W. V., Dos Santos, R. A. S., & Simões e Silva, A. C. (2008). The renin-angiotensin system and diabetes: an update. *Vascular Health and Risk Management, 4*(4), 787–803. https://doi.org/10.2147/VHRM.S1905

Ribeiro, J., Davids, K., Silva, P., Coutinho, P., Barreira, D., & Garganta, J. (2021). Talent Development in Sport Requires Athlete Enrichment: Contemporary Insights from a Nonlinear Pedagogy and the Athletic Skills Model. *Sports Medicine, 51*(6), 1115–1122. https://doi.org/10.1007/s40279-021-01437-6

Rigat, B., Hubert, C., Alhenc-Gelas, F., Cambien, F., Corvol, P., & Soubrier, F. (1990). An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. *The Journal of Clinical Investigation, 86*(4), 1343–1346. https://doi.org/10.1172/JCI114844

Rizzo, M., Gensini, F., Fatini, C., Manetti, P., Pucci, N., Capalbo, A., Vono, M. C. R., & Galanti, G. (2003). ACE I/D polymorphism and cardiac adaptations in adolescent athletes. *Medicine and Science in Sports and Exercise, 35*(12), 1986–1990. https://doi.org/10.1249/01.MSS.0000098993.51693.0B

Rodríguez-Romo, G., Ruiz, J. R., Santiago, C., Fiuza-Luces, C., González-Freire, M., Gómez-Gallego, F., Morán, M., & Lucia, A. (2010). Does the ACE I/D polymorphism, alone or in combination with the ACTN3 R577X polymorphism, influence muscle power phenotypes in young, non-athletic adults? *European Journal of Applied Physiology, 110*(6), 1099–1106. https://doi.org/10.1007/s00421-010-1608-2

Roessler, R., Donath, L., Verhagen, E., Junge, A., Schweizer, T., Faude, O., Rössler, R., Donath, L., Verhagen, E., Junge, A., Schweizer, T., & Faude, O. (2014). Exercise-Based Injury Prevention in Child and Adolescent Sport: A Systematic Review and Meta-Analysis. *Sports Medicine, 44*(12), 1733–1748. https://doi.org/10.1007/s40279-014-0234-2

Rogol, A. D., Roemmich, J. N., & Clark, P. A. (2002). Growth at puberty. *Journal of Adolescent Health, 31*(6), 192–200. https://doi.org/10.1016/S1054-139X(02)00485-8

Roltsch, M. H., Brown, M. D., Hand, B. D., Kostek, M. C., Huberty, A., Ferrell, R. E., & Hagberg, J. M. (2002). No Association Between ACE Insertion/Deletion Polymorphism and Maximal Exercise Cardiovascular Hemodynamics In Young Women. *Medicine & Science in Sports & Exercise, 34*(5), S7. https://doi.org/10.1097/00005768-200205001-00041

Romero, R., Kuivaniemi, H., Tromp, G., & Olson, J. M. (2002). The design, execution, and interpretation of genetic association studies to decipher complex diseases. *American Journal of Obstetrics and Gynecology, 187*(5), 1299–1312. https://doi.org/10.1067/mob.2002.128319

Roth, S. M., Walsh, S., Liu, D., Metter, E. J., Ferrucci, L., & Hurley, B. F. (2008). The ACTN3 R577X nonsense allele is under-represented in elite-level strength athletes. *European Journal of Human Genetics, 16*(3), 391–394. https://doi.org/10.1038/sj.ejhg.5201964

Ruiz, J. R., Arteta, D., Buxens, A., Artieda, M., Gómez-Gallego, F., Santiago, C., Yvert, T., Morán, M., & Lucia, A. (2010). Can we identify a power-oriented polygenic profile? *Journal of Applied Physiology, 108*(3), 561–566. https://doi.org/10.1152/japplphysiol.01242.2009

Ruiz, J. R., Gómez-Gallego, F., Santiago, C., González-Freire, M., Verde, Z., Foster, C., & Lucia, A. (2009). Is there an optimum endurance polygenic profile? *The Journal of Physiology, 587*(7), 1527–1534. https://doi.org/10.1113/jphysiol.2008.166645

Ruiz, J. R., Santiago, C., Yvert, T., Muniesa, C., Díaz-Ureña, G., Bekendam, N., Fiuza-Luces, C., Gómez-Gallego, F., Femia, P., & Lucia, A. (2013). ACTN3 genotype in Spanish elite swimmers: No "heterozygous advantage." *Scandinavian Journal of Medicine & Science in Sports, 23*(3), e162–e167. https://doi.org/10.1111/sms.12045

Rumpf, M. C., & Cronin, J. (2012). Injury Incidence, Body Site, and Severity in Soccer Players Aged 6–18 Years: Implications for Injury Prevention. *Strength and Conditioning Journal, 34*(1), 20–31. https://doi.org/10.1519/SSC.0b013e31821a9833

Ryan-Moore, E., Mavrommatis, Y., & Waldron, M. (2020). Systematic Review and Meta-Analysis of Candidate Gene Association Studies With Fracture Risk in Physically Active Participants. *Frontiers in Genetics, 11*(June). https://doi.org/10.3389/fgene.2020.00551

Ryan, D., Lewin, C., Forsythe, S., & McCall, A. (2018). Developing World-Class Soccer Players: An Example of the Academy Physical Development Program From an English Premier League Team. *Strength & Conditioning Journal, 40*(3), 2–11. https://doi.org/10.1519/SSC.0000000000000340

Saavedra, C., Lagassé, P., Bouchard, C., & Simoneau, J. A. (1991). Maximal anaerobic performance of the knee extensor muscles during growth. *Medicine and Science in Sports and Exercise, 23*(9), 1083–1089. http://www.ncbi.nlm.nih.gov/pubmed/1943630

Salanti, G., Sanderson, S., & Higgins, J. P. T. (2005). Obstacles and opportunities in meta-analysis of genetic association studies. *Genetics in Medicine, 7*(1), 13–20. https://doi.org/10.1097/01.GIM.0000151839.12032.1A

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-Size Indices for Dichotomized Outcomes in Meta-Analysis. *Psychological Methods, 8*(4), 448–467. https://doi.org/10.1037/1082-989X.8.4.448

Santiago, C., Ruiz, J. R., Muniesa, C. A., González-Freire, M., Gómez-Gallego, F., & Lucia, A. (2010). Does the polygenic profile determine the potential for becoming a world-class athlete? Insights from the sport of rowing. *Scandinavian Journal of Medicine & Science in Sports, 20*(1), e188–e194. https://doi.org/10.1111/j.1600-0838.2009.00943.x

Saunders, C. J., September, A. V., Xenophontos, S. L., Cariolou, M. A., Anastassiades, L. C., Noakes, T. D., & Collins, M. (2007). No Association of the ACTN3 Gene R577X Polymorphism with Endurance Performance in Ironman Triathlons. *Annals of Human Genetics, 71*(6), 777–781. https://doi.org/10.1111/j.1469-1809.2006.00385.x

Schuh-Renner, A., Grier, T. L., Canham-Chervak, M., Hauschild, V. D., Roy, T. C., Fletcher, J., & Jones, B. H. (2017). Risk factors for injury associated with low, moderate, and high mileage road marching in a U.S. Army infantry brigade. *Journal of Science and Medicine in Sport, 20*, S28–S33. https://doi.org/10.1016/j.jsams.2017.07.027

Scott, R. A., Irving, R., Irwin, L., Morrison, E., Charlton, V., Austin, K., Tladi, D., Deason, Mi., Headley, S. A., Kolkhorst, F. W., Yang, N., North, K., & Pitsiladis, Y. P. (2010). ACTN3 and ACE Genotypes in Elite Jamaican and US Sprinters. *Medicine & Science in Sports & Exercise, 42*(1), 107–112. https://doi.org/10.1249/MSS.0b013e3181ae2bc0

Scott, R. A., Moran, C., Wilson, R. H., Onywera, V., Boit, M. K., Goodwin, W. H., Gohlke, P., Payne, J., Montgomery, H., & Pitsiladis, Y. P. (2005). No association between Angiotensin Converting Enzyme (ACE) gene variation and endurance athlete status in Kenyans. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 141*(2), 169–175. https://doi.org/10.1016/j.cbpb.2005.05.001

Sedgwick, P. (2014). Retrospective cohort studies: advantages and disadvantages. *BMJ, 348*(jan24 1), g1072–g1072. https://doi.org/10.1136/bmj.g1072

Semagn, K., Babu, R., Hearne, S., & Olsen, M. (2014). Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. *Molecular Breeding, 33*(1), 1–14. https://doi.org/10.1007/s11032-013-9917-x

September, A. V, Cook, J., Handley, C. J., van der Merwe, L., Schwellnus, M. P., & Collins, M. (2009). Variants within the COL5A1 gene are associated with Achilles tendinopathy in two populations. *British Journal of Sports Medicine, 43*(5), 357–365. https://doi.org/10.1136/bjsm.2008.048793

September, A. V, Schwellnus, M. P., Collins, M., & Gibson, W. T. (2007). Tendon and ligament injuries: the genetic component. *British Journal of Sports Medicine, 41*(4), 241–246. https://doi.org/10.1136/bjsm.2006.033035

Sessa, F., Chetta, M., Petito, A., Franzetti, M., Bafunno, V., Pisanelli, D., Sarno, M., Iuso, S., & Margaglione, M. (2011). Gene polymorphisms and sport attitude in Italian athletes. *Genetic Testing and Molecular Biomarkers, 15*(4), 285–290. https://doi.org/10.1089/gtmb.2010.0179

Seto, J. T., Quinlan, K. G. R., Lek, M., Zheng, X. F., Garton, F., MacArthur, D. G., Hogarth, M. W., Houweling, P. J., Gregorevic, P., Turner, N., Cooney, G. J., Yang, N., & North, K. N. (2013). ACTN3 genotype influences muscle performance through the regulation of calcineurin signaling. *Journal of Clinical Investigation, 123*(10), 4255–4263. https://doi.org/10.1172/JCI67691

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ, 349*(jan02 1), g7647–g7647. https://doi.org/10.1136/bmj.g7647

Shang, X., Li, Z., Cao, X., Xie, C., Gu, M., Chen, P., Yang, X., & Cai, J. (2015). The association between the ACTN3 R577X polymorphism and noncontact acute ankle sprains. *Journal of Sports Sciences, 33*(17), 1775–1779. https://doi.org/10.1080/02640414.2015.1012098

Sherar, L. B., Mirwald, R. L., Baxter-Jones, A., & Thomis, M. (2005). Prediction of adult height using maturity-based cumulative height velocity curves. *Journal of Pediatrics, 147*(4), 508–514. https://doi.org/10.1016/j.jpeds.2005.04.041

Sierra, A. P. R., Lima, G. H. O., da Silva, E. D., Maciel, J. F. de S., Benetti, M. P., de Oliveira, R. A., Martins, P. F. de O., Kiss, M. A. P., Ghorayeb, N., Newsholme, P., Pesquero, J. B., & Cury-Boaventura, M. F. (2019). Angiotensin-Converting Enzyme Related-Polymorphisms on Inflammation, Muscle and Myocardial Damage After a Marathon Race. *Frontiers in Genetics, 10*(October), 1–12. https://doi.org/10.3389/fgene.2019.00984

Silva, M. S. M., Bolani, W., Alves, C. R., Biagi, D. G., Lemos, J. R., da Silva, J. L., de Oliveira, P. A., Alves, G. B., de Oliveira, E. M., Negrão, C. E., Krieger, J. E., Dias, R. G., & Pereira, A. C. (2015). Elimination of Influences of the ACTN3 R577X Variant on Oxygen Uptake by Endurance Training in Healthy Individuals. *International Journal of Sports Physiology and Performance, 10*(5), 636–641. https://doi.org/10.1123/ijspp.2014-0205

Silventoinen, K., Magnusson, P. K. E., Tynelius, P., Kaprio, J., & Rasmussen, F. (2008). Heritability of body size and muscle strength in young adulthood: a study of one million Swedish men. *Genetic Epidemiology, 32*(4), 341–349. https://doi.org/10.1002/gepi.20308

Simmons, R., Doma, K., Sinclair, W., Connor, J., & Leicht, A. (2021). Acute Effects of Training Loads on Muscle Damage Markers and Performance in Semi-elite and Elite Athletes: A Systematic Review and Meta-analysis. *Sports Medicine, 51*(10), 2181–2207. https://doi.org/10.1007/s40279-021-01486-x

Smeets, J. S. J., Horstman, A. M. H., Vles, G. F., Emans, P. J., Goessens, J. P. B., Gijsen, A. P., van Kranenburg, J. M. X., & van Loon, L. J. C. (2019). Protein synthesis rates of muscle, tendon, ligament, cartilage, and bone tissue in vivo in humans. *PLOS ONE, 14*(11), e0224745. https://doi.org/10.1371/journal.pone.0224745

Sohani, Z. N., Meyre, D., de Souza, R. J., Joseph, P. G., Gandhi, M., Dennis, B. B., Norman, G., & Anand, S. S. (2015). Assessing the quality of published genetic association studies in meta-analyses: The quality of genetic studies (Q-Genie) tool. *BMC Genetics, 16*(1), 1–8. https://doi.org/10.1186/s12863-015-0211-2

Sohani, Z. N., Sarma, S., Alyass, A., De Souza, R. J., Robiou-Du-Pont, S., Li, A., Mayhew, A., Yazdi, F., Reddon, H., Lamri, A., Stryjecki, C., Ishola, A., Lee, Y. K., Vashi, N., Anand, S. S., & Meyre, D. (2016). Empirical evaluation of the Q-Genie tool: A protocol for assessment of effectiveness. *BMJ Open, 6*(6), 1–8. https://doi.org/10.1136/bmjopen-2015-010403

Somerville, R. P. T., Oblander, S. A., & Apte, S. S. (2003). Matrix metalloproteinases: Old dogs with new tricks. *Genome Biology, 4*(6). https://doi.org/10.1186/gb-2003-4-6-216

Sonoda, T., Takada, J., Iba, K., Asakura, S., Yamashita, T., & Mori, M. (2012). Interaction between ESRα polymorphisms and environmental factors in osteoporosis. *Journal of Orthopaedic Research, 30*(10), 1529–1534. https://doi.org/10.1002/jor.22083

Soomro, N., Sanders, R., Hackett, D., Hubka, T., Ebrahimi, S., Freeston, J., & Cobley, S. (2016). The Efficacy of Injury Prevention Programs in Adolescent Team Sports. *The American Journal of Sports Medicine, 44*(9), 2415–2424. https://doi.org/10.1177/0363546515618372

Southam, L., Rodriguez-Lopez, J., Wilkins, J. M., Pombo-Suarez, M., Snelling, S., Gomez-Reino, J. J., Chapman, K., Gonzalez, A., & Loughlin, J. (2007). An SNP in the 5′-UTR of GDF5 is associated with osteoarthritis susceptibility in Europeans and with in vivo differences in allelic expression in articular cartilage. *Human Molecular Genetics, 16*(18), 2226–2232. https://doi.org/10.1093/hmg/ddm174

Specker, B., Thiex, N. W., & Sudhagoni, R. G. (2015). Does Exercise Influence Pediatric Bone? A Systematic Review. *Clinical Orthopaedics & Related Research, 473*(11), 3658–3672. https://doi.org/10.1007/s11999-015-4467-7

Stastny, P., Lehnert, M., De Ste Croix, M., Petr, M., Svoboda, Z., Maixnerova, E., Varekova, R., Botek, M., Petrek, M., Kocourkova, L., & Cięszczyk, P. (2019). Effect of COL5A1, GDF5, and PPARA Genes on a Movement

Screen and Neuromuscular Performance in Adolescent Team Sport Athletes. *Journal of Strength and Conditioning Research, 26*, 1. https://doi.org/10.1519/JSC.0000000000003142

Stępień-Słodkowska, M., Ficek, K., Eider, J., Leońska-Duniec, A., Maciejewska-Karłowska, A., Sawczuk, M., Zarębska, A., Jastrzębski, Z., Grenda, A., Kotarska, K., Cięszczyk, P., Stępień-Słodkowska, M., Ficek, K., Eider, J., Leońska-Duniec, A., Maciejewska-Karłowska, A., Sawczuk, M., Zarębska, A., Jastrzębski, Z., … Cięszczyk, P. (2013). The +1245g/t polymorphisms in the collagen type I alpha 1 (col1a1) gene in polish skiers with anterior cruciate ligament injury. *Biology of Sport, 30*(1), 57–60. https://doi.org/10.5604/20831862.1029823

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A.-W., Churchill, R., Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., … Higgins, J. P. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ, 355*(October), i4919. https://doi.org/10.1136/bmj.i4919

Sterne, J., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rucker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ, 343*(jul22 1), d4002–d4002. https://doi.org/10.1136/bmj.d4002

Storm, E. E., & Kingsley, D. M. (1999). GDF5 coordinates bone and joint formation during digit development. *Developmental Biology, 209*(1), 11–27. https://doi.org/10.1006/dbio.1999.9241

Strandberg, S., Nordström, P., Lorentzon, R., & Lorentzon, M. (2003). Vitamin D receptor start codon polymorphism (FokI) is related to bone mineral density in healthy adolescent boys. *Journal of Bone and Mineral Metabolism, 21*(2), 109–113. https://doi.org/10.1007/s007740300018

Stranger, B. E., Forrest, M. S., Dunning, M., Ingle, C. E., Thorne, N., Redon, R., Bird, C. P., Grassi, A. De, Tyler-smith, C., Carter, N., Scherer, S. W., Tavaré, S., Deloukas, P., Hurles, M. E., Dermitzakis, E. T., & Bari, B. (2009). Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science, 315*(5813), 848–853. https://doi.org/10.1126/science.1136678.Relative

Strøyer, J., Hansen, L., & Klausen, K. (2004). Physiological Profile and Activity Pattern of Young Soccer Players during Match Play. *Medicine and Science in Sports and Exercise, 36*(1), 168–174. https://doi.org/10.1249/01.MSS.0000106187.05259.96

Sun, S. S., Schubert, C. M., Chumlea, W. C., Roche, A. F., Kulin, H. E., Lee, P. A., Himes, J. H., & Ryan, A. S. (2002). National Estimates of the Timing of Sexual Maturation and Racial Differences Among US Children. *Pediatrics, 110*(5), 911–919. https://doi.org/10.1542/peds.110.5.911

Suuriniemi, M., Kovanen, V., Mahonen, A., Alén, M., Wang, Q., Lyytikäinen, A., & Cheng, S. (2006). COL1A1 Sp1 polymorphism associates with bone density in early puberty. *Bone, 39*(3), 591–597. https://doi.org/https://doi.org/10.1016/j.bone.2006.02.053

Suuriniemi, M., Mahonen, A., Kovanen, V., Alén, M., & Cheng, S. (2003). Relation of Pvu II site polymorphism in the COL1A2 gene to the risk of fractures in prepubertal Finnish girls. *Physiological Genomics, 14*(3), 217–224. https://doi.org/10.1152/physiolgenomics.00070.2003

Swain, M., Kamper, S. J., Maher, C. G., Broderick, C., McKay, D., & Henschke, N. (2018). Relationship between growth, maturation and musculoskeletal conditions in adolescents: a systematic review. *British Journal of Sports Medicine, 52*(19), 1246–1252. https://doi.org/10.1136/bjsports-2017-098418

Taberner, M., O'Keefe, J., Flower, D., Phillips, J., Close, G., Cohen, D. D., Richter, C., & Carling, C. (2020). Interchangeability of position tracking technologies; can we merge the data? *Science and Medicine in Football, 4*(1), 76–81. https://doi.org/10.1080/24733938.2019.1634279

Tabor, H. K., Risch, N. J., & Myers, R. M. (2002). Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nature Reviews Genetics, 3*(5), 391–397. https://doi.org/10.1038/nrg796

Tajima, O., Ashizawa, N., Ishii, T., Amagai, H., Mashimo, T., Liu, L. J., Saitoh, S., Tokuyama, K., & Suzuki, M. (2000). Interaction of the effects between vitamin D receptor polymorphism and exercise training on bone metabolism. *Journal of Applied Physiology, 88*(4), 1271–1276. http://search.ebscohost.com/login.aspx?direct=true&db=s3h&AN=SPHS-661572&site=ehost-live

Tang, H., Quertermous, T., Rodriguez, B., Kardia, S. L. R., Zhu, X., Brown, A., Pankow, J. S., Province, M. A., Hunt, S. C., Boerwinkle, E., Schork, N. J., & Risch, N. J. (2005). Genetic Structure, Self-Identified Race/Ethnicity, and Confounding in Case-Control Association Studies. *The American Journal of Human Genetics, 76*(2), 268–275. https://doi.org/10.1086/427888

Tang, L., Cheng, G.-L., & Xu, Z.-H. (2013). Association between estrogen receptor alpha gene (ESR1) PvuII (C/T) and XbaI (A/G) polymorphisms and hip fracture risk: evidence from a meta-analysis. *PloS One, 8*(12), e82806. https://doi.org/10.1371/journal.pone.0082806

Tanner, J. M., & Whitehouse, R. H. (1976). Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. *Archives of Disease in Childhood, 51*(3), 170–179. https://doi.org/10.1136/adc.51.3.170

Tanner, J. M., Whitehouse, R. H., & Takaishi, M. (1966). Standards from birth to maturity for height, weight, height velocity, and weight velocity: British children, 1965. I. *Archives of Disease in Childhood, 41*(219), 454–471. https://doi.org/10.1136/adc.41.219.454

Tanrıverdi, H., Evrengul, H., Kaftan, A., Dursunoglu, D., Turgut, S., Akda, B., & Kılıç, M. (2005). Effects of Angiotensin-Converting Enzyme Polymorphism on Aortic Elastic Parameters in Athletes. *Cardiology, 104*(3), 113–119. https://doi.org/10.1159/000087243

Tarumi, T., Yamabe, T., Fukuie, M., Kimura, R., Zhu, D. C., Ohyama-Byun, K., Maeda, S., & Sugawara, J. (2021). Proximal Aortic Compliance in Young Male Endurance Athletes: An MRI Study. *Medicine and Science in Sports and Exercise, 53*(3), 543–550. https://doi.org/10.1249/MSS.0000000000002508

Taylor, R. R., Mamotte, C. D. S., Fallon, K., & van Bockxmeer, F. M. (1999). Elite athletes and the gene for angiotensin-converting enzyme. *Journal of Applied Physiology, 87*(3), 1035–1037. https://doi.org/10.1152/jappl.1999.87.3.1035

Tee, J. C., McLaren, S. J., & Jones, B. (2020). Sports Injury Prevention is Complex: We Need to Invest in Better Processes, Not Singular Solutions. *Sports Medicine, 50*(4), 689–702. https://doi.org/10.1007/s40279-019-01232-4

Teunissen, J. W., Rommers, N., Pion, J., Cumming, S. P., Rössler, R., D'Hondt, E., Lenoir, M., Savelsbergh, G. J. P., & Malina, R. M. (2020). Accuracy of maturity prediction equations in individual elite male football players. *Annals of Human Biology, 47*(4), 409–416. https://doi.org/10.1080/03014460.2020.1783360

Thakkinstian, A., McElduff, P., D'Este, C., Duffy, D., & Attia, J. (2005). A method for meta-analysis of molecular association studies. *Statistics in Medicine, 24*(9), 1291–1306. https://doi.org/10.1002/sim.2010

Tharabenjasin, P., Pabalan, N., & Jarjanazi, H. (2019). Association of the ACTN3 R577X (rs1815739) polymorphism with elite power sports: A meta-analysis. *PloS One, 14*(5), e0217390. https://doi.org/10.1371/journal.pone.0217390

The 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68–74. https://doi.org/10.1038/nature15393

Thomas, D. C., & Witte, J. S. (2002). Point: population stratification: a problem for case-control studies of candidate-gene associations? *Cancer Epidemiology, Biomarkers & Prevention : A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology, 11*(6), 505–512. http://www.ncbi.nlm.nih.gov/pubmed/12050090

Thomis, M. A. I., Huygens, W., Heuninckx, S., Chagnon, M., Maes, H. H. M., Claessens, A. L., Vlietinck, R., Bouchard, C., & Beunen, G. P. (2004). Exploration of myostatin polymorphisms and the angiotensin-converting enzyme insertion/deletion genotype in responses of human muscle to strength training. *European Journal of Applied Physiology, 92*(3), 267–274. https://doi.org/10.1007/s00421-004-1093-6

Thompson, P. D., Arena, R., Riebe, D., & Pescatello, L. S. (2013). ACSM's New Preparticipation Health Screening Recommendations from ACSM's Guidelines for Exercise Testing and Prescription, Ninth Edition. In *Current Sports Medicine Reports* (Ninth, Vol. 12, Issue 4). Lippincott Williams & Wilkins. https://doi.org/10.1249/JSR.0b013e31829a68cf

Tierney, P. J., Young, A., Clarke, N. D., & Duncan, M. J. (2016). Match play demands of 11 versus 11 professional football using Global Positioning System tracking: Variations across common playing formations. *Human Movement Science, 49*, 1–8. https://doi.org/10.1016/j.humov.2016.05.007

Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J., & Richards, J. B. (2018). Genetic architecture: the shape of the genetic contribution to human traits and disease. *Nature Reviews Genetics, 19*(2), 110–124. https://doi.org/10.1038/nrg.2017.101

Tinkle, B. T., & Wenstrup, R. J. (2005). A genetic approach to fracture epidemiology in childhood. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics, 139C*(1), 38–54. https://doi.org/10.1002/ajmg.c.30073

Tobina, T., Michishita, R., Yamasawa, F., Zhang, B., Sasaki, H., Tanaka, H., Saku, K., & Kiyonaga, A. (2010). Association between the angiotensin I-converting enzyme gene insertion/deletion polymorphism and endurance running speed in Japanese runners. *The Journal of Physiological Sciences, 60*(5), 325–330. https://doi.org/10.1007/s12576-010-0100-4

Torres-Costoso, A., López-Muñoz, P., Martínez-Vizcaíno, V., Álvarez-Bueno, C., & Cavero-Redondo, I. (2020). Association Between Muscular Strength and Bone Health from Children to Young Adults: A Systematic Review and Meta-analysis. *Sports Medicine, 50*(6), 1163–1190. https://doi.org/10.1007/s40279-020-01267-y

Toth, M., Chvyrkova, I., Bernardo, M. M., Hernandez-Barrantes, S., & Fridman, R. (2003). Pro-MMP-9 activation by the MT1-MMP/MMP-2 axis and MMP-3: role of TIMP-2 and plasma membranes. *Biochemical and Biophysical Research Communications, 308*(2), 386–395. https://doi.org/10.1016/S0006-291X(03)01405-0

Towlson, C., Salter, J., Ade, J. D., Enright, K., Harper, L. D., Page, R. M., & Malone, J. J. (2021). Maturity-associated considerations for training load, injury risk, and physical performance in youth soccer: One size does not fit all. *Journal of Sport and Health Science, 10*(4), 403–412. https://doi.org/10.1016/j.jshs.2020.09.003

Trajanoska, K., Morris, J. A., Oei, L., Zheng, H.-F., Evans, D. M., Kiel, D. P., Ohlsson, C., Richards, J. B., & Rivadeneira, F. (2018). Assessment of the genetic and clinical determinants of fracture risk: genome wide association and mendelian randomisation study. *BMJ, 362*, k3225. https://doi.org/10.1136/bmj.k3225

Tsianos, G., Sanders, J., Dhamrait, S., Humphries, S., Grant, S., & Montgomery, H. (2004). The ACE gene insertion/deletion polymorphism and elite endurance swimming. *European Journal of Applied Physiology, 92*(3), 360–362. https://doi.org/10.1007/s00421-004-1120-7

Tuncel, G., Temel, S. G., & Ergoren, M. C. (2019). Strong association between VDR FokI (rs2228570) gene variant and serum vitamin D levels in Turkish Cypriots. *Molecular Biology Reports, 46*(3), 3349–3355. https://doi.org/10.1007/s11033-019-04796-6

Tzaphlidou, M. (2008). Bone Architecture: Collagen Structure and Calcium/Phosphorus Maps. *Journal of Biological Physics, 34*(1–2), 39–49. https://doi.org/10.1007/s10867-008-9115-y

Utesch, T., Bardid, F., Büsch, D., & Strauss, B. (2019). The Relationship Between Motor Competence and Physical Fitness from Early Childhood to Early Adulthood: A Meta-Analysis. *Sports Medicine, 49*(4), 541–551. https://doi.org/10.1007/s40279-019-01068-y

Valdes, A. M., Evangelou, E., Kerkhof, H. J. M., Tamm, A., Doherty, S. A., Kisand, K., Tamm, A., Kerna, I., Uitterlinden, A., Hofman, A., Rivadeneira, F., Cooper, C., Dennison, E. M., Zhang, W., Muir, K. R., Ioannidis, J. P. A., Wheeler, M., Maciewicz, R. A., van Meurs, J. B., … Doherty, M. (2011). The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance. *Annals of the Rheumatic Diseases, 70*(5), 873–875. https://doi.org/10.1136/ard.2010.134155

Välimäki, V.-V., Alfthan, H., Lehmuskallio, E., Löyttyniemi, E., Sahi, T., Suominen, H., & Välimäki, M. J. (2005). Risk factors for clinical stress fractures in male military recruits: a prospective cohort study. *Bone, 37*(2), 267–273. https://doi.org/10.1016/j.bone.2005.04.016

Van Der Sluis, A., Elferink-Gemser, M. T., Coelho-E-Silva, M. J., Nijboer, J. A., Brink, M. S., & Visscher, C. (2014). Sport injuries aligned to Peak Height Velocity in talented pubertal soccer players. *International Journal of Sports Medicine, 35*(4), 351–355. https://doi.org/10.1055/s-0033-1349874

van Ginkel, S., Ruoss, S., Valdivieso, P., Degens, H., Waldron, S., de Haan, A., & Flück, M. (2016). ACE inhibition modifies exercise-induced pro-angiogenic and mitochondrial gene transcript expression. *Scandinavian Journal of Medicine & Science in Sports, 26*(10), 1180–1187. https://doi.org/10.1111/sms.12572

Van Hooren, B., & De Ste Croix, M. (2020). Sensitive Periods to Train General Motor Abilities in Children and Adolescents. *Strength and Conditioning Journal, 42*(6), 7–14. https://doi.org/10.1519/SSC.0000000000000545

van Wilgen, C. P., & Verhagen, E. A. L. M. (2012). A qualitative study on overuse injuries: The beliefs of athletes and coaches. *Journal of Science and Medicine in Sport, 15*(2), 116–121. https://doi.org/10.1016/j.jsams.2011.11.253

Varley, I., Greeves, J. P., Sale, C., Friedman, E., Moran, D. S., Yanovich, R., Wilson, P. J., Gartland, A., Hughes, D. C., Stellingwerff, T., Ranson, C., Fraser, W. D., & Gallagher, J. A. (2016). Functional polymorphisms in the P2X7 receptor gene are associated with stress fracture injury. *Purinergic Signalling, 12*(1), 103–113. https://doi.org/10.1007/s11302-016-9495-6

Varley, I., Hughes, D. C., Greeves, J. P., Stellingwerff, T., Ranson, C., Fraser, W. D., & Sale, C. (2015). RANK/RANKL/OPG pathway: genetic associations with stress fracture period prevalence in elite athletes. *Bone, 71*, 131–136. https://doi.org/10.1016/j.bone.2014.10.004

Varley, I., Hughes, D. C., Greeves, J. P., Stellingwerff, T., Ranson, C., Fraser, W. D., & Sale, C. (2018). The association of novel polymorphisms with stress fracture injury in Elite Athletes: Further insights from the SFEA cohort. *Journal of Science and Medicine in Sport, 21*(6), 564–568. https://doi.org/10.1016/j.jsams.2017.10.038

Varley, I., Patel, S., Williams, A. G., & Hennis, P. J. (2017). The current use, and opinions of elite athletes and support staff in relation to genetic testing in elite sport within the UK. *Biology of Sport, 35*(1), 13–19. https://doi.org/10.5114/biolsport.2018.70747

Vaughan, D., Brogioli, M., Maier, T., White, A., Waldron, S., Rittweger, J., Toigo, M., Wettstein, J., Laczko, E., & Flück, M. (2016). The Angiotensin Converting Enzyme Insertion/Deletion Polymorphism Modifies Exercise-Induced Muscle Metabolism. *PLOS ONE, 11*(3), e0149046. https://doi.org/10.1371/journal.pone.0149046

Venckunas, T., Skurvydas, A., Brazaitis, M., Kamandulis, S., Snieckus, A., & Moran, C. N. (2012). Human alpha-actinin-3 genotype association with exercise-induced muscle damage and the repeated-bout effect. *Applied Physiology, Nutrition, and Metabolism = Physiologie Appliquee, Nutrition et Metabolisme, 37*(6), 1038–1046. https://doi.org/10.1139/h2012-087

Vincent, B., Windelinckx, A., Nielens, H., Ramaekers, M., Van Leemputte, M., Hespel, P., & Thomis, M. A. (2010). Protective role of alpha-actinin-3 in the response to an acute eccentric exercise bout. *Journal of Applied Physiology (Bethesda, Md. : 1985), 109*(2), 564–573. https://doi.org/10.1152/japplphysiol.01007.2009

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era--concepts and misconceptions. *Nature Reviews. Genetics, 9*(4), 255–266. https://doi.org/10.1038/nrg2322

Vlahovich, N., Fricker, P. A., Brown, M. A., & Hughes, D. (2016). Ethics of genetic testing and research in sport: a position statement from the Australian Institute of Sport. *British Journal of Sports Medicine, 51*(1), bjsports-2016-096661. https://doi.org/10.1136/bjsports-2016-096661

Wackerhage, H. (2014). Molecular Exercise Physiology. Routledge. https://doi.org/10.4324/9780203132142

Wackerhage, H., Miah, A., Harris, R., Montgomery, H., & Williams, A. G. (2009). Genetic research and testing in sport and exercise science: a review of the issues. *Journal of Sports Sciences, 27*(11), 1109–1116. https://doi.org/10.1080/02640410903114364

Wackerhage, H., & Schoenfeld, B. J. (2021). Personalized, Evidence-Informed Training Plans and Exercise Prescriptions for Performance, Fitness and Health. *Sports Medicine, 51*(9), 1805–1813. https://doi.org/10.1007/s40279-021-01495-w

Wackerhage, H., & Woods, N. (2002). Exercise-induced signal transduction and gene regulation in skeletal muscle. *Journal of Sports Science & Medicine, 1*(4), 103–114. http://search.ebscohost.com/login.aspx?direct=true&db=s3h&AN=SPHS-865670&site=ehost-live

Waldron, M., Harding, J., Barrett, S., & Gray, A. (2021). A New Foot-Mounted Inertial Measurement System in Soccer: Reliability and Comparison to Global Positioning Systems for Velocity Measurements During Team Sport Actions. *Journal of Human Kinetics, 77*(1), 37–50. https://doi.org/10.2478/hukin-2021-0010

Waldron, M., & Murphy, A. (2013). A Comparison of Physical Abilities and Match Performance Characteristics Among Elite and Subelite Under-14 Soccer Players. *Pediatric Exercise Science, 25*(3), 423–434. https://doi.org/10.1123/pes.25.3.423

Walsh, S., Liu, D., Metter, E. J., Ferrucci, L., & Roth, S. M. (2008). ACTN3 Genotype is Associated with Muscle Phenotypes in Women across the Adult Age Span. *Medicine & Science in Sports & Exercise, 40*(Supplement), S191. https://doi.org/10.1249/01.mss.0000322289.16416.12

Wang, C., Li, H., Chen, K., Wu, B., & Liu, H. (2017). Association of polymorphisms rs1800012 in COL1A1 with sports-related tendon and ligament injuries: a meta-analysis. *Oncotarget, 8*(16), 27627–27634. https://doi.org/10.18632/oncotarget.15271

Wang, C., Vargas, J. T., Stokes, T., Steele, R., & Shrier, I. (2020). Analyzing Activity and Injury: Lessons Learned from the Acute:Chronic Workload Ratio. *Sports Medicine, 50*(0123456789), 1243–1254. https://doi.org/10.1007/s40279-020-01280-1

Wang, G., Mikami, E., Chiu, L.-L., DE Perini, A., Deason, M., Fuku, N., Miyachi, M., Kaneoka, K., Murakami, H., Tanaka, M., Hsieh, L.-L., Hsieh, S., Caporossi, D., Pigozzi, F., Hilley, A., Lee, R., Galloway, S., Gulbin, J., Rogozkin, V., … Pitsiladis, Y. (2013). Association analysis of ACE and ACTN3 in elite Caucasian and East Asian swimmers. *Medicine and Science in Sports and Exercise, 45*(5), 892–900. https://doi.org/10.1249/MSS.0b013e31827c501f

Wang, Q., & Seeman, E. (2009). Growth-related cortical fragility at metaphyseal regions. *IBMS BoneKEy, 6*(11), 429–433. https://doi.org/10.1138/20090408

Wang, Q., Wang, X.-F., Iuliano-Burns, S., Ghasem-Zadeh, A., Zebaze, R., & Seeman, E. (2010). Rapid growth produces transient cortical weakness: A risk factor for metaphyseal fractures during puberty. *Journal of Bone and Mineral Research, 25*(7), 1521–1526. https://doi.org/10.1002/jbmr.46

Wang, T.-T., Tavera-Mendoza, L. E., Laperriere, D., Libby, E., Burton MacLeod, N., Nagai, Y., Bourdeau, V., Konstorum, A., Lallemant, B., Zhang, R., Mader, S., & White, J. H. (2005). Large-Scale in Silico and Microarray-Based Identification of Direct 1,25-Dihydroxyvitamin D3 Target Genes. *Molecular Endocrinology, 19*(11), 2685–2695. https://doi.org/10.1210/me.2005-0106

Waterman, B. R., Gun, B., Bader, J. O., Orr, J. D., & Belmont, P. J. (2016). Epidemiology of Lower Extremity Stress Fractures in the United States Military. *Military Medicine, 181*(10), 1308–1313. https://doi.org/10.7205/MILMED-D-15-00571

Watson, J. D., & Crick, F. H. (1953). The structure of DNA. *Cold Spring Harbor Symposia on Quantitative Biology, 18*, 123–131. https://doi.org/10.1101/SQB.1953.018.01.020

Webborn, N., Williams, A. G., McNamee, M., Bouchard, C., Pitsiladis, Y., Ahmetov, I., Ashley, E., Byrne, N., Camporesi, S., Collins, M., Dijkstra, P., Eynon, N., Fuku, N., Garton, F. C., Hoppe, N., Holm, S., Kaye, J., Klissouras, V., Lucia, A., … Wang, G. (2015). Direct-to-consumer genetic testing for predicting sports performance and talent identification: Consensus statement. *British Journal of Sports Medicine, 49*(23), 1486–1491. https://doi.org/10.1136/bjsports-2015-095343

Wenstrup, R. J., Florer, J. B., Brunskill, E. W., Bell, S. M., Chervoneva, I., & Birk, D. E. (2004). Type V collagen controls the initiation of collagen fibril assembly. *Journal of Biological Chemistry, 279*(51), 53331–53337. https://doi.org/10.1074/jbc.M409622200

Wenstrup, R. J., Florer, J. B., Davidson, J. M., Phillips, C. L., Pfeiffer, B. J., Menezes, D. W., Chervoneva, I., & Birk, D. E. (2006). Murine Model of the Ehlers-Danlos Syndrome. *Journal of Biological Chemistry, 281*(18), 12888–12895. https://doi.org/10.1074/jbc.M511528200

Wentz, L., Liu, P.-Y., Haymes, E., & Ilich, J. Z. (2011). Females have a greater incidence of stress fractures than males in both military and athletic populations: a systemic review. *Military Medicine, 176*(4), 420–430. https://doi.org/10.7205/milmed-d-10-00322

Wik, E. H., Lolli, L., Chamari, K., Materne, O., Di Salvo, V., Gregson, W., & Bahr, R. (2020a). Injury patterns differ with age in male youth football: a four-season prospective study of 1111 time-loss injuries in an elite national academy. *British Journal of Sports Medicine, 55*(14), 794–800. https://doi.org/10.1136/bjsports-2020-103430

Wik, E. H., Martínez-Silván, D., Farooq, A., Cardinale, M., Johnson, A., & Bahr, R. (2020b). Skeletal maturation and growth rates are related to bone and growth plate injuries in adolescent athletics. *Scandinavian Journal of Medicine & Science in Sports, 30*(5), 894–903. https://doi.org/10.1111/sms.13635

Williams, A. G., Day, S. H., Folland, J. P., Gohlke, P., Dhamrait, S., & Montgomery, H. E. (2005). Circulating angiotensin converting enzyme activity is correlated with muscle strength. *Medicine and Science in Sports and Exercise, 37*(6), 944–948. https://doi.org/10.1249/01.mss.0000166577.42935.4e

Williams, A. G., & Folland, J. P. (2008). Similarity of polygenic profiles limits the potential for elite human physical performance. *The Journal of Physiology, 586*(1), 113–121. https://doi.org/10.1113/jphysiol.2007.141887

Williams, A. G., Rayson, M., Jubb, M., World, M., Woods, D., Hayward, M., Martin, J., Humphries, S., & Montgomery, H. (2000). The ACE gene and muscle performance. *Nature, 403*, 614. https://doi.org/10.1038/35001141

Williams, A. G., Wackerhage, H., & Day, S. H. (2016). Genetic Testing for Sports Performance, Responses to Training and Injury Risk: Practical and Ethical Considerations. In *Medicine and Sport Science* (Vol. 61, pp. 105–119). https://doi.org/10.1159/000445244

Williams, A. M., & Reilly, T. (2000). Talent identification and development in soccer. *Journal of Sports Sciences, 18*(9), 657–667. https://doi.org/10.1080/02640410050120041

Windelinckx, A., De Mars, G., Beunen, G., Aerssens, J., Delecluse, C., Lefevre, J., & Thomis, M. A. I. (2007). Polymorphisms in the vitamin D receptor gene are associated with muscle strength in men and women. *Osteoporosis International, 18*(9), 1235–1242. https://doi.org/10.1007/s00198-007-0374-4

Wong, P.-L., Chamari, K., Dellal, A., & Wisløff, U. (2009). Relationship Between Anthropometric and Physiological Characteristics in Youth Soccer Players. *Journal of Strength and Conditioning Research, 23*(4), 1204–1210. https://doi.org/10.1519/JSC.0b013e31819f1e52

Woods, D., Hickman, M., Jamshidi, Y., Brull, D., Vassiliou, V., Jones, A., Humphries, S., & Montgomery, H. (2001). Elite swimmers and the D allele of the ACE I/D polymorphism. *Human Genetics, 108*(3), 230–232. https://doi.org/10.1007/s004390100466

Wormhoudt, R., Savelsbergh, G. J. P., Teunissen, J. W., & Davids, K. (2017). The Athletic Skills Model: Optimizing Talent Development Through Movement. Education (1st ed.). Routledge. https://doi.org/10.4324/9781315201474

Yamin, C., Amir, O., Sagiv, M., Attias, E., Meckel, Y., Eynon, N., Sagiv, M., & Amir, R. E. (2007). ACE ID genotype affects blood creatine kinase response to eccentric exercise. *Journal of Applied Physiology, 103*(6), 2057–2061. https://doi.org/10.1152/japplphysiol.00867.2007

Yang, N., Schindeler, A., McDonald, M. M., Seto, J. T., Houweling, P. J., Lek, M., Hogarth, M., Morse, A. R., Raftery, J. M., Balasuriya, D., MacArthur, D. G., Berman, Y., Quinlan, K. G. R., Eisman, J. A., Nguyen, T. V., Center, J. R., Prince, R. L., Wilson, S. G., Zhu, K., … North, K. N. (2011). α-Actinin-3 deficiency is associated with reduced bone mass in human and mouse. *Bone, 49*(4), 790–798. https://doi.org/10.1016/j.bone.2011.07.009

Yanovich, R., Friedman, E., Milgrom, R., Oberman, B., Freedman, L., & Moran, D. S. (2012). Candidate gene analysis in israeli soldiers with stress fractures. *Journal of Sports Science & Medicine, 11*(1), 147–155. http://search.ebscohost.com/login.aspx?direct=true&db=s3h&AN=74441987&site=ehost-live

Yilmaz, D., Ersoy, B., Bilgin, E., Gümüşer, G., Onur, E., & Pinar, E. D. (2005). Bone mineral density in girls and boys at different pubertal stages: relation with gonadal steroids, bone formation markers, and growth parameters. *Journal of Bone and Mineral Metabolism, 23*(6), 476–482. https://doi.org/10.1007/s00774-005-0631-6

Yin, Y., & Wang, Y. (2017). Association of BMP-14 rs143383 ploymorphism with its susceptibility to osteoarthritis. *Medicine, 96*(42), e7447. https://doi.org/10.1097/MD.0000000000007447

Zhang, B., Tanaka, H., Shono, N., Miura, S., Kiyonaga, A., Shindo, M., & Saku, K. (2003). The I allele of the angiotensin-converting enzyme gene is associated with an increased percentage of slow-twitch type I fibers in human skeletal muscle. *Clinical Genetics, 63*(2), 139–144. http://www.ncbi.nlm.nih.gov/pubmed/12630962

Zhao, L., Chang, Q., Huang, T., & Huang, C. (2016). Prospective cohort study of the risk factors for stress fractures in Chinese male infantry recruits. *The Journal of International Medical Research, 44*(4), 787–795. https://doi.org/10.1177/0300060516639751

Zondervan, K. T., & Cardon, L. R. (2007). Designing candidate gene and genome-wide case–control association studies. *Nature Protocols, 2*(10), 2492–2501. https://doi.org/10.1038/nprot.2007.366

Zouhal, H., Coso, J. Del, Jayavel, A., Tourny, C., Ravé, G., Jebabli, N., Clark, C. C. T., Barthélémy, B., Hackney, A. C., & Abderrahman, A. Ben. (2021). Association between ACTN3 R577X genotype and risk of non-contact injury in trained athletes: A systematic review. *Journal of Sport and Health Science, 3*. https://doi.org/10.1016/j.jshs.2021.07.003

CHAPTER 9: Appendices

9.1 Appendix 1 – Ethical approval

Ethical approval for the research in this thesis was provided by the St Mary's University Twickenham Ethics Sub-Committee (No. SMEC_2019-20_002).

St Mary's University Twickenham London

4 November 2019 SMEC_2019-20_002

Ed Ryan-Moore (SHAS): 'Evaluation of selected genes and environmental factors on the development of elite youth footballers'

Dear Ed

University Ethics Sub-Committee

Thank you for re-submitting your ethics application for consideration.

I can confirm that all required amendments have been made and that you therefore have ethical approval to undertake your research.

Yours sincerely

Dr Jamie North
Chair, Ethics Sub-Committee

Cc Dr Mark Waldron, Dr Yiannis Mavrommatis

St Mary's University, Waldegrave Road, Strawberry Hill, Twickenham, London TW1 4SX
Switchboard 020 8240 4000, Fax 020 8240 4255, www.stmarys.ac.uk

9.2 Appendix 2 – Participant information, consent and assent forms

9.2.1 Participant information and consent form (age 8-12 years)

How do genes affect injury and physical performance?

My name is Ed Ryan-Moore, I am an U9-16s fitness coach at Fulham Football Club and I am collecting some information to answer a question I have.

I would like to find out some information about your genes. Genes have instructions on how to make you!

I would like to find out how these genes that make you affect injury and physical performance.

If you want to take part, I will ask you to spit into a plastic tube so I can look at your genes. Then I will keep a record of all the times you get injured along with your physical testing and height and weight information.

If you want to stop at any point just tell me.

It is up to you if you would like to take part. YOU DON'T HAVE TO!

If you do, please fill in the form below.

I would like to provide some saliva in a plastic tube Edward Ryan-Moore will give me and to take part in the project with all my injuries, physical testing, height and weight recorded.

Please write your name……

Please return this form to Ed Ryan-Moore as soon as possible

9.2.2 Participant information and consent form (age 13-17 years)

St Mary's University
Waldegrave Road
Strawberry Hill
Twickenham
TW1 4SX

Project Title: ***Evaluation of Selected Genes and Environmental Factors on the Development of Elite Youth Footballers.***

My name is Ed Ryan-Moore, I am a U9-16s fitness coach at Fulham Football Club and I am doing some research about how your genes, which have instructions on how to make you, affect your risk of injury and physical testing scores.

We already measure your injuries and physical testing scores but to look at your genes I will ask you to provide a saliva in a plastic tube. I will then compare the differences in your genes to your injuries and physical testing scores.

I will be the only person who knows what your genes are, **and I won't tell anyone else at Fulham Football Club about these**. You do not have to participate, and it will not affect your involvement with Fulham Football Club in any way. If you decide to take part and want to stop at any point just tell me.

It is up to you if you would like to take part. If you do, please fill in the form below.

I would like to provide saliva in a plastic tube that Edward Ryan-Moore will give me and to take part in the project with all my injuries, physical testing, height and weight recorded.

Please write your name……

Please return this form to your Ed Ryan-Moore as soon as possible

9.2.3 Adult participant information sheet

St Mary's University
Waldegrave Road
Strawberry Hill
Twickenham
TW1 4SX

Project Title: Evaluation of Selected Genes and Environmental Factors on the Development of Elite Youth Footballers.

INFORMATION SHEET FOR PARTICIPANT PARENTS / GAURDIANS

Before you take part in this study you must understand why the research is being carried out and exactly what it will involve. Please take the time to read the following information sheet carefully and do not hesitate to ask if there is anything that is unclear or that you would like more information on. Take time to decide whether or not you wish to take part.

The Research Project

Injuries can result in substantial loss of time from training and matches to elite footballers. Physical traits can help to guard against injuries, and we all have different innate strengths and weaknesses. Understanding the potential influence that genes may have on the risk of injury and physical characteristics can help provide targeted interventions which are specific to the individual to reduce the injury risk and maximise physical potential. The purpose of this study is to see how selected genes and environmental factors affect injuries and physical performance.

I would like to extend an invitation to participate in this study as part of my PhD studies. The main researcher will be myself, Edward Ryan-Moore U9-16s Sports Scientist at Fulham Football Club and PhD Student at St Mary's University, under the supervision and guidance of Dr Mark Waldron and Dr Yiannis Mavrommatis of St Mary's University.

Personal injury, physical testing and anthropometric data are routinely collected by Fulham Football Club as part of normal practise. Fulham Football Club hold personal injury, physical testing and anthropometric data which will be analysed in this study. **The genetic information collected as part of this study will not be shared with anyone at Fulham Football Club**. Participation or non-participation in the study will have no impact on your involvement with Fulham Football Club. A summary of anonymised results will aim to be published as scientific literature and used towards the completion of my PhD research. Funding for my research is being provided by Fulham Football Club and St Mary's University. For further information please contact Edward Ryan-Moore via mobile (07859907536) or email (111453@live.stmarys.ac.uk)

Participation in the Research Project

You have been invited to take part in this study as you are a registered player at Fulham Football Club. **Participation is not a requirement** and will have no influence on your involvement with Fulham Football Club. If you agree to take part both you and your son will be asked to complete a consent form. You may withdraw from the project at any time, without reason, by contacting myself via email or mobile.

Should you agree to participate in the study all that will be required for your participation, beyond normal activities at Fulham Football Club, is to **provide a salivary sample.** This takes roughly 5 minutes and will be used for DNA extraction. The DNA will then be analysed in relation to injury and physical data. Injury data for your son along with other physical performance data is already recorded at Fulham Football Club so will not require any extra commitment from yourself beyond normal participation at Fulham Football Club. As the research develops the DNA sample provided may be used to analyse new genetic variations and/or comparative variables in the future. This will require no further time or commitment from you or your son and is included in consent to participate in this study.

There are no additional risks to your health or wellbeing on top of those associated with normal participation in a football and the findings of the genetic testing will be anonymised so that only the main researcher can identify individual results. Individual genetic information will be kept confidential. **The individual genetic results of this study will not be available to other Fulham Football Club staff and will have no influence on your treatment within the club without further consent.**

To ensure the saliva sample is viable for DNA extraction we ask that participants avoid eating, drinking caffeinated beverages, or using mouthwash or toothpaste in the 2 hours prior to sample collection.

The collective data from the study will be collated and analysed as part of a PhD project with the aim of publishing the findings in a scientific journal. A copy of the findings will be made available free of charge to all participants upon request. Taking part in the study will help us to gain a greater understanding of your specific genetic building blocks which may allow us to provide targeted interventions to reduce injury risk and maximise physical performance potential. In total, the amount of extra time given to this study will be no longer than 15 minutes to provide a sample correctly. A code will be generated which anonymises your son's data to all but the main investigator to ensure confidentiality.

YOU WILL BE GIVEN A COPY OF THIS FORM TO KEEP TOGETHER WITH A COPY OF YOUR CONSENT FORM

9.2.4 Adult participant consent form

St Mary's University Twickenham London

Name of Participant: _____

Title of the project: ***Evaluation of Selected Genes and Environmental Factors on the Development of Elite Youth Footballers.***

Main investigator and contact details: Edward Ryan-Moore
Mobile = 07859907536
Email = 111453@live.stmarys.ac.uk

Members of the research team: Dr Mark Waldron, Dr Jamie North and Dr Yiannis Mavrommatis.

1. I agree to taking part in the above research. I have read the Participant Information Sheet which is attached to this form. I understand what my role will be in this research, and all my questions have been answered to my satisfaction.
2. I understand that I am free to withdraw from the research at any time, for any reason and without prejudice.
3. I have been informed that the confidentiality of the information I provide will be safeguarded.
4. I am free to ask any questions at any time before and during the study.
5. I have been provided with a copy of this form and the Participant Information Sheet.

Data Protection: I agree to the University processing personal data which I have supplied. I agree to the processing of such data for any purposes connected with the Research Project as outlined to me.

Name of parent (print)……

Signed……..…… Date…….……..…..

If you wish to withdraw from the research, please complete the form below and return to the main investigator named above.

Title of Project: _____

I WISH TO WITHDRAW FROM THIS STUDY

Name of Participant: _____

Signed: _____ Date: _____

9.2.5 Parental information sheet

St Mary's University
Waldegrave Road
Strawberry Hill
Twickenham
TW1 4SX

Project Title: Evaluation of Selected Genes and Environmental Factors on the Development of Elite Youth Footballers.

INFORMATION SHEET FOR PARTICIPANT PARENTS / GAURDIANS

Before you take part in this study you must understand why the research is being carried out and exactly what it will involve. Please take the time to read the following information sheet carefully and do not hesitate to ask if there is anything that is unclear or that you would like more information on. Take time to decide whether or not you wish to take part.

The Research Project

Injuries can result in substantial loss of time from training and matches to elite youth footballers. Physical traits can help to guard against injuries, and we all have different innate strengths and weaknesses. Understanding the potential influence that genes may have on the risk of injury and physical characteristics can help provide targeted interventions which are specific to the individual to reduce the injury risk and maximise physical potential. The purpose of this study is to see how selected genes and environmental factors affect injuries and physical performance.

I would like to extend an invitation to you and your son to participate in this study as part of my PhD studies. The main researcher will be myself, Edward Ryan-Moore U9-16s Sports Scientist at Fulham Football Club and PhD Student at St Mary's University, under the supervision and guidance of Dr Mark Waldron and Dr Yiannis Mavrommatis of St Mary's University.

Personal injury, physical testing and anthropometric data are routinely collected by Fulham Football Club as part of normal practise. Fulham Football Club hold personal injury, physical testing and anthropometric data which will be analysed in this study. **The genetic information collected as part of this study will not be shared with anyone at Fulham Football Club**. Participation or non-participation in the study will have no impact on your involvement with Fulham Football Club. A summary of anonymised results will aim to be published as scientific literature and used towards the completion of my PhD research. Funding for my research is being provided by Fulham Football Club and St Mary's University. For further information please contact Edward Ryan-Moore via mobile (07859907536) or email (111453@live.stmarys.ac.uk)

Participation in the Research Project

You have been invited to take part in this study as your son is a registered player at the Fulham Football Club Academy and is between the ages of 8 and 17 years of age. **Participation is not a**

requirement and will have no influence on your involvement with Fulham Football Club. If you agree to take part both you and your son will be asked to complete a consent form. You may withdraw from the project at any time, without reason, by contacting myself via email or mobile.

Should you agree to participate in the study all that will be required for your participation, beyond normal activities at Fulham Football Club, is for your son to **provide a salivary sample.** This takes roughly 5 minutes and will be used for DNA extraction. The DNA will then be analysed in relation to injury and physical data. Injury data for your son along with other physical performance data is already recorded at Fulham Football Club so will not require any extra from you or your son beyond normal participation at Fulham Football Club. As the research develops the DNA sample provided may be used to analyse new genetic variations and/or comparative variables in future research. This will require no further time or commitment from you or your son and is included in consent to participate in this study.

There are no additional risks to your son's health or wellbeing on top of those associated with normal participation in a football academy programme and the findings of the genetic testing will be anonymised so that only the main researcher can identify individual results. Individual genetic information will be kept confidential. **The individual genetic results of this study will not be available to other Fulham Football Club staff and will have no influence on your son's treatment within the club without further consent.**

To ensure the saliva sample is viable for DNA extraction we ask that participants avoid eating, drinking caffeinated beverages, or using mouthwash or toothpaste in the 2 hours prior to sample collection.

The collective data from the study will be collated and analysed as part of a PhD project with the aim of publishing the findings in a scientific journal. A copy of the findings will be made available free of charge to all participants upon request. Taking part in the study will help us to gain a greater understanding of your son's specific genetic building blocks which may allow us to provide targeted interventions to reduce injury risk and maximise physical performance potential. In total, the amount of extra time given to this study from your son will be no longer than 15 minutes to provide a sample correctly. A code will be generated which anonymises your son's data to all but the main investigator to ensure confidentiality.

YOU WILL BE GIVEN A COPY OF THIS FORM TO KEEP TOGETHER WITH A COPY OF YOUR CONSENT FORM

9.2.6 Parental consent form

St Mary's University Twickenham London

Name of Participant: _____

Title of the project: ***Evaluation of Selected Genes and Environmental Factors on the Development of Elite Youth Footballers.***

Main investigator and contact details: Edward Ryan-Moore
Mobile = 07859907536
Email = 111453@live.stmarys.ac.uk

Members of the research team: Dr Mark Waldron, Dr Jamie North and Dr Yiannis Mavrommatis.

1. I agree to my child taking part in the above research. I have read the Participant Information Sheet which is attached to this form. I understand what my child's role will be in this research, and all my questions have been answered to my satisfaction.
2. I understand that I am free to withdraw my child from the research at any time, for any reason and without prejudice.
3. I have been informed that the confidentiality of the information I and my child provide will be safeguarded.
4. I am free to ask any questions at any time before and during the study.
5. I have been provided with a copy of this form and the Participant Information Sheet.

Data Protection: I agree to the University processing personal data which I and my child have supplied. I agree to the processing of such data for any purposes connected with the Research Project as outlined to me.

Name of parent (print)……..

Signed……..…… Date…….....

If you wish to withdraw your child from the research, please complete the form below and return to the main investigator named above.

Title of Project: _____

I WISH TO WITHDRAW MY CHILD FROM THIS STUDY

Name of Participant: _____

Name of Parent _____

Signed: _____ Date: _____