Genetic Variation, Protein Composition and Potential Influences on Tendon Properties in Humans

B.P. Foster*, C.I. Morse, G.L. Onambele, I.I. Ahmetov and A.G. Williams

Institute for performance research, Manchester Metropolitan University, Crewe, UK

Abstract: Sequence variations in genes that code for proteins involved in homeostatic processes within tendons may influence tendon mechanical properties. Since variants of the four genes COL5A1, TNC, MMP3 and GDF5 have been implicated in the aetiopathogenesis of tendinopathies, which is ultimately characterised by abnormal structural and regulatory processes, sequence variations in these four genes may also influence how the tendon functions mechanically, even in the absence of tendinopathy. For example, two reports of association between variation in the COL5A1 gene and measures of flexibility complement reported associations between genotype and incidence of tendinopathy. Non-genetic factors such as age, body mass and physical activity status influence risk of tendon injury and physical performance potential independently from genomics, and also in gene-environment interactions. However, these non-genetic factors are often not considered in genetic association studies, probably due to their retrospective nature. Further research examining COL5A1, TNC, MMP3 and GDF5, as well as other genes that may influence the maintenance of tendon homeostasis such as COL1A1 which regulates the production of collagen type 1, the most abundant structural component of tendon is encouraged. Establishing the genetic basis of tendon properties in asymptomatic populations may advance understanding of some aspects relevant to physical performance and of the aetiology of tendinopathies. To improve understanding, accurate and reproducible assessments of tendon properties are required. However, no valid and reliable assessments of tendon properties, such as those involving in vivo ultrasound imaging techniques, have yet been applied to genetic association studies in humans.

Keywords: Genetic association studies, humans, sequence variants, tendinopathies, tendon, tendon properties.

1. TENDON PROPERTIES

Historically, tendons were considered bands of connective tissue with relatively no dynamic function [1-2]. More recently, numerous studies have shown that tendons provide an integral interface for transmission of forces from muscle to the bone in order to produce moments about joints, hence the mechanical properties of tendons determine the degree of joint motion in direct response to these forces [3]. From a biomechanical standpoint, external loads that act on the body are resisted by internal structures such as tendons, which undergo deformation. The degree of deformation of strain produced is related to the stress caused by these external loads and the material that it acts upon. Thus, knowledge of the mechanical properties of musculoskeletal tissues, including tendons, can assist in understanding injury risk and physical performance capabilities. The primary mechanical properties of relevance to human physical performance [4-8] include stiffness (the tendon force-displacement relationship) and the elastic modulus (the slope of the stress-strain curve in the elastic deformation region).

A biomechanical viewpoint only provides us with an analysis of forces and their ‘action’ or ‘effects’ on such structures and their function. To understand the mechanical properties of tendons, one needs to explore its structure - i.e. both its global characteristic dimensions (e.g. cross-sectional area and length) as well as its internal structures (e.g. cross-link density) - and its dynamic function in greater detail, by examining its biochemical components. The structure or morphology of tendon, and thus its function, is controlled by cells called tenocytes that maintain the extracellular matrix (ECM) [9] and ultimately its material and mechanical properties.

The tenocytes ‘sense’ and respond to mechanical loads deriving from external forces. The sensitivity of this response has recently been found to be mediated by a ‘mechanostat’ set point, in vitro [9]. This preset threshold is governed by complex interactions between the cell’s cytoskeleton and the ECM [9-13], ultimately giving rise to gene regulation and control over the tendon’s structure and function at a molecular level. Fundamentally, the expression of functional gene products or proteins is regulated, which provides the basis for maintenance and changes in structure and function. Mechanical signalling from external loads [9, 11, 12, 14, 15] causes gene expression of various proteins involved in tendon homeostasis to vary greatly between and within animal and human populations. However, even if mechanical loading is controlled, the abundance of various proteins can still vary. This is where genetic variation is likely to influence observed/measurable differences in protein content, and thus material and mechanical properties.

Tendon pathologies or tendinopathies (including tendinosis and tendinitis) are primarily degenerative conditions that may or may not be associated with signs of inflammation [16-17]. Since research in this area is relatively
extensive, reference to research studies on tendinopathies will form a small element of this review. For example, recent work has associated tendinopathies with genetic variants in proteins that serve important structural and functional roles in tendon. However, in this review, attempts will also be made to consider how those same molecular characteristics may influence tendon mechanical properties *per se*. That is to say, the same gene variants and differential gene expression of these same proteins may directly influence tendon mechanical properties either with or quite separately from their influence on incidence of injury. Indeed, there is conflicting evidence from some pathologic studies regarding mechanical properties. Some studies report no significant difference in mechanical properties in patients with tendon pathology from healthy subjects [18-22], but other studies reporting significant weakening of material and altered mechanical tendon properties, with both increased stiffness [23] and decreased stiffness [24-26] observed in patients. The inconsistent results regarding human *in vivo* mechanical properties between studies on tendon pathology may be a consequence of different tendon types under investigation, (total collagen levels vary between tendon types in normal tendons [27, 28]) as well as varying subject age and sex. The latter two factors of age [29-32] and sex [33-35] are frequently reported to be highly influential on tendon material and mechanical properties.

Seven genes will be discussed in this review. The seven genes and how they relate to the five tendon proteins they produce are shown in Table 1. Four of the genes listed in Table 1 (*COL5A1, TNC, MMP3, GDF5*) have genetic variants reported in recent tendinopathy studies that may predispose individuals to such conditions [36-41]. In fact, all the genes that contain sequence variants shown to be associated with tendon pathology to date, encode for proteins directly involved in biological processes within tendons. Fig. (1) shows the key structural proteins found in tendon, including those associated with tendon pathologies or musculotendinous range of motion. Therefore, these genes can also be considered as candidate genes for association with fundamental tendon properties. Two studies have already reported an association between a polymorphism in one of these genes and measures of musculotendinous range of motion in humans [41, 42]. Table 2 summarises the genetic association studies that have identified a polymorphic association with tendon pathologies/musculotendinous range of motion in humans. In addition to the four genes associated with tendinopathies (*COL5A1, TNC, MMP3, GDF5*), *COL5A2* will also be discussed in this review because, like *COL5A1*, it codes for a protein which is a fundamental component of the Col V molecule (quaternary protein). Type I collagen (*Col I*), and its two coding genes *COL1A1* and *COL1A2*, will also be discussed because it forms the major structural component of tendon, even though genetic variation has not yet been associated with tendinopathies or tendon properties. Associations have, however, been observed between genetic variation in *COL1A1* and risk of ligament injury [43, 44].

This article will review genes and proteins that are known to be related to tendon structure and/or the dynamic nature of tendon homeostasis, and the review will describe how variations within these genes may affect human tendon mechanical properties such as stiffness and elastic modulus. There is substantial evidence that tendon mechanical properties such as stiffness influence the capacity of the muscle-tendon unit to produce force during exercise of various kinds [4-8]. For example, it has been shown that a relatively stiffer muscle-tendon unit is significantly related to maximal concentric and isometric muscle performance [4]. On the other hand, during cyclical contractions low tendon stiffness (high tendon compliance) has been shown to improve muscle power and efficiency [7]. Consequently, genotype associations with tendon mechanical properties are highly likely to also influence exercise performance in a more general sense.

### 2. GENES AND PROTEINS OF INTEREST

#### 2.1. Collagen Type V

**2.1.1. Structure of Protein and Genes**

Col V is a widely distributed quantitatively minor fibrillar collagen forming between 1-3% of total collagen content of tendon ECM [45], although evidence suggests that in functional terms it is a major collagen of developing connective tissues [46]. Col V can assemble into a diverse number of molecular forms but all contain a pro α(V) chain. This pro α1(V) chain is encoded for by the *COL5A1* gene (9q34.3) and comprises 66 exons distributed over 203.07 kilobases (kb) of genomic DNA. There is a pro α2(V) chain which is encoded for by the *COL5A2* gene (2q32.2), comprising of 54 exons and 147.98kb [47]. Together these chains form the heterotrimer protein structure of Col V ([α1(V)]2:α2(V)), which is ubiquitous in human tendons.

| Table 1. Tendon Proteins, Genes of Focus, and Abbreviations Addressed in this Review |
|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Protein                        | Abbreviation of Protein Used in this Review | Genes of Focus in this Review | Abbreviation of Gene of Focus |
| Type V collagen                | Col V                                  | collagen, type V, alpha 1      | *COL5A1*                        |
|                                |                                       | collagen, type V, alpha 2      | *COL5A2*                        |
| Tenascin C                     | Ten C                                  | tenascin C                     | *TNC*                           |
| Matrix metalloproteinase-3     | MMP-3                                  | matrix metalloproteinase 3     | *MMP3*                          |
| Growth/differentiation factor 5 | GDF-5                                  | growth differentiation factor 5 | *GDF5*                          |
| Type I collagen                | Col I                                  | collagen, type I, alpha 1      | *COL1A1*                        |
|                                |                                       | collagen, type I, alpha 2      | *COL1A2*                        |

Note: the two collagen proteins comprise numerous protein chains and therefore are dependent on more than one gene.
2.1.2. Function

Col V plays a functionally important role in tendon via its relationship with Col I fibrils, in that it is thought to copolymerise with Col I fibrils to form heterotypic fibres, and thereby organises and regulates the diameter of these fibres [48-51] as well as forming intermolecular cross-links with Col I fibrils [52]. Interaction of Col V with Col I in in vitro self-assembly assays has shown a decrease in the diameter of fibrils with increasing amounts of Col V [53], possibly due to an increase in nucleation sites in the thin filaments of Col V for a given quantity of Col I [49]. These sites serve as steric hindrances for the addition of Col I molecules, through amino terminal domains which project out, thereby regulating lateral growth and hence diameter [51, 52].

Both fibril diameter [29, 54, 55] and extent of cross-linking [56-59] are positively correlated with mechanical
properties of tendon such as stiffness and Young’s Modulus, in animal models. Additionally, this correlation has been reported in vivo in relation to fibril diameter in humans [55]. There appears to be an optimum level of permanent cross-linking, above which there is a decrease in mechanical strength [53, 60, 61]. However, these studies investigated age-related declines in mechanical properties and the same relationship with cross-linking may not exist in a cross-section of a younger population. Reduced Col V content has been reported to compromise the diameter of the collagen fibril in vitro cultures [62, 63] and consequently may reduce the material properties of tendon such as maximum stress and linear modulus [53].

2.1.3. Evidence of Polymorphic Associations with Tendon Pathologies

The first study to report an association between variation in the COL5A1 gene and tendon pathology [39] identified the COL5A1 gene as an ideal candidate genetic marker of Achilles tendinopathies because it is on the same locus of genomic DNA as the ABO gene, which has been reported to be associated with tendon injuries in Hungarian and Finnish patients [64, 65]. Two restriction fragment length polymorphisms (RFLPs) were identified within the 3’ untranslated region (UTR) of the COL5A1 gene (BstUI and DpnII) that had no known role in the expression or function of Col V. An association was found between the BstUI RFLP (rs12722) and Achilles tendon pathology (ATP), and more specifically chronic tendinopathy without rupture (the C allele was protective against ATP). As the authors rightly stated, this association does not show conclusively that Col V is involved in the development of these pathologies, and it is of course likely that there are numerous genetic variants that contribute to the overall heritability of such conditions [66]. It is also possible that non-genetic factors influenced the results of the study described in this paragraph [39], as body mass and physical activity were not controlled participant selection criteria.

A subsequent study investigated the same variant of the COL5A1 gene but in two separate white populations in South Africa and Australia [38] and the results generally concurred with the initial association study [39], in that rs12722 (BstUI RFLP) within the 3’ UTR was associated with Achilles tendinopathies and associated with individuals who possess a ‘T’ allele at this locus. Thus, individuals who were homozygous for the ‘C’ allele were apparently less likely to develop the condition. The authors intelligently attempted to investigate a combination of additional markers or neighbouring alleles in the same sequence region of the 3’ UTR of the COL5A1 gene, known as an ‘inferred’ haplotype, in order to provide more information as to the predisposing causative factor. The haplotype consisting of markers rs12722 and rs3196378 (alleles ‘T’ and ‘C’ respectively) was
significantly overrepresented in the South African tendinopathy group but not in the Australian group. The DNA sequence that contains the ‘C’ allele at the rs3196378 marker forms part of a miRNA recognition sequence. miRNA are key regulators of gene expression at a posttranscriptional level by inhibiting translation or inducing mRNA cleavage [67, 68]. Consequently, protein expression may be modified and in this instance one could speculate that COL5A1 expression may be altered, leading to suboptimal levels of Col V protein and ultimately a compromised collagen fibre and healing process. Regarding COL5A2, no polymorphisms as of yet have been associated with tendon pathologies or function.

2.1.4. Possible Influences on Tendon Properties

Two recent studies considered a measurable in vivo phenotype that may link the COL5A1 gene variant to tendon injuries, by investigating flexibility - a possible immediate phenotype [41, 42]. Flexibility is an established determining factor for patellar tendinopathies in active populations [69, 70]. Genetics has been reported to contribute substantially to the variability of certain flexibility phenotypes [71, 72], and although that is not the case in other studies [73, 74], associations between gene variants and flexibility are therefore plausible. COL5A1 was hypothesised to be associated with flexibility following reports that mutations in the COL5A1 gene have been implicated in Ehlers Danlos syndrome, a condition characterised by joint hypermobility [75]. These disease-associated rare mutations may produce non-functional COL5A1 and COL5A2 alleles leading to haploinsufficiency of COL5A1 and COL5A2 mRNA, predictably resulting in the synthesis of around half the amount of normal Col V protein [75-78]. Phenotypically, this may result in abnormally large collagen fibrils [79] and impaired mechanical properties of tendon. In the first study, an association was reported between the common polymorphism rs12722 and flexibility [41]. Individuals heterozygous (CT) for this genotype were less flexible than homozygous individuals of either allele, however the study sample contained significant heterogeneity in terms of tendon injury history - i.e. participants with tendinopathies, history of rupture, and no history of tendon injury were combined in the genotyping results. Furthermore, measures used to quantify flexibility, such as an instrumental standing leg raise [80] and a trunk flexion sit and reach test [81], are rather crude measures of the function of the muscle-tendon unit as a whole, and certainly do not provide precise data on the mechanical properties of the tendon per se. Regarding the instrumental standing leg raise test, subjectivity [82] and specifically the perception of pain onset, as well as abnormal defence reactions [83], could be particularly relevant to the participants with history of Achilles problems. However, the main critical comment on the instrumental standard leg raise test is that the measured phenotype is a composite of various factors including muscle and tendon stiffness and does not take into account the dimensions of the structures so the mechanical properties cannot be determined. Additionally, any kind of sit-and-reach test is a composite measure of various factors contributing to ‘flexibility’, including muscle-tendon unit stiffness [84], limb lengths and proportions [85]. Thus, we conclude that while their approach was a useful step in the study of the genetics of flexibility and range of motion, the mechanics of the tendon per se were clearly not determined by Collins et al., [41] so potential associations between tendon properties and genes coding for proteins expressed in tendon could not be investigated directly in that study. The previous critique is accentuated by a report showing the mechanical properties of the series elastic component (tendon-aponeurosis) are independent of the parallel elastic component (passive muscle stiffness) in vivo [86].

In the second study, Brown et al., [42] investigated the COL5A1 rs12722 polymorphism and sit-and-reach performance in a healthy and physically active cohort (325 Caucasian subjects). Individuals homozygous for the ‘C’ allele had greater flexibility, but this was only observed in the older (≥35 years) subjects where sex and COL5A1 genotype accounted for approximately 23% of the variance. As per the previous study, some factors which may affect the material and mechanical properties of the tendon were not considered. For example, circulating oestrogen was not assessed in the female subjects - chronic oestrogen levels can influence tendon stiffness [87]. Also, a lack of detailed information regarding the habitual physical activity levels of the older subjects was another limitation noted by Brown et al., [42] themselves, because tendon stiffness can increase [15, 88] with higher physical activity and decrease [89, 90] with lower activity.

In conclusion for Col V, it is indeed possible that variations in the genes that encode for the molecular components of Col V may influence a tendon’s material and mechanical properties, although such gene variants have not yet been shown to influence tendon properties per se. Since Col V expression levels appear critical in determining a tendon’s fibre structure through diameter and cross-linking [52, 63], several testable hypotheses regarding genetic variants and mechanical properties of tendon such as stiffness, maximal strain and elastic modulus are likely to be tested in the coming years.

2.2. Tenascin-C

2.2.1. Structure of Protein and Gene

Ten C is an ECM glycoprotein consisting of 6 monomers, expressed in tissues bearing high tensile stress, such as tendons and myotendinous and osteotendinous junctions [91] and as with the collagen fibre, is an important structural component of tendon [92, 93]. This protein is known to be sensitive to mechanical loading in vitro [94]. It is encoded by the TNC gene (9q33.1), which comprises 28 exons spanning 97.63 kb of genomic DNA [47].

2.2.2. Function

Ten C is a structural component of tendons, yet is an elastic protein [95], and as it is expressed in mechanically loaded tendons it may contribute to increased elasticity of the ECM [92, 96, 97] via increased gene expression in response to stretch [98]. In addition to its structural roles, Ten C performs various regulatory roles within the ECM. Due to its modular structure, the protein is able to interact with various other proteins involved in ECM homeostasis, as well as playing an important role in regulating cell-matrix interactions [99]. It is also believed that Ten C plays an invaluable role in regulating proper alignment and
organisation of the collagen fibres in vitro [95]. Therefore, when Ten C is expressed it may contribute to an increased crimp angle - a region-specific morphological feature of collagen fibres associated with the mechanical properties of tendon.

Ten C may also be a strong candidate for involvement in the aetiology of musculoskeletal injuries, as expression of the TNC gene has shown to be altered with human tendinopathies, determined by immunoblotting [100, 101]. This may be particularly relevant when engaging in intensive activity following inactivity. Inactivity decreases tendon stiffness in vivo, particularly at the muscular end of human tendon [89]. Yet Ten C expression relies on mechanical loading and as it is an ‘elastic’ protein, a relative decrease in stiffness due to inactivity or insufficient mechanical loading makes intuitive sense if Ten C is more highly expressed. Indeed, the myotendinous interface has been reported to be mechanically the most vulnerable site for injury [102, 103]. Thus, during reloading after inactivity the overall extensibility (strain to failure) would decrease increasing the risk of tendon rupture, due to experiencing greater strains for a given load [104].

2.2.3. Evidence of Polymorphic Associations with Tendon Pathologies

One study investigated a guanine-thymine (GT) dinucleotide repeat polymorphism for potential association with the risk of incurring both chronic Achilles tendinopathies and Achilles tendon rupture [37]. This polymorphism is a tandem repeat, consisting of a 2 base pair sequence repeated a varying number of times within a non-coding region (intron 17). Variants containing 12 and 14 GT repeats were overrepresented in subjects with tendinopathies, while variants containing 13 and 17 repeats were underrepresented. The control group had been active in high impact sports for a considerable time (11.5 years) and were currently engaging in ~5 hours per week, and so their apparent resistance to tendinopathies was unlikely to be due to a significantly lower exposure to high impact loading. Thus, a genetic influence may indeed exist, although replication of these data would be a valuable development. Furthermore, whether the TNC polymorphism is involved in causative mechanisms is still debatable.

2.2.4. Possible Influences on Tendon Properties

Even though the GT repeat polymorphism in intron 17 is not part of the coding sequence, intronic variations may influence the binding of proteins involved in gene transcription, thus affecting gene expression. As Ten C expression has been reported to be up-regulated in certain pathological conditions [105-107], it could be postulated that the 12 and 14 GT repeats within intron 17 of the TNC gene may overexpress Ten C, increasing the elastic properties of the myotendinous unit, as well as reducing the ultimate tendon strain to failure for a given load [97]. Thus, TNC is a candidate gene with regards to determining the degree of passive stiffness/compliance of tendon.

2.3. Matrix Metalloproteinase 3

2.3.1. Structure of Protein and Gene

MMP-3 (otherwise known as stromelysin-1) is part of a group of 5 domain structures of zinc-dependent enzymes known as Matrix Metalloproteinases (MMPs), characterised according to the type of zinc binding. Structurally, the MMP-3 protein constitutes a multi domain structure made up of a propeptide, a catalytic N-terminal domain and a haemopexin-like C-terminal, all of which combine to form functional MMP-3 which has the capacity to interact with its substrates [108]. The MMP-3 protein is encoded for by the MMP3 gene (11q22.2) which is 10 exons in length and covers 7.79kb [47].

2.3.2. Function

MMP-3 plays a crucial role in the normal development, repair and remodelling of connective tissues, and ultimately plays regulatory roles in maintaining ECM homeostasis, through proteolytic activity. This is achieved by catalysing the degradation of both ECM and non-ECM proteins [109-111]. MMP-3 hydrolyses multiple substrates including different types of intact fibrillar collagens, proteoglycans and a wide range of ECM components [112]. The ECM is in a state of dynamic equilibrium between synthesis and degradation [113], and its gene expression has shown to be increased by mechanical loading in vitro [114, 115]. Recently, ECM regulation has been shown to be determined by a combination of duration and magnitude of the mechanical stimulus in vitro [13], and in an in vivo rodent model [116], which potentially represents the impact of differing forms of voluntary exercise on tissue remodelling processes in humans.

The expression level of MMP-3, which may be regulated at the transcriptional, translational, or posttranslational levels by interaction with inhibitors [117], appears to differ between normal and highly stressed tendons, as well as tendons displaying pathological characteristics, determined by histological analyses. In highly stressed tendons, expression levels are elevated compared to normal tendons in animal models [13, 118, 119], which is thought to represent a repair or maintenance function that may be associated with an underlying degenerative process [117]. In addition, ‘stress-shielding’ or load deprived tendon has shown to increase the expression of MMP-3 mRNA in relation to normal cadaveric tendon tissue samples, determined in vitro [118, 120, 121]. Collectively, these studies point toward a ‘U’ shape relationship between load and MMP-3 expression. The extremes of this relationship may cause a loss of mechanical function, which may be related to the subtle degradation of ECM components, notably those involved in cross-linking and/or stabilisation of the tendon structure [121], such as minor collagens including Col V as well as proteoglycans.

In contrast, a lower level of MMP-3 expression compared to normal tissue samples has been reported in human tendon displaying pathological characteristics [100, 117, 122-124]. These observations may represent a failure of the normal matrix remodelling process [125]. It should be noted that even in normal human tendon, there is a significant difference between sexes where males have twice the amount of resting mRNA expression levels of MMP-3 compared to females [126], which may indicate an impaired ECM maintenance and a weakening of the material properties in females, leading to increased injury susceptibility [127].
Even though the evidence is generally consistent across studies, with regards to the MMP-3 gene expression in different mechanical environments and in pathologies, there are some complexities. For instance, increased MMP-3 mRNA expression does not mean that a given amount of MMP-3 protein will be produced due to post-transcriptional and post-translational regulation [112, 125].

2.3.3. Evidence of Polymorphic Associations with Tendon Pathologies

Gene variants have been investigated in the MMP3 gene which have the potential to substantially alter its expression [128], particularly the 5A/6A polymorphism within the promoter region of human MMP3. This polymorphism has been associated with a number of pathological states [129-131]. The association between gene variants in MMP3 and tendon pathology was first postulated when immunohistochemically detectable MMP-3 protein was lower in a ‘normal’ region of Achilles tendon tissue in patients with a degenerate core region nearby, compared to normal control tissue [100]. This suggests these patients with tendinosis were predisposed to developing the condition due to inherently reduced MMP-3 protein levels.

One study to date has reported an association between variation in the MMP3 gene and Achilles tendinopathy [40]. Three SNP’s spanning most of the gene were identified as being potentially informative, as they are part of all four major haplotypes within the MMP3 gene (one exon SNP - rs679620, two intron SNP’s - rs591058, rs650108). All three MMP3 variants were found to be associated with Achilles tendinopathy individually, and as inferred haplotypes - particularly between the rs679620 and rs591058 gene variants. These two variants were found to be in almost perfect linkage disequilibrium. In contrast, the ‘ATG’ inferred haplotype containing all three SNP’s were significantly underrepresented in the tendinopathy group compared to the control group, suggesting this combination has a protective effect against the development of Achilles tendinopathy.

Raleigh et al., [40] were the first to demonstrate an interaction between variants on two different genes, vis-à-vis the development of Achilles tendinopathy (all three SNP’s of the MMP3 gene and the marker rs12722 of the 3’ UTR region of COL5A1 gene). The rs679620 marker of the MMP3 gene and the rs12722 marker of the 3’ UTR region of COL5A1 gene represent the best pair of genotypes for estimating the risk for Achilles tendinopathy, with the ‘G+T’ allele combination associated with tendinopathies. However, the authors do not address how the MMP3 variants alone, or as haplotypes and inferred haplotypes between different genes, cause an increased/decreased risk of tendinopathies. Nevertheless, they do suggest that the rs679620 variant of MMP3, which is a non-synonymous polymorphism, may influence the downstream function of the mature MMP-3 enzyme and its activation [129] due to the subtle change in the amino acid coding and its interaction with other amino acids (‘G’ allele=glutamate, ‘A’ allele=lysine). The ‘G’ allele may encourage elevated levels of MMP-3 expression via increased MMP-3 activation, as a result of altered interaction with other amino acids in the propeptide region.

2.3.4. Possible Influences on Tendon Properties

As the COL5A1 BstUI RFLP was shown to be associated with human flexibility, the MMP3 rs679620 variant was investigated for this same association, though no association was evident [132]. The precise rationale for investigating a link between this gene variant and flexibility is unclear, although as a link was previously identified between the MMP3 gene variant and Achilles tendon injuries [40] and as flexibility has been reported to be a possible risk factor for these injuries [133], the investigation seems justified. It must be noted that the flexibility phenotype assessed was a measure of musculoskeletal passive flexibility, which encompasses tendons, ligaments, joint capsules, aponeuroses and fascia sheaths, as well as the muscle and not necessarily just the tendon. Thus, as previously mentioned in section 2.1.4 there are limitations to the techniques used for measuring flexibility in these studies.

As the mechanical properties of tendons are primarily a function of the ECM, and because a majority of ECM components are substrates for the proteolytic activities of MMP-3 [109], it may be that MMP-3 expression would contribute to the material integrity, and thus tendon mechanical properties. It may be that elevated expression levels of the MMP3 gene indicate a degenerative environment, putting the ECM in a state of imbalance with a greater rate of degradation compared to synthesis. The tendon homeostatic abilities would thus be compromised and substrates involved in cross-linking and stabilisation of the collagen fibril (Col V) may be degraded, ultimately weakening the material properties and resulting in a reduction in matrix stiffness [56-57].

2.4. Growth Differentiation Factor-5

2.4.1. Structure of Protein and Gene

GDF-5 is a member of the transforming growth factor (TGF) super-family, encoded for by the GDF5 gene (20q11.22) of which its entire coding region comprises 4 exons and is approximately 21.42kb in length [47]. Structurally, it is a ‘dimer’ consisting of two monomers interlinked by disulfide bonds. Mature forms of the protein are approximately 110-140 amino acids in length and seven cysteine amino acid residues are involved in creating its rigid structure [134].

2.4.2. Function

GDF-5 is involved in maintenance, growth and repair of bones, cartilage and musculoskeletal soft tissues including tendon [135-137]. When GDF-5 was first investigated for its possible role in tendon biology, it was found to possess a unique ability to induce a tendon-like tissue rather than cartilage and bone, when implanted intra-muscularly in rats [138]. Further investigations found a significant role of GDF-5 in tendon within rodent models with induced Achilles tendon injuries. Firstly, GDF-5 was found to enhance tendon healing and tensile strength of the tendon when implanted on collagen sponges, in a dose-dependent manner [139]. Further studies examined the ultrastructural, compositional and mechanical characteristics of the Achilles tendon in rodents deficient in GDF-5, and found the maximum load to failure decreased possibly due to
significantly less collagen, an increase in irregularly shaped Col I fibrils and compromised material behaviour (decrease in strength and stiffness) [140-142]. Therefore unsurprisingly, GDF-5 has been shown to increase mechanical strength in these rodent models [143-146]. These studies are further supported at a cellular and gene level by studies showing an improved collagen organisation with GDF-5 treatment [147, 148], as well as an increased expression of genes and synthesis of the components of tendon ECM, in particular, Col I, Ten C and MMP-3 [149].

2.4.3. Evidence of Polymorphic Associations with Tendon Pathologies

The human GDF5 gene contains mutations known to cause a number of rare inherited disorders, including acromesomelic chondrodysplasia of the Hunter-Thompson and Grebe types as well as Du Pan Syndrome, all of which are characterised by musculoskeletal abnormalities, including shortened limb bones, brachydactyly and severe joint dislocations [150-152]. The hypothesised involvement of GDF5 in tendon pathologies derives from this evidence, as it was postulated that the observed joint dislocations may be attributed to abnormalities in tendons [140]. As well as genetic mutations within the GDF5 gene, a functional promoter SNP (rs143383; T/C) of the 5’ UTR of the GDF5 gene has been associated with multifactorial disorders, such as osteoarthritis at different joint locations across different ethnic groups [153-156], congenital dislocation of the hip [157], as well as total body height, hip axis length and fracture risk [158, 159], and lumbar disc degeneration [160]. In articular cartilage of individuals with osteoarthritis, there was a 12% lower expression of GDF-5 associated with the ‘T’ allele at this SNP marker compared to the ‘C’ allele [161]. So, a reduction in the expression of the GDF5 gene associated with the ‘T’ allele may contribute to tendon pathologies, and this has been investigated by one study.

In a case-control study, an association was reported between the GDF5 SNP rs143383 referred to above and the risk of ATP [36]. Individuals of ‘TT’ genotype were found to have approximately twice the risk of developing ATP within an Australian population independently and when combined with a South African population, which probably means it is less likely to be a false positive observation. No significant association between genotype and higher risk was shown in the South African cohort alone, although the observed odds ratio was still similar (~1.7). The relatively small sample size of the Australian tendinopathy group (n=59) as well as the different physical characteristics (body mass and BMI) between the ATP and control groups of both populations, are perhaps limitations of the study. However, the odds ratios and confidence intervals observed suggest a robust association and these findings complement those studies demonstrating the impact on gene expression of the GDF5 variant in question [161-163].

2.4.4. Possible Influences on Tendon Properties

It may be hypothesised that the material properties of the tendon are compromised in the presence of the rs143383 ‘T’ allele variant, i.e. a reduction in tensile strength and stiffness. GDF-5 may be involved in collagen cross-linking by promoting the proteolytic activation of lysyl oxidase [164] as well as mediating the collagen structure and organisation, by increasing the thickness of collagen fibrils [140]. Collagen cross-linking and diameter per se, have been shown to increase tendon matrix stiffness in animal models [54, 56, 57] and humans [55], so the ‘T’ allele variant of GDF5 may hinder these processes and thus reduce tendon stiffness. Therefore, if gene variants within the GDF5 gene influence tendon material and mechanical properties, it is likely to be via mediating the growth of other structures, in particular Col I fibrils.

2.5. Collagen Type 1

2.5.1. Structure of Protein and Genes

Collagen is the main protein constituent of tendon tissue and is reported to make up 65-75% of a tendon’s dry mass [165] in cadaver tissue and more recently ~90% of in vivo patella tendon biopsies in humans [166]. Collagen comprises fibrillar collagen molecules containing more than 95% Col 1 [27]. This collagen protein is encoded for by the COL1A1 gene (17q21.33), which constitutes 52 exons and is 18.34kb in length, and to a lesser extent by the COL1A2 gene (7q21.3), which constitutes 52 exons also, and is 36.67kb in length [47]. The COL1A1 gene encodes for the alpha (α) 1 chain, while the COL1A2 gene encodes for the α 2 chain. Two α 1 chains and one α 2 chain combine to form a heterotrimer protein structure.

2.5.2. Function

Col I is a major protein constituent contributing significantly to the structural integrity of soft tissues such as cruciate ligaments, joint capsules and tendons, via the formation of strong parallel bundles of fibres. Col I fibrils and fibres are well recognised to be involved in tensile strength and the stiffness of tendon matrix, based on its intra- and inter- molecular cross links, orientation, density, diameter and length, all of which have been shown to affect the mechanical properties of the tendon as a whole in animal models [54-59, 167, 168].

2.5.3. Possible Polymorphic Associations with Tendon Pathologies

Mutations as well as single nucleotide polymorphisms in the COL1A1 gene, particularly a SNP affecting the Sp1 binding site in the first intron of the COL1A1 gene (+1245; G/T; rs1800012), have been associated with lower bone mineral density and osteoarthritis [169-176] as well as being implicated in the disease osteogenesis imperfecta (OI) which is characterised by fragile collagen structures [177-179]. Additionally, a point mutation in the COL1A2 gene (nucleotide position 1121) that substitutes serine or cysteine for glycine residues (C-to-T transition and G-to-T transversion, respectively), also leads to the OI phenotype [180]. Consequently, sequence variants such as these, and others that may have a less clinically evident but still important influence, might be associated with soft tissue injuries. Indeed, associations have been reported between SNP’s in the COL1A1 gene at the intronic Sp1 transcription factor binding site and the risk of cruciate ligament ruptures and shoulder dislocations [43, 44], as well as upper limb muscle strength in elderly men [175]. These associations may be mediated through reduced Col 1 content or a weaker form of Col I, but as of yet no genetic association has been made with tendon pathologies or tendon properties.
2.5.4. Possible Influences on Tendon Properties

No association has yet been reported between variation in the \textit{COL1A1} gene and tendon pathologies or properties. It is known that a SNP within the intronic Sp1 binding site (rs1800012) increases transcriptional activity of the \textit{COL1A1} gene, resulting in abnormal ratios of the \textit{α1(1)} protein relative to \textit{α2(1)}, which possibly gives rise to weaker heterotrimers being formed (three \textit{α1(1)} chains) instead of the conventional heterotrimers (two \textit{α1(1)} and one \textit{α2(1)} chains) [178]. It has also been reported that two polymorphisms in the proximal promoter of \textit{COL1A1} are in linkage disequilibrium with the Sp1 polymorphism [181], and in fact form an extended haplotype with the Sp1 polymorphism to regulate \textit{COL1A1} transcription. This is achieved by affecting the binding affinity of important regulating factors, such as Sp1, with the ‘T’ allele at the Sp1 binding site found to have a higher DNA binding affinity than the ‘G’ allele [171]. Consequently, individuals who carry a ‘T’ allele instead of a ‘G’ at this SNP, highly express \textit{COL1A1} and thus possess a greater proportion of the weaker \textit{α1} heterotrimers, may be more likely to have a compromised tendon internal structure. It has also been suggested that overproduction of Col I \textit{α1} chains in tendon might result in a higher tensile strength [44], although that statement contradicts the mechanism just outlined and is not expanded upon by the authors.

It is unlikely that tendon properties are affected solely by the Sp1 polymorphism in intron 1 of the \textit{COL1A1} gene. It is more likely that the extended haplotype influences the transcription of \textit{COL1A1} [171] and the material quality of the Col I fibril, with individuals carrying a ‘T’ allele at the Sp1 polymorphism ultimately producing higher gene activity, which might contribute to a more adversely affected Col I fibril. In addition, other genes such as those already reviewed need to be considered at the same time.

CONCLUSION

Referring to the studies examining gene variants and their associations with tendinopathies, it appears there is no single causative gene variant that predisposes individuals to tendinopathies. This suggests tendinopathies are likely to be polygenic. This concurs with the expectation that a multitude of genes, their associated proteins, and their heterogeneous interactions are required to maintain normal tendon structure and homeostasis through development, regeneration and normal function [62, 168]. Therefore, it is likely, even after controlling for other parameters such as gender, age and habitual physical activity, that the intrinsic material (structural and regulatory) and mechanical properties of a tendon, quite apart from tendinopathies, are similarly influenced by polygenics.

Caution must be taken when interpreting the findings of these genetic association studies on tendinopathies. Investigations into the genetic factors involved in their aetiology, thus far, are very much in their infancy [182]. To date, all studies examining the genetic factors involved in tendinopathies investigate the pathology of the Achilles tendon, so whether these findings are applicable to other types of tendon is unknown. Also, the studies fail to control for all pertinent environmental factors, which may influence inter-individual variability in risk of tendinopathy. Non-genetic factors such as age, body mass and physical activity of participants, all of which may affect the risk of incurring tendon injuries (independent of genetic predisposition), are usually not controlled for when recruiting individuals for these studies. This is mainly because the studies are retrospective in nature, rather than prospective, in that the pathological condition was evident before genetic variables were considered. Lastly, these genotype-phenotype associations have only been reported in people of one form of geographic ancestry (white Caucasian), and it remains to be investigated whether these associations will be observed in peoples of other geographic ancestry. Additionally, to improve the strength of these relationships intra-ancestral, it would be beneficial to conduct twin/family studies to provide estimated inter-individual variability that is inherited, which has not yet been done [182].

Having further highlighted the role of ECM proteins and their possible link to tendinopathies, it appears they do not act as one single entity but subtly interact to form interlinked structures (Fig. 1) and govern dynamic processes within the ECM. Col V fibrils combine with Col I fibrils to regulate the diameter of the fibres [49] with Ten C playing an invaluable role in regulating the proper alignment and organisation of the collagen fibres [95]. MMP-3 may degrade minor collagens such as Col V, which may alter the cross-linking and stabilisation of the tendon structure [112]. And lastly, an improved collagen organisation and cross-linking density was observed with the addition of GDF-5 [147, 164]. The common theme in these associations is the integrity of the collagen fibre in conjunction with organisation, diameter and cross-linking, all of which have been linked to tendon properties such as tendon matrix stiffness [54, 183].

It is therefore reasonable to suggest that variants in genomic DNA sequence within these and other relevant proteins are likely to contribute to observed phenotypic variations in the tendon, most notably the mechanical properties, which may have implications for physical performance capabilities and the risk of incurring musculoskeletal injuries. However, no study has yet attempted to investigate genetic influences upon tendon properties \textit{per se} in an asymptomatic population.

FUTURE

When investigating a gene variant’s influences on tendon properties, it is important to negate factors other than genetics that are likely to contribute to tendon phenotypes. To establish a valid and reliable association between a gene variant and tendon mechanical properties, experimental error must be minimised and appropriate phenotype measurements utilised, which has not been adequately achieved in the previous genetic association studies investigating flexibility [41, 42]. It would be more appropriate to measure the overall stiffness of the muscle-tendon unit to assess flexibility using, for example, passive isokinetic dorsiflexion adopted by Morse et al. [184] to assess the human gastrocnemius muscle-tendon unit. This comprehensive \textit{in vivo} assessment utilises techniques such as dynamometry, electromyography, electromyography (EMG) and ultrasonography. The mechanical properties of the tendon itself can be assessed \textit{in vivo}, which would be in line with the objectives of associating genetics with tendon properties. A thorough and highly reliable assessment is detailed by Pearson and
Onambele [185] with respect to the tendon compliance of the patella tendon, in thatodynamometry, EMG and ultrasonography were utilised as well as the force-displacement relationship to calculate tendon mechanical stiffness. Therefore, an accurate, reproducible and non-invasive assessment of tendon properties in vivo is required to maximise the ability to detect a genetic contribution to the interindividual variability in mechanical properties of human tendon.

From a genetic perspective, a powerful approach to find significant DNA polymorphisms associated with tendon phenotypes could be to perform genome-wide association studies (GWAS). Geneticists have developed genotyping arrays (often called SNP chips) that can now assay up to 2 million variants simultaneously which, due to linkage disequilibrium, capture a substantial proportion of total genomic variability [186]. GWAS involves testing a comprehensive catalogue of common genetic variants, and can be applied to a case-control study (to find those variants associated with a medical condition or other extreme phenotype, such as patients with tendinopathies or athletes successful in sports requiring high flexibility) or to a genotype-phenotype association study (to find those variants associated with a phenotype measured on a continuous scale). By testing all common variants, one could pinpoint key genes and shed light on underlying mechanisms. Three key results have emerged from GWAS: (1) most traits can be influenced by a large number of loci; (2) the vast majority of the common variants at these loci have a moderate effect, increasing risk by 10–50% (similar to effects of many environmental risk factors); and (3) the loci include most of the genes found by linkage analysis, but reveal many more genes not previously implicated [187]. Over the next decade, it would be desirable to conduct genetic studies of thousands of patients with tendinopathies (in case-control studies) and of healthy subjects with measures of flexibility (genotype-phenotype studies), with appropriate combinations of GWAS and sequencing. In turn, intensive functional studies will be required to characterize the genes and pathways, and to construct animal models that mimic human tendon physiology.

ACKNOWLEDGEMENTS

This research has been supported in part by an International Joint Project grant from the Royal Society, London, UK.

CONFLICT OF INTEREST

Declared none.

REFERENCES


Foster et al.
Tendon Properties in Humans


Pearson SJ, Onambele GN. Influence of time of day on tendon compliance and estimations of voluntary activation levels. Muscle Nerve 2006; 33: 792-800.


Received: November 14, 2011 Revised: January 23, 2012 Accepted: January 27, 2012

© Foster et al.; Licensee Bentham Open.
This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.