COLLAGEN GENE SEQUENCE VARIANTS IN EXERCISE-RELATED TRAITS

Kevin O'Connell,[1, D] Colleen J. Saunders,[1, D] Malcolm Collins[2, 1, D]

[1] MRC/UCT Research Unit for Exercise Science and Sports Medicine of the Department of Human Biology, Faculty of Health Sciences, University of Cape Town, South Africa
[2] South African Medical Research Council in Cape Town, South Africa
[D] Manuscript Preparation

Address for correspondence:
Prof. Malcolm Collins
MRC/UCT Research Unit for Exercise Science and Sports Medicine,
University of Cape Town, PO Box 115, Newlands, 7725, South Africa
Phone: +27 21 650 4574
Fax: +27 21 686 7530
E-mail: malcolm.collins@uct.ac.za

Abstract. Collagens are major structural proteins of tendons, ligaments and other components of musculoskeletal tissues. Rare mutations in many of the genes, which encode for the collagen α-chains, result in serious musculoskeletal disorders, highlighting the importance of this protein family in the normal structure and function of musculoskeletal tissues. Since these rare mutations cause severe disorders, it has been proposed that a lack of biological redundancy exists within the collagen fibril, and that collagen-encoding genes are therefore ideal candidates for association with less severe exercise-related traits. This review identifies a number of collagen gene variants which are associated with various exercise-related traits. Based on the evidence outlined in this review, we propose that a general genetic continuum exists for collagen genes and their associated traits. At one end of this general continuum model, a single mutation within one or more collagen genes will result in severe Mendelian disorders. At the other end of the continuum, functional variants within these collagen genes collectively contribute to the aetiology of anomalous multifactorial connective tissue traits, which arise as a result of the interaction of genetic and non-genetic factors which modulate physiological responses to environmental stimuli.

Key words: Achilles tendinopathy, ACL, performance, flexibility, ligament

An Overview of Collagens and the Structure of the Collagen Fibril

Collagens are a family of twenty-eight structurally and functionally diverse proteins which consist of three polypeptide α-chains wound in a characteristic uninterrupted or interrupted triple helical structure (Kadler et al. 2007; Liu et al. 1995). The majority of these proteins are important for tissue assembly or maintenance (Kadler et al. 2007). Furthermore, many are the major macromolecular building blocks of the collagen fibril, which is the structural component of tendons, ligaments, bone, cartilage and connective tissue structures in muscle and other tissues. Non-cartilaginous connective tissues generally contain a different set of collagen proteins when compared

to cartilage, however some collagens do occur in both (Table 1, Figure 1) (Banos et al. 2008; Kadler et al. 2007). Since the collagen fibril within the musculoskeletal system is an important structure in exercise performance and injury, this review will focus predominately on the structural collagens of the non-cartilage and cartilage fibril, as well as those collagens that are directly associated with these fibrils.

The collagen family can be divided into two groups, based on their structure and function: (i) the fibrillar, or fibril-forming, collagens which form the fibrillar scaffolding for the extra-cellular matrix and, (ii) the non-fibrillar collagens which include, amongst others, the fibril associated collagens with interrupted triple helices (FACITs), network forming collagens, short chain collagens and beaded filament collagens (Table 1) (Kadler et al. 2007; Riley 2005).

Table 1. The classification of some of the structural collagens within the non-cartilage and non-cartilage fibril, together with those that associate with the fibril or form networks around the fibril forming cells

Tissue Distribution	Fibrillar Collagens		Non-fibrillar Collagens	
	Major	Minor	FACITs[1]	Others
Tendons Ligaments Bone Muscle Other non-cartilaginous tissues Fibrocartilage	Type I Collagen Type III Collagen	Type V Collagen	Type XII Collagen Type XIV Collagen	Type IV Collagen[2] Type VI Collagen[3] Type VIII Collagen[4]
Cartilage Fibrocartilage	Type II Collagen	Type XI Collagen[5] Type XXVII Collagen[6]	Type IX Collagen	Type IV Collagen[2] Type X Collagen[4]

[1] Fibril associated collagens with interrupted triple helices.
[2] Network forming collagens.
[3] Beaded filament collagens.
[4] Short chain collagens.
[5] Also expressed in the developing tendon.
[6] Not a classical fibrillar collagen, expressed primarily in cartilage and sites of transition from cartilage to bone.

The classical fibrillar collagens are further sub-divided into the major (types I, II, III) and minor (types V, XI) collagens (Kadler et al. 2007). The predominant major fibrillar collagen is type I collagen, consisting of two α1 (I) and one α2 (I) chains. It is responsible for the hierarchical structure and mechanical strength of non-cartilage connective tissue (Eriksen et al. 2002; Hoffmann and Gross 2007; Rees et al. 2009). Type II collagen, a homotrimer consisting of three α1 (II) chains, is the equivalent main structural collagen in cartilage (Kannu et al. 2010). Type III collagen, which consists of three α1 (III) chains, is also a major fibrillar collagen which forms heterotypic fibrils together with type I collagen, and is important in healing and during fibrillogenesis (Banos et al. 2008; Liu et al. 1995). It is thought that type III collagen regulates the diameter of type I collagen fibrils during development and healing by limiting lateral growth (Banos et al. 2008). Tissues with elastic properties, such as skin and arteries, also have a higher type III collagen content than less elastic tissues (Gelse et al. 2003). Type V collagen is a minor fibrillar collagen which is co-expressed with types I and III collagen, and also plays a role in the regulation of type I collagen fibril diameter (Birk 2001; van der Rest and Garrone 1991). The predominant isoform of type V collagen is a heterotrimer of two α1 (V) and one α2 (V) chains, however other isoforms, which may contain α3 (V) chains, also exist (Birk 2001). Type XI collagen is the minor fibrillar collagen co-expressed with type II collagen in the cartilage (Kadler et al. 2007; van der Rest and Garrone 1991). Recent investigations on the function of type V collagen in tendon development

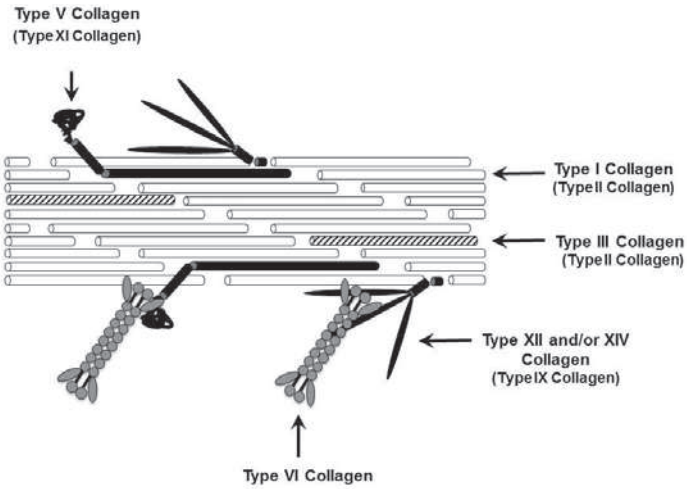

Figure 1. A schematic diagram of the collagen fibril, adapted from Collins et al. (2011). The major fibrillar type I (solid white cylinders) and III (hatched cylinders) collagen molecules comprise the majority of the fibril. The minor fibrillar type V collagen (solid black cylinders with globular ends) is embedded within the fibril with a protruding amino end. Type XII and XIV collagen (solid black cylinders with forked ends) belong to the fibril-associated collagens with interrupted triple helices (FACITs) sub-family. Beaded non-fibrillar type VI collagen (solid grey circles) interacts, at the surface of the fibril, with the other collagens that make up the fibril. Collagen types listed in parentheses are the cartilage specific collagens equivalent to those they are listed under. The proteins in this diagram are not drawn to scale.

suggest a synergistic role with type XI collagen in the regulation of fibrillogenesis and growth of mature tendon fibrils (Wenstrup et al. 2011). Type XXVII collagen, although not a classical fibrillar collagen, is nevertheless also a fibrillar collagen expressed primarily in cartilage and sites of transition from cartilage to bone, but also in skeletal muscle (Boot-Handford et al. 2003; Hjorten et al. 2007; Pace et al. 2003).

The FACITs, which include, amongst others, types IX, XII and XIV (Gelse et al. 2003), are important non-fibrillar collagens in mediating cell-matrix interactions between the collagen fibres and cell surfaces (Riley 2005), and during stabilisation of the attached collagen triple helices (Boudko et al. 2010; Ricard-Blum and Ruggiero 2005). Type VI collagen, a non-fibrillar beaded filament collagen, also facilitates fibrillar interaction with the extracellular matrix through interactions with types I, V, and XII collagens (Minamitani et al. 2004; Symoens et al. 2011; van der Rest and Garrone 1991). In addition, abnormal collagen fibrillogenesis in *COL6A1* deficient mice highlights the potential importance that type VI collagen may have in this biological process (Izu et al. 2010). Although type IV collagen is not a structural component of the fibril, it does form a major part of the basement membrane which surrounds and anchors the cellular component of musculoskeletal tissues, and maintains tissue architecture during development and wound healing (Kadler et al. 2007; Timpl and Brown 1996). In particular, the basement membranes are connected to cells by receptors which specifically bind to type IV collagen and other proteins. The short-chain collagen types VIII and X form networks in the basement membranes of non-cartilaginous and cartilaginous tissues respectively (Kadler et al. 2007). The non-fibrillar collagens are further reviewed by Ricard-Blum et al. (2005).

Typically, the polypeptide α-chains of the fibrillar collagens comprise of approximately 1000 amino acids in an uninterrupted repeating triplet, which contains a glycine amino acid in every third position with proline,

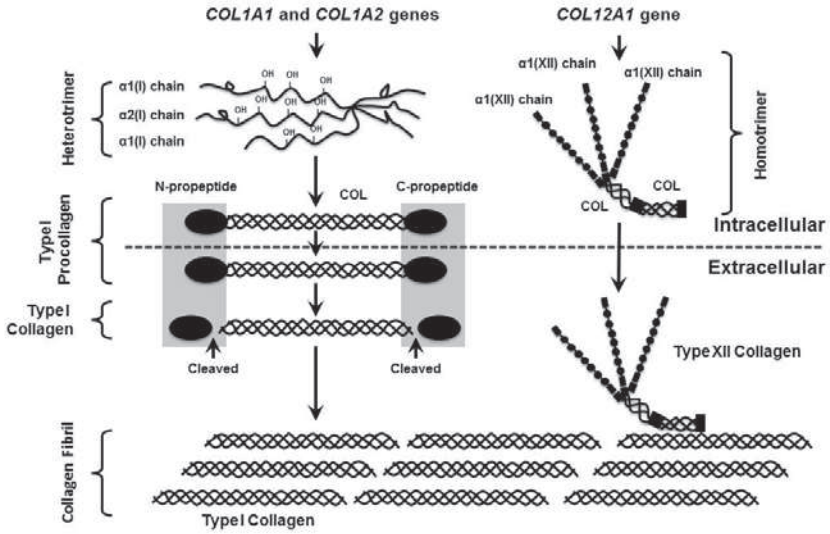

Figure 2. A schematic diagram of collagen fibril assembly. The polypeptide α-chains of the fibrillar collagens comprise an uninterrupted repeating triplet, which contains a glycine amino acid in every third position with proline and 4-hydroxyproline residues in the other two positions (Gly-X-Y). This is indicated by the –OH groups on the type I collagen α-chains. These repeating triplets form the triple helical domains in all the collagens. In the major fibrillar collagens, including type I collagen, the globular carboxy- and amino-terminal domains flanking the triple helix are post-translationally cleaved prior to their aggregation into collagen fibrils. Post-translational cleavage of the amino-terminal domain does not occur in some of the minor fibrillar (type V collagen) and non-fibrillar (type XII and XIV collagens) collagens. Type I and type XII collagen have been used as examples for the collagen fibril assembly process. The molecules and proteins in this diagram are not drawn to scale.

and 4-hydroxyproline residues in the other two positions (Gly-X-Y) (Kadler et al. 2007). These repeating triplets are essential for the formation of the triple helical domains in all the collagens (Figure 2) (Kadler et al. 2007). Most collagens occur as homotrimers of three identical α-chains, however heterotrimers of two or three different α-chains are also common (Kadler et al. 2007). These α-chains are wound together in a triple helix and, in the major fibrillar collagens, the globular carboxy- and amino-terminal domains flanking the triple helix are posttranslationally cleaved, prior to their aggregation into collagen fibrils (Figure 2) (Banos et al. 2008; Kadler et al. 2007; O'Brien 1997). The synthesis of the collagen proteins has been extensively reviewed by Banos et al. (2008).

Collagen Related Disorders

The major structural domains of collagen proteins are typically highly conserved across species, and each polypeptide α-chain is encoded by a specific gene (Table 2). In the case of homotrimeric collagens, all three α-chains are encoded by the same gene. For example, the α1 chains of type II collagen are encoded by the *COL2A1* gene. In heterotrimeric collagens, each different α-chain is encoded by a different gene. The *COL1A1* and *COL1A2* genes, for example, encode the α1 and α2 chains of type I collagen respectively. Rare mutations in many of these genes result in serious musculoskeletal disorders including, amongst others, osteogenesis imperfecta (Pace et al. 2003; Pollitt et al. 2006), chondrodysplasias (Warman et al. 1993) and Ehlers-Danlos syndrome (Byers et al.

1997; Malfait et al. 2005; Malfait et al. 2010; Malfait and De 2005; Smith et al. 1997) (Table 2). Symptoms of these disorders include bone fragility, joint hypermobility, skin hyperextensability and abnormal wound healing (Malfait et al. 2010; Pollitt et al. 2006). These severe clinical disorders highlight the importance of this protein family in the normal structure and function of musculoskeletal tissues. Since rare mutations within most of the genes encoding the collagens in the fibril cause severe disorders, it has been proposed that there is a lack of biological redundancy within the collagen fibril. Many of these collagen-encoding genes are therefore ideal candidates for association with less severe exercise-related traits (Ribbans and Collins 2013).

A Genetic Continuum

In addition to these severe rare Mendelian disorders, common variants within the collagen genes are also associated with a multitude of anomalous connective tissue traits (Table 2). Therefore, it is possible that a continuum of associated traits may exist for the range of genetic variation within these genes, encoding the fibrillar and non-fibrillar collagens. An example of this genetic continuum was recently proposed to describe traits associated

Table 2. Disorders resulting from rare mutations or common variants in the collagen gene(s) that encode for the non-cartilage and cartilage fibril, as well as the collagen network around the fibril producing cells

Type	Gene(s)	Mendelian Disorder(s)[1]	Complex Multifactorial Disorder(s)
I	COL1A1 COL1A2	Osteogenesis imperfecta, EDS[2], Infantile Cortical Hyperostosis	Osteoporotic Fractures, Osteoarthritis, Myocardial Infarction, Lumbar Disc Disease, Stress Urinary Incontinence
II	COL2A1	Collagenopathy, types II and III	Mild Spondyloepiphyseal Dysplasia Congenita, Nonsyndromic Cleft Palate, Osteoarthritis, Myopia
III	COL3A1	EDS (Hypermobility type; type 3), EDS (Vascular type; type 4)	Mitral Valve Prolapse, Pelvic Organ Prolapse, Cervical Artery Dissection, Rupture of Intracranial Aneurysms
IV	COL4A1 COL4A6	Porencephaly, Alport Syndrome, Hematuria, Brain Small Vessel Disease	Arterial Stiffness
V	COL5A1 COL5A2 COL5A3	EDS (Classic type; types 1 and 2)	Central Corneal Thickness, Ischemic Heart Disease, Cervical Artery Dissection, Rupture of Intracranial Aneurysms
VI	COL6A1 COL6A2 COL6A3	Bethlem Myopathy, Ullrich Congenital Muscular Dystrophy	Ossification of the Posterior Longitudinal Ligament (OPLL[3]) and the Ligamentum Flavum, Diffuse Idiopathic Skeletal Hyperostosis
VIII	COL8A1 COL8A2	Posterior Polymorphous Corneal Dystrophy, Fuchs' Endothelial Corneal Dystrophy	Primary Open Angle Glaucoma, Central Corneal Thickness
IX	COL9A1 COL9A2 COL9A3	Epiphyseal Dysplasia, Stickler Syndrome	Osteoarthristis, Lumbar Disc Degeneration
X	COL10A1	Metaphyseal Chondrodysplasia	Macular Degeneration, Osteoarhtritis
XI	COL11A1 COL11A2 COL2A1	Fibrochondrogenesis, Marshall Syn., Stickler Syn., Weissenbacher-Zweymuller Syn., Deafness, Otospondylomegaepiphyseal Dysplasia	Lumbar Disc Herniation, Limbus Vertebra, Lumbar Spine Stenosis, OPLL, Rheumatoid Arthritis
XII	COL12A1	Developmental Delay, Mild Dysmorphism, Lax Connective Tissue	Abnormal Corneal Endothelial Maturation
XIV	COL14A1	Punctate Palmoplantar Keratoderma, Recurrent Strokes	Abnormal Corneal Endothelial Maturation
XXVII	COL27A1	None identified to date	Tourette's Syndrome

[1] As listed in the OMIM database (OMIM. Accessed: 05/2013).
[2] EDS, Ehlers-Danlos Syndrome.
[3] OPLL, Ossification of the Posterior Longitudinal Ligament.

with variation within the *COL5A1* gene (Collins and Posthumus 2011). In this model, two functional copies of the *COL5A1* gene are required for normal development, structure and function of the connective tissue. Mutations that inactivate both copies of *COL5A1* result in death *in utero* (Wenstrup et al. 2006), while mutations, which inactivate a single copy of the gene (haploinsufficiency) will result in classical types of Ehlers-Danlos syndrome (EDS) (Malfait et al. 2010). At this end of the continuum the mutations in *COL5A1* result in death or EDS, which occur regardless of environmental exposure and independent of other non-genetic risk factors (Collins and Posthumus 2011). At the opposite end of the continuum, functional common variants within *COL5A1* contribute to more complex, less severe traits. These multifactorial traits arise as a result of the interaction between genetic and non-genetic factors, modifying physiological responses to environmental exposures (Collins and Posthumus 2011).

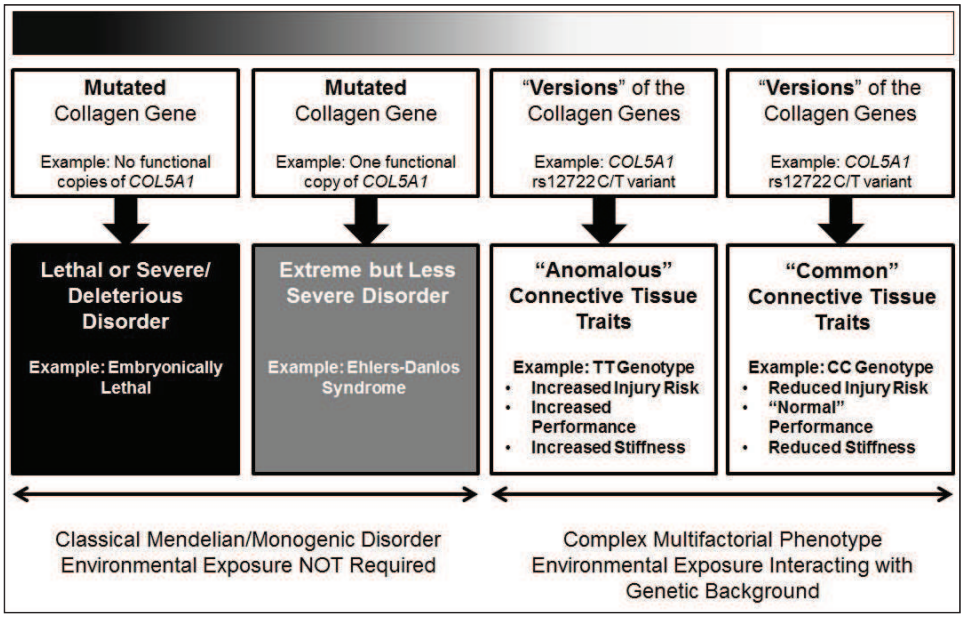

Figure 3. A proposed general genetic continuum for collagen genes, adapted from Collins et al. (2011). The black shading represents the lethal or severe traits due to mutations in the collagen genes. At this end of the continuum a single mutation results in the disorder. The white shading represents the most beneficial "versions" of the collagen genes. At this end of the continuum variants within collagen genes collectively contribute to the aetiology of the trait. The *COL5A1* gene and rs12722 gene variant are used as examples at each stage of the continuum.

Lethal mutations have also been described within other fibrillar collagen encoding genes, such as *COL1A1*, *COL1A2* and *COL2A1*, and severe musculoskeletal tissue disorders caused by non-lethal mutations, within most of the genes encoding components of the collagen fibril, have also been reported (Table 2). As with the *COL5A1* genetic continuum, these mutations result in death or severe disorders which occur regardless of environmental exposure and independent of other non-genetic risk factors (Figure 3). At the opposite end of the continuum, common variants within the same collagen genes contribute to less severe traits, such as increased susceptibility

to musculoskeletal tissue injuries, by modifying the effects of environmental exposures (Figure 3). Since the structure of the musculoskeletal system is an important factor in performance traits, we hypothesise that common variants within the collagen genes may also be associated with athletic ability. Like musculoskeletal tissue injuries, elite sporting performance is also a multifactorial trait resulting from the combination of numerous genetic and environmental factors, which interact with one another in a poorly understood but complex manner to mold athletic performance (Tucker and Collins 2012). This hypothesis is supported by the experimental studies investigating common variants in genes encoding fibrillar collagens with exercise-related traits, and will be expanded on in the following sections of this review.

Genes Encoding Fibrillar Collagens

Type I Collagen

Type I collagen, encoded by the *COL1A1* and *COL1A2* genes, is the predominant fibrillar collagen in non-cartilage connective tissues (Eriksen et al. 2002; Hoffmann and Gross 2007; Rees et al. 2009). Several polymorphic regulatory elements have been identified within the promoter and first intron of the *COL1A1* gene, which result in differing transcription factor binding affinities and in turn regulate α1 (I) chain production (Jin et al. 2009; Mann et al. 2001). One of the most extensively investigated functional polymorphic transcription factor binding sites is the Sp-1 binding site variant (rs1800012, G/T) within the first intron of the *COL1A1* gene. The G allele of this variant produces normal levels of *COL1A1* mRNA and the heterotrimeric protein consisting of two α1 (I) and one α2 (I) chains. However, the T allele has been shown to functionally increase Sp-1 binding to its binding element within the first intron, thereby increasing *COL1A1* mRNA and type I collagen protein translation, and resulting in formation of α1 (I) homotrimers. It has been proposed that tissues consisting of both hetero- and homotrimeric type I collagen are produced, which have different mechanical properties to the normal tissue only consisting of heterotrimers (Jin et al. 2009; Mann et al. 2001).

The association of this functional Sp-1 binding site variant with cruciate ligament injuries has been investigated in Swedish (Khoschnau et al. 2008), South African (Posthumus et al. 2009a) and Polish (Ficek et al. 2012) cohorts. In all three studies, the rare TT genotype was under-represented in participants with diagnosed cruciate ligament (predominately ACL) ruptures when compared to apparently healthy control participants with no history of cruciate ligament injuries (Ficek et al. 2012; Khoschnau et al. 2008; Posthumus et al. 2009a). In a combined analysis of all three studies, which all had similar genotype distributions, the TT genotype was significantly over-represented ($p < 0.0001$, odds ratio 18.4, 95.0% confidence interval 2.5 to 136.6) in the control groups (24 TT, 4.0%; 157 GT, 26.3% and 417 GG, 69.7%) when compared to the cruciate ligament rupture groups (1 TT, 0.2%; 133 GT, 30.2% and 307 GG, 69.6%) (Collins et al. 2010; Ficek et al. 2012; Khoschnau et al. 2008; Posthumus et al. 2009a)

In the original Swedish study, the TT genotype (0.1% shoulder dislocations versus 3.7% controls) was also reported to be significantly under-represented in participants with shoulder dislocations (Khoschnau et al. 2008). The most recent investigation of the Sp1 binding site variant (rs1800012) reported an interaction between this variant and another polymorphic regulatory element in the *COL1A1* promoter (rs1107946, G/T, position −1997) in modulating risk of ACL injury (Ficek et al. 2012). Specifically, the haplotype constructed from the rs1800012 T and rs1107946 G alleles, both known to functionally increase expression of *COL1A1*, was over-represented in the control participants when compared to individuals with cruciate ligament ruptures (Ficek et al. 2012) (Jin et al. 2009).

In addition, the *COL1A1* Sp1 binding site variant was also investigated for the risk of both acute Achilles tendon ruptures (n = 41) and chronic Achilles tendinopathy (n = 83) in a South African cohort (Posthumus et al. 2009c). Although no significant associations were identified for either the acute or chronic injuries, it is interesting to note that the TT genotype was absent in individuals diagnosed with acute Achilles tendon ruptures (0.0% ruptures, 2.4% tendinopathy and 4.8% controls) (Posthumus et al. 2009c). Furthermore, the studies have shown an increase in type I collagen mRNA expression in tendinopathic tendons (Riley 2004), while differential expression of type I collagen in anterior cruciate ligaments is associated with the degree of healing after a rupture (Lo et al. 2003). These findings highlight the potential differences in the genetic and biochemical components contributing to a risk of acute and chronic injuries, and require further investigation.

In addition to this injury mechanism specific association, it is also interesting to note that the TT genotype is associated with increased risk in a number of other pathologies (Jin et al. 2011). This highlights that certain gene variants may affect the aetiology of traits in different ways. The reasons for these differences remain unknown, however it may be the result of interactions with other gene variants or environmental stimuli.

Type V Collagen

As mentioned previously, type V collagen is known to interact with types I and III collagen and plays a role in the regulation of collagen fibrillogenesis (Birk 2001; van der Rest and Garrone 1991). The rs12722 C/T variant within the 3'-untranslated region (UTR) of the *COL5A1* gene, which encodes the α1 chain of type V collagen, was the first collagen gene variant reported to be associated with Achilles tendon injuries (Mokone et al. 2006). Specifically, the CC genotype was significantly over-represented in physically active healthy control participants with no self-reported history of Achilles tendon injury, when compared to participants with clinically diagnosed chronic Achilles tendinopathy in independent South African and Australian cohorts (Mokone et al. 2006; September et al. 2009). Several other variants, rs71746744 (-/AGGG), rs169399 (ATCT/-) and rs1134170 (A/T), within the *COL5A1* 3'-UTR were also found to independently associate with the risk of Achilles tendinopathy in the same Caucasian South African and Australian cohorts (Abrahams et al. 2013). Specifically, the *COL5A1* rs71746744 AGGG/AGGG, rs169399 -/- and rs1134170 TT genotypes were significantly over-represented in the participants with clinically diagnosed Achilles tendinopathy when compared to the controls in both cohorts, as well as in the combined cohorts (Abrahams et al. 2013).

Posthumus et al. (2009b) further investigated the *COL5A1* rs12722 C/T variant for an association with the risk of ACL injuries in a Caucasian South African cohort. The CC genotype was also significantly over-represented in apparently healthy female control participants (27.3%, 23 of 84) when compared with female participants with ACL ruptures confirmed at surgery (5.4%, 2 of 37), but not in a male cohort (Posthumus et al. 2009b). The reasons for this gender specific association remain unknown and require additional research.

Since the 3'-UTR of eukaryotic genes contains elements which are emerging as important post-transcriptional regulators (Mazumder et al. 2003; Xie et al. 2005), it has been suggested that the *COL5A1* 3'-UTR plays a role in the regulation of the *COL5A1* mRNA stability and by implication type V collagen production. In support of this, two major functional allelic forms of the *COL5A1* 3'-UTR, namely the C- and T-forms, have been identified and show significant differences in mRNA stability, with the T-form having increased stability (Laguette et al. 2011). The T- and C-allelic forms were predominately identified in participants diagnosed with Achilles tendinopathy and asymptomatic controls respectively (Laguette et al. 2011). The differences in the sequence of the two functional forms of the 3'-UTR

are determined by seven variants which include the four variants, rs12722 (C/T), rs71746744 (-/AGGG), rs169399 (ATCT/-) and rs1134170 (A/T), shown to be independently associated with chronic Achilles tendinopathy (Laguette et al. 2011). Based on these findings, Collins and Posthumus (2011) proposed that the relative content of type V collagen in tendons, ligaments and other tissues alters the fibril diameters and packing density within these tissues. This will alter their mechanical properties, susceptibility to injury and other exercise-related traits. To further explore this hypothesis, *COL5A1* 3'-UTR variants were investigated for associations with other exercise-related traits, namely range of motion (ROM) measurements, endurance performance and exercise-associated muscle cramps (EAMC).

Mutations within *COL5A1* cause classic EDS, a condition that is characterized by joint hypermobility (Malfait et al. 2010). Furthermore, it has also been proposed that *COL5A1* is associated with benign joint hypermobility syndrome (BJHS) (Grahame 2009). Altered musculotendinous flexibility, which is defined as ''the ability to move a joint through its complete range of motion'' (ROM), has been cited as an intrinsic risk factor for many of these injuries (Knapik et al. 1991; Knapik et al. 1992; Krivickas and Feinberg 1996). Although flexibility can be improved through regular stretching, it is at least in part also determined genetically (Hakim et al. 2004). Since variants within *COL5A1* may affect tissue architecture and biomechanical properties (Collins and Posthumus 2011), variants within this gene were investigated for associations with ROM measurements. The *COL5A1* rs12722 variant was associated with lower limb ROM measurements in a mixed injured/uninjured cohort (Collins et al. 2009). In addition, a follow-up study in an independent, larger cohort of apparently healthy physically active Caucasian South Africans reported that the *COL5A1* rs12722 CC genotype protected participants against an age-related decline in sit and reach (SR) ROM (Brown et al. 2011b). Since this functional variant was associated with SR ROM, it was recently investigated for effects on the mechanical properties of human tendons *in vivo* (Kubo et al. 2013). Specifically individuals with the *COL5A1* rs12722 CC genotype recorded significantly greater measurements of maximal knee extensor tendon elongation and strain when compared to their TC or TT counterparts (Kubo et al. 2013). No significant differences were recorded for the plantar flexors (Kubo et al. 2013). This study highlights that the level of effect of gene variants may differ between tissue structures, possibly due to differences in environmental stimuli, and adds further evidence to the proposed genetic continuum model outlined in Figure 3 (Collins and Posthumus 2011; Kubo et al. 2013).

Reduced ROM, specifically SR ROM, has also been associated with improved endurance running economy and performance (Craib et al. 1996; Jones 2002). This relationship is proposed to result from reduced aerobic demand due to increased elastic return, and may result in improved performance (Jones 2002). Given that the *COL5A1* rs12722 variant is associated with SR ROM (Brown et al. 2011b; Collins et al. 2009), it is not surprising that this variant was also associated with endurance running performance. The TT genotype of this variant was associated with faster run times in South African Ironman Triathletes, while triathletes with the CC genotype were over-represented in the slowest tertile for the run stage of this triathlon (Posthumus et al. 2011). The TT genotype was also associated with faster performance in a 56km ultra-marathon road race (Brown et al. 2011a). Furthermore, when these endurance runners were divided into four quadrants (Fast and inflexible, fast and flexible, slow and inflexible, slow and flexible) based on both their race finishing times and SR ROM measures, runners with the *COL5A1* rs12722 T allele were over-represented in the fast and inflexible quadrant (Brown et al. 2011a). It is interesting to note that variations in range of motion as well as increased intensity and/or duration of exercise are proposed to modulate the risk of exercise-related muscle cramps (EAMC). As such *COL5A1* rs12722, together with

other collagen gene variants, were investigated for an association with the risk of EAMC (O'Connell et al. 2013). The rs12722 TT genotype was significantly over-represented in athletes with a self-reported history of EAMC when compared to athletes with no self-reported life-long history of EAMC (O'Connell et al. 2013).

Type XI Collagen

Type XI collagen has been suggested to interact with type V collagen in the regulation of fibrillogenesis during tendon development (Wenstrup et al. 2011); and sequence variants within the genes encoding the α-chains of type XI collagen have also been investigated for association with AT in Caucasian South African and Australian populations (Hay et al. 2013). Although there were no independent genotype associations with the risk of chronic Achilles tendinopathy for the *COL11A1* rs3753841 (T/C), *COL11A1* rs1676486 (C/T) and *COL11A2* rs1799907 (T/A) functional variants, significant gene-gene interactions in modulating the risk of Achilles tendinopathy were observed (Hay et al. 2013). Specifically, the TCT-inferred pseudo-haplotype constructed from rs3753841, rs1676486 and rs1799907 was significantly over-represented in participants with clinically diagnosed Achilles tendinopathy (25.9%) when compared to control participants (17.1%) in a combined Caucasian South African and Australian cohort (Hay et al. 2013). Similarly, the TCT pseudo-haplotype was significantly over-represented in the Achilles tendinopathy participants when South African and Australian cohorts were analysed separately (Hay et al. 2013). Furthermore, the TCT (AGGG) pseudo-haplotype constructed from the type XI collagen genes and independently associated *COL5A1* rs71746744 (-/AGGG) variant was also significantly over-represented in the Achilles tendinopathy participants (25.2%) compared to the controls (9.1%) (Hay et al. 2013).

The proposed functional effects of the variants contributing to these inferred pseudo-haplotypes have been previously reported (Hay et al. 2013). Based on these proposed functions, it was hypothesised that this interaction between *COL11A1*, *COL11A2* and *COL5A1* may modulate a risk of chronic Achilles tendinopathy by affecting mRNA stability, thereby altering type V and/or XI collagen production, which in turn regulates collagen fibril diameter and changes the biomechanical properties of the collagen fibril (Hay et al. 2013). In addition, this interaction further corroborates the evidence supporting integrated roles of type V and type XI collagen in tendon fibrillogenesis (Wenstrup et al. 2011).

Type XXVII Collagen

As previously mentioned, the fibrillar type XXVII collagen is expressed primarily in cartilage and sites of transition from cartilage to bone, but also in skeletal muscle (Boot-Handford et al. 2003; Hjorten et al. 2007; Pace et al. 2003). Recently, Saunders et al. (2013) investigated variants within the *COL27A1* and *TNC* genes as risk factors for chronic Achilles tendinopathy in Caucasian South African and Australian populations. Both of these genes are situated close to *COL5A1* on chromosome 9, and a variant within the *TNC* gene was previously associated with chronic Achilles tendinopathy (Mokone et al. 2005; Saunders et al. 2013). Although no independent associations were identified for *COL27A1* rs946053 (G/T), this variant was implicated in a haplotype, with the *TNC* rs13321 and rs2104772 variants, associated with a risk of chronic Achilles tendinopathy (Saunders et al. 2013). This finding highlights the possibility that variants in *COL27A1* and its neighbouring genes also play a role in modulating risk of chronic Achilles tendinopathy, and suggests that *COL27A1* variants may be useful candidates in the investigation of other exercise-related traits.

Genes Encoding the FACITs and other Non-fibrillar Collagens

Type XII and XIV Collagen

FACIT collagens, such as type XII and XIV collagens, mediate cell-matrix interactions between the collagen fibres and cell surfaces (Riley 2005), and assist with stabilisation of the attached collagen triple helices (Boudko et al. 2010; Ricard-Blum and Ruggiero 2005). Furthermore, these collagens play a similar role to type V collagen in regulating collagen fibril assembly and diameter (Young et al. 2002). As such, September et al. (2008) investigated variants within both *COL12A1* and *COL14A1* for associations with a risk of ATin in SA population. However, no significant associations between *COL12A1* rs240736 (T/C), *COL12A1* rs970547 (G/A), *COL14A1* rs4870723 (A/C) *COL14A1* rs1563392 (T/A) and a risk of AT were observed (September et al. 2008). The *COL12A1* rs970547 variant was also investigated for an association with a risk of ACL rupture in a Caucasian South African cohort (Posthumus et al. 2010). Interestingly, a significant association was again only identified after gender stratification of the participants. The *COL12A1* rs970547 AA genotype was significantly over-represented in female participants with clinically diagnosed ACL ruptures when compared to apparently healthy female control participants (Posthumus et al. 2010).

Type VI Collagen

The studies identifying *COL5A1* rs12722 as a marker for endurance running performance suggest that other collagen gene variants may also be used as markers of endurance performance (Brown et al. 2011a; Posthumus et al. 2011). Given that *COL6A1* null mice run, on average, significantly less than wild-type mice (Bonaldo et al. 1998), the *COL6A1* gene was also investigated for an association with endurance performance. Caucasian South African triathletes with the *COL6A1* rs35796750 TT genotype completed the cycling stage of the triathlon significantly faster than triathletes with a TC or CC genotype (O'Connell et al. 2011). No associations were identified between *COL6A1* rs35796750 and performance in either the swimming or running stages of the triathlon. Further analysis of this variant and others within the *COL6A1* gene in more traditional cycling events is required to confirm this association.

Revisiting the Genetic Continuum

This review has identified a number of collagen gene variants which are associated with various exercise-related traits, such as endurance performance, risk of chronic Achilles tendinopathy and ACL rupture. Furthermore, mutations and common variants within these collagen genes are shown to cause a number of severe Mendelian disorders and are associated with several other complex multifactorial conditions, respectively (Table 2). Therefore, we propose that a general genetic continuum, as outlined in figure 3, exists for collagen genes and their associated traits. At one end of this general continuum model, a single mutation within one or more collagen genes will result in lethal or severe Mendelian disorders (Figure 3). At the other end of the continuum, functional variants within these collagen genes collectively contribute to the aetiology of anomalous multifactorial connective tissue traits. These multifactorial traits arise as a result of the interaction of genetic and non-genetic factors which modulate physiological responses to environmental stimuli (Figure 3).

Future Research

Further research is required to expand the genetic continuum model for collagen genes in exercise-related traits. One of the main limitations to this model is the lack of experimental data investigating the functional effects of these associated variants. Understanding the functional effects of these variants will not only corroborate the genetic associations already observed with multifactorial connective tissue traits, but also aid in elucidating the biological mechanisms through which these variants exert their effects. Although there is an increasing number of sequence variants shown to be associated with anomalous multifactorial connective tissue traits, further research is required to fully understand how these variants interact with one another and environmental risk factors to either mask or modify risk of exercise-related traits, as well as the extent to which environmental factors modulate genetic risk through epigenetics (Collins and Raleigh 2009; Ribbans and Collins 2013; September et al. 2006).

Acknowledgements

This study was supported in part by funds from the National Research Foundation (NRF) of South Africa, University of Cape Town, and the South African Medical Research Council. The grant holder acknowledges that opinions, findings and conclusions or recommendations expressed in any publication generated by the NRF supported research are that of the author(s), and that the NRF accepts no liability whatsoever in this regard. Author MC and Dr Alison September have filed patents on the application of specific sequence variations, which include some of those mentioned in this manuscript, related to risk assessment of Achilles tendon and anterior cruciate ligament injuries.

References

Abrahams Y., Laguette M.J., Prince S., Collins M. Polymorphisms within the *COL5A1* 3'-UTR That Alters mRNA Structure and the MIR608 Gene are Associated with Achilles Tendinopathy. Ann. Hum. Genet. 2013; 77 (3): 204–214.

Banos C.C., Thomas A.H., Kuo C.K. Collagen fibrillogenesis in tendon development: current models and regulation of fibril assembly. Birth Defects Res. C. Embryo. Today. 2008; 84 (3): 228–244.

Birk D.E. Type V collagen: heterotypic type I/V collagen interactions in the regulation of fibril assembly. Micron. 2001; 32 (3): 223–237.

Bonaldo P., Braghetta P., Zanetti M., Piccolo S., Volpin D., Bressan G.M. Collagen VI deficiency induces early onset myopathy in the mouse: an animal model for Bethlem myopathy. Hum. Mol. Genet. 1998; 7 (13): 2135–2140.

Boot-Handford R.P., Tuckwell D.S., Plumb D.A., Rock C.F., Poulsom R. A novel and highly conserved collagen (pro(alpha) 1 (XXVII)) with a unique expression pattern and unusual molecular characteristics establishes a new clade within the vertebrate fibrillar collagen family. J. Biol. Chem. 2003; 278 (33): 31067–31077.

Boudko S.P., Zientek K.D., Vance J., Hacker J.L., Engel J., Bachinger H.P. The NC2 domain of collagen IX provides chain selection and heterotrimerization. J. Biol. Chem. 2010; 285 (31): 23721–23731.

Brown J.C., Miller C.J., Posthumus M., Schwellnus M.P., Collins M. The *COL5A1* Gene, Ultra-Marathon Running Performance, and Range of Motion. International Journal of Sports Physiology and Performance. 2011a; 6 (4): 485–496.

Brown J.C., Miller C.J., Schwellnus M.P., Collins M. Range of motion measurements diverge with increasing age for *COL5A1* genotypes. Scandinavian Journal of Medicine and Science in Sports. 2011b; 21 (6): e266–e272.

Byers P.H., Duvic M., Atkinson M., Robinow M., Smith L.T., Krane S.M., Greally M.T., Ludman M., Matalon R., Pauker S., Quanbeck D., Schwarze U. Ehlers-Danlos syndrome type VIIA and VIIB result from splice-junction mutations or genomic deletions that involve exon 6 in the COL1A1 and COL1A2 genes of type I collagen. Am. J. Med. Genet. 1997; 72 (1): 94–105.

Collins M., Mokone G.G., September A.V., Van der Merwe L., Schwellnus M.P. The *COL5A1* genotype is associated with range of motion measurements. Scandinavian Journal of Medicine and Science in Sports. 2009; 19 (6): 803–810.

Collins M., Posthumus M. Type V collagen genotype and exercise-related phenotype relationships: a novel hypothesis. Exercise and Sport Sciences Reviews. 2011; 39 (4): 191–198.

Collins M., Posthumus M., Schwellnus M.P. The COL1A1 gene and acute soft tissue ruptures. Br. J. Sports Med. 2010; 44 (14): 1063–1064.

Collins M., Raleigh S.M. Genetic risk factors for musculoskeletal soft tissue injuries. Med. Sport Sci. 2009; 54: 136–149.

Craib M.W., Mitchell V.A., Fields K.B., Cooper T.R., Hopewell R., Morgan D.W. The association between flexibility and running economy in sub-elite male distance runners. Medicine and Science in Sports and Exercise. 1996; 28 (6): 737–743.

Eriksen H.A., Pajala A., Leppilahti J., Risteli J. Increased content of type III collagen at the rupture site of human Achilles tendon. J. Orthop. Res. 2002; 20 (6): 1352–1357.

Ficek K., Cieszczyk P., Kaczmarczyk M., Maciejewska-Karlowska A., Sawczuk M., Cholewinski J., Leonska-Duniec A., Stepien-Slodkowska M., Zarebska A., Stepto N.K., Bishop D.J., Eynon N. Gene variants within the COL1A1 gene are associated with reduced anterior cruciate ligament injury in professional soccer players. J. Sci Med Sport. 2012: 1440-2440.

Gelse K., Poschl E., Aigner T. Collagens-structure, function, and biosynthesis. Adv. Drug Deliv. Rev. 2003; 55 (12): 1531–1546.

Grahame R. Joint hypermobility syndrome pain. Curr. Pain Headache Rep. 2009; 13 (6): 427–433.

Hakim A.J., Cherkas L.F., Grahame R., Spector T.D., MacGregor A.J. The genetic epidemiology of joint hypermobility: a population study of female twins. Arthritis Rheum. 2004; 50 (8): 2640–2644.

Hay M., Patricios J., Collins R., Branfield A., Cook J., Handley C.J., September A.V., Posthumus M., Collins M. Association of type XI collagen genes with chronic Achilles tendinopathy in independent populations from South Africa and Australia. Br. J. Sports Med. 2013; 47 (9): 569–574.

Hjorten R., Hansen U., Underwood R.A., Telfer H.E., Fernandes R.J., Krakow D., Sebald E., Wachsmann-Hogiu S., Bruckner P., Jacquet R., Landis W.J., Byers P.H., Pace J.M. Type XXVII collagen at the transition of cartilage to bone during skeletogenesis. Bone. 2007; 41 (4): 535–542.

Hoffmann A., Gross G. Tendon and ligament engineering in the adult organism: mesenchymal stem cells and gene-therapeutic approaches. Int. Orthop. 2007; 31 (6): 791–797.

Izu Y., Ansorge H.L., Zhang G., Soslowsky L.J., Bonaldo P., Chu M.L., Birk D.E. Dysfunctional tendon collagen fibrillogenesis in collagen VI null mice. Matrix Biology. 2010; 30 (1): 53–61.

Jin H., Evangelou E., Ioannidis J.P., Ralston S.H. Polymorphisms in the 5' flank of COL1A1 gene and osteoporosis: meta-analysis of published studies. Osteoporos.Int. 2011; 22 (3): 911–921.

Jin H., van't Hof R.J., Albagha O.M., Ralston S.H. Promoter and intron 1 polymorphisms of COL1A1 interact to regulate transcription and susceptibility to osteoporosis. Hum. Mol. Genet. 2009; 18 (15): 2729–2738.

Jones A.M. Running economy is negatively related to sit-and-reach test performance in international-standard distance runners. International Journal of Sports Medicine. 2002; 23 (1): 40–43.

Kadler K.E., Baldock C., Bella J., Boot-Handford R.P. Collagens at a glance. J.Cell Sci. 2007; 120 (Pt 12): 1955–1958.

Kannu P., Bateman J.F., Randle S., Cowie S., du S.D., McGrath S., Edwards M., Savarirayan R. Premature arthritis is a distinct type II collagen phenotype. Arthritis Rheum. 2010; 62 (5): 1421–1430.

Khoschnau S., Melhus H., Jacobson A., Rahme H., Bengtsson H., Ribom E., Grundberg E., Mallmin H., Michaelsson K. Type I collagen alpha1 Sp1 polymorphism and the risk of cruciate ligament ruptures or shoulder dislocations. Am. J. Sports Med. 2008; 36 (12): 2432–2436.

Knapik J.J., Bauman C.L., Jones B.H., Harris J.M., Vaughan L. Preseason strength and flexibility imbalances associated with athletic injuries in female collegiate athletes. Am. J. Sports Med. 1991; 19 (1): 76–81.

Knapik J.J., Jones B.H., Bauman C.L., Harris J.M. Strength, flexibility and athletic injuries. Sports Med. 1992; 14 (5): 277–288.

Krivickas L.S., Feinberg J.H. Lower extremity injuries in college athletes: relation between ligamentous laxity and lower extremity muscle tightness. Arch. Phys. Med. Rehabil. 1996; 77 (11): 1139–1143.

Kubo K., Yata H., Tsunoda N. Effect of gene polymorphisms on the mechanical properties of human tendon structures. Springer Plus. 2013; 2 (1): 343.

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

Liu S.H., Yang R.S., al-Shaikh R., Lane J.M. Collagen in tendon, ligament, and bone healing. A current review. Clin. Orthop. Relat Res. 1995; (318): 265–278.

Lo I.K., Marchuk L., Hart D.A., Frank C.B. Messenger ribonucleic acid levels in disrupted human anterior cruciate ligaments. Clin. Orthop.Relat Res. 2003; 407: 249–258.

Malfait F., Coucke P., Symoens S., Loeys B., Nuytinck L., De P.A. The molecular basis of classic Ehlers-Danlos syndrome: a comprehensive study of biochemical and molecular findings in 48 unrelated patients. Hum. Mutat. 2005; 25 (1): 28–37.

Malfait F., De P.A. Molecular genetics in classic Ehlers-Danlos syndrome. Am. J. Med. Genet. C. Semin. Med. Genet. 2005; 139C (1): 17–23.

Malfait F., Wenstrup R.J., De P.A. Clinical and genetic aspects of Ehlers-Danlos syndrome, classic type. Genet. Med. 2010; 12 (10): 597–605.

Mann V., Hobson E.E., Li B., Stewart T.L., Grant S.F., Robins S.P., Aspden R.M., Ralston S.H. A COL1A1 Sp1 binding site polymorphism predisposes to osteoporotic fracture by affecting bone density and quality. J. Clin. Invest. 2001; 107 (7): 899–907.

Mazumder B., Seshadri V., Fox P.L. Translational control by the 3'-UTR: the ends specify the means. Trends Biochem. Sci. 2003; 28 (2): 91–98.

Minamitani T., Ikuta T., Saito Y., Takebe G., Sato M., Sawa H., Nishimura T., Nakamura F., Takahashi K., Ariga H., Matsumoto K. Modulation of collagen fibrillogenesis by tenascin-X and type VI collagen. Exp. Cell Res. 2004; 298 (1): 305–315.

Mokone G.G., Gajjar M., September A.V., Schwellnus M.P., Greenberg J., Noakes T.D., Collins M. The guanine-thymine dinucleotide repeat polymorphism within the tenascin-C gene is associated with achilles tendon injuries. The American Journal of Sports Medicine. 2005; 33 (7): 16–21.

Mokone G.G., Schwellnus M.P., Noakes T.D., Collins M. The *COL5A1* gene and Achilles tendon pathology. Scand J. Med. Sci. Sports. 2006; 16 (1): 19–26.

O'Brien M. Structure and metabolism of tendons. Scand. J. Med. Sci. Sports. 1997; 7 (2): 55–61.

O'Connell K., Posthumus M., Collins M. *COL6A1* gene and Ironman triathlon performance. Int. J. Sports Med. 2011; 32 (11): 896–901.

O'Connell K, Posthumus M, Schwellnus M.P., Collins M. Collagen genes and exercise-associated muscle cramping. Clin. J. Sport Med. 2013; 23 (1): 64–69.

Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD). Accessed: 05/2013. World Wide Web URL: www.omim.org.

Pace J.M., Corrado M., Missero C., Byers P.H. Identification, characterization and expression analysis of a new fibrillar collagen gene, *COL27A1*. Matrix Biol. 2003; 22 (1): 3–14.

Pollitt R., McMahon R., Nunn J., Bamford R., Afifi A., Bishop N., Dalton A. Mutation analysis of COL1A1 and COL1A2 in patients diagnosed with osteogenesis imperfecta type I–IV. Hum. Mutat. 2006; 27 (7): 716.

Posthumus M., Schwellnus M.P., Collins M. The *COL5A1* gene: a novel marker of endurance running performance. Medicine and Science in Sports and Exercise. 2011; 43 (4): 584–589.

Posthumus M., September A.V., Keegan M., O'Cuinneagain D., van der Merwe W., Schwellnus M.P., Collins M. Genetic risk factors for anterior cruciate ligament ruptures: COL1A1 gene variant. Br. J. Sports Med. 2009a; 43 (5): 352–356.

Posthumus M., September A.V., O'Cuinneagain D., van der Merwe W., Schwellnus M.P., Collins M. The *COL5A1* Gene Is Associated With Increased Risk of Anterior Cruciate Ligament Ruptures in Female Participants. Am. J. Sports Med. 2009b; 37 (11): 2234–2240.

Posthumus M., September A.V., O'Cuinneagain D., van der Merwe W., Schwellnus M.P., Collins M. The association between the *COL12A1* gene and anterior cruciate ligament ruptures. Br. J. Sports Med. 2010; 44 (16): 1160–1165.

Posthumus M., September A.V., Schwellnus M.P., Collins M. Investigation of the Sp1-binding site polymorphism within the COL1A1 gene in participants with Achilles tendon injuries and controls. J. Sci. Med. Sport. 2009c; 12 (1): 184–189.

Rees S.G., Dent C.M., Caterson B. Metabolism of proteoglycans in tendon. Scand. J. Med. Sci. Sports. 2009; 19 (4): 470–478.

Rest van der M., Garrone R. Collagen family of proteins. FASEB Journal. 1991; 5 (13): 2814–2823.

Ribbans W.J., Collins M. Pathology of the tendo Achillis: do our genes contribute? Bone Joint J. 2013; 95-B (3): 305–313.

Ricard-Blum S., Ruggiero F. The collagen superfamily: from the extracellular matrix to the cell membrane. Pathologie-Biologie. (Paris). 2005; 53 (7): 430–442.

Riley G. The pathogenesis of tendinopathy. A molecular perspective. Rheumatology (Oxford). 2004; 43 (2): 131–142.

Riley G. Chronic tendon pathology: molecular basis and therapeutic implications. Expert. Rev. Mol. Med. 2005; 7 (5): 1–25.

Saunders C.J., Van der Merwe L., Posthumus M., Cook J., Handley C.J., Collins M., September A.V. Investigation of variants within the *COL27A1* and TNC genes and Achilles tendinopathy in two populations. J. Orthop. Res. 2013; 31 (4): 632–637.

September A.V., Cook J., Handley C.J., Van der Merwe L., Schwellnus M.P., Collins M. Variants within the *COL5A1* gene are associated with achilles tendinopathy in two populations. British Journal of Sports Medicine. 2009; 43: 357–365.

September A.V., Mokone G.G., Schwellnus M.P., Collins M. Genetic risk factors for Achilles tendon injuries. International Sport Med. Journal. 2006; 7 (3): 201–215.

September A.V., Posthumus M., Van der Merwe L., Schwellnus M., Noakes T.D., Collins M. The *COL12A1* and *COL14A1* genes and Achilles tendon injuries. International Journal of Sports Medicine. 2008; 29 (3): 257–263.

Smith L.T., Schwarze U., Goldstein J., Byers P.H. Mutations in the COL3A1 gene result in the Ehlers-Danlos syndrome type IV and alterations in the size and distribution of the major collagen fibrils of the dermis. J. Invest Dermatol. 1997; 108 (3): 241–247.

Symoens S., Renard M., Bonod-Bidaud C., Syx D., Vaganay E., Malfait F., Ricard-Blum S., Kessler E., Van L.L., Coucke P., Ruggiero F., De P.A. Identification of binding partners interacting with the alpha1-N-propeptide of type V collagen. Biochem. J. 2011; 433 (2): 371–381.

Timpl R., Brown J.C. Supramolecular assembly of basement membranes. Bioessays. 1996; 18 (2): 123–132.

Tucker R. and Collins M. What makes champions? A review of the relative contribution of genes and training to sporting success. Br. J. Sports Med. 2012; 46 (8): 555–561.

Warman M.L., Abbott M., Apte S.S., Hefferon T., McIntosh I., Cohn D.H., Hecht J.T., Olsen B.R., Francomano C.A. A type X collagen mutation causes Schmid metaphyseal chondrodysplasia. Nat. Genet. 1993; 5 (1): 79–82.

Wenstrup R.J., Florer J.B., Davidson J.M., Phillips C.L., Pfeiffer B.J., Menezes D.W., Chervoneva I., Birk D.E. Murine model of the Ehlers-Danlos syndrome. *COL5A1* haploinsufficiency disrupts collagen fibril assembly at multiple stages. J. Biol. Chem. 2006; 281 (18): 12888–12895.

Wenstrup R.J., Smith S.M., Florer J.B., Zhang G., Beason D.P., Seegmiller R.E., Soslowsky L.J., Birk D.E. Regulation of collagen fibril nucleation and initial fibril assembly involves coordinate interactions with collagens V and XI in developing tendon. J. Biol. Chem. 2011; 286 (23): 20455–20465.

Xie X., Lu J., Kulbokas E.J., Golub T.R., Mootha V., Lindblad-Toh K., Lander E.S., Kellis M. Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. Nature. 2005; 434 (7031): 338–345.

Young B.B., Zhang G., Koch M., Birk D.E. The roles of types XII and XIV collagen in fibrillogenesis and matrix assembly in the developing cornea. J. Cell Biochem. 2002; 87 (2): 208–220.

Cite this article as: O'Connell K., Saunders C.J., Collins M. Collagen gene sequence variants in exercise-related traits. Centr Eur J Sport Sci Med. 2013; 1: 3–17.