A REVIEW OF STUDIES ABOUT THE GENES ENCODING THE COLLAGEN PROTEINS IN THE CONTEXT OF THE ANTERIOR CRUCIATE LIGAMENT RUPTURE

Piotr Szumiło

College of Physical Education and Sport, Szczecin University, Poland

Address for correspondence:
Piotr Szumiło
University of Szczecin, College of Physical Education and Sport
Felczaka 3c, 71–412 Szczecin, Poland
E-mail: coal@o2.pl

Abstract. ACL rupture is a common injury in professional sport as well as recreation. It happens most often during deceleration, lateral pivoting, or landing tasks. Most often it is a non-contact mechanism during which the knee is exposed to large forces. The main component of the ACL, reaching 75% of the content, is collagen. Ligaments are made of several types of collagen, which perform different functions. It has been proved that various variants of collagen genes and their interaction with other genes may significantly influence the risk of injury to the ACL. This publication contains a review of studies about polymorphisms of collagen genes in the context of ACL rupture.

Key words: ACL rupture, collagen genes, polymorphisms

Introduction

Sport and recreation are an important part of human life. Physical actvity is taken up by people of all ages and all abilities. Unfortunately, injuries are the worst aspect of sport and recreation. Damage to the soft tissues, such as tendons or ligaments, are common both in physical recreation as well as competitive sports (Collins et al. 2010). The knee is the most exposed to injury part of the body (Flørenes et al. 2010). Injuries to the knee have most commonly affected footballers and skiers as well as handball and volleyball players. One of the most serious knee injuries is rupture of anterior cruciate ligament (ACL) and this usually applies to handball and volleyball players (Majewski et al. 2006). Scandinavian studies show that this problem affects 34 people per 100,000 population, and in the risk group (sportsmen) the number is 85 per 100,000 inhabitants (Renstrom et al. 2008).

Ligaments are composed of densely arranged collagen fiber strands attached at both ends to the bone. Anterior cruciate ligaments are collagenous structures consisting of water and fibro-cartilaginous specific proteins, which build collagen fibrils (Hoffmann et al. 2007). A small quantity of all cells are cells called fibroblasts, which

are responsible for synthesis of the extracellular matrix surrounding (Trulsson et al. 2010). Collagen is the major component of ligaments and accounts for approximately 75% of the dry weight. The main structural components of this ligament are collagens type I, III–VI, XII and XIV, but also proteoglycans, such as decorin, lumican and versican, and glycoproteins, such as elastin, tenascin C and COMP (Frank 2004).

Collagen I (COL1A1)

Collagen I accounts for 85% of all collagens. Collagen I is a heterotrimer consisting of two α1 and one α2 chains. The α1 chain is encoded for by the *COL1A1* gene. *COL1A1* consists of 51 exons and is located on chromosome 17q21 (Zhang et al. 2011). *COL1A1* contains a functional Sp1 binding site polymorphism (single nucleotide polymorphism rs1800012) within its first intron (Mann et al. 2001). This polymorphism has recently been associated with cruciate ligament (CL) ruptures, shoulder dislocation (SD), anterior cruciate ligament (ACL) ruptures and Achilles tendon ruptures in two populations (Khoschnau et al. 2008; Posthumus et al. 2008, 2009). Moreover, previous studies suggested that the dysfunction of type I collagen genes was associated with disorders such as osteogenesis imperfecta, systemic diseases with scleral thinning, and myopia (Yang et al. 2011; Marijanović et al. 2010; McBrien et al. 2003). The G to T substitution in an intronic Sp1 binding site (rs 1800012), resulting in increased affinity for the transcription factor Sp1 and increased gene expression, has been one of the most extensively investigated polymorphisms within this gene (Meyer et al. 2008). Studies revealed that the rare TT genotype was shown to be underrepresented in subjects with acute soft tissue ruptures (Khoschnau et al. 2008).

In their studies, Khoschnau et al. (2008) identified one subject each with a TT genotype from 233 (0.4%) subjects with CL ruptures and 126 (0.8%) subjects with SD. Similar research conducted on the population of South Africa showed that the rare TT genotype was significantly underrepresented in the ACL rupture group (0 out of 117, 0%), compared with the controls (6 out of 130, 4.6%). In addition, rare TT genotype was absent in 117 subjectswith ACL ruptures (Posthumus et al. 2009). Two subjects (2.4%) with a TT genotype were identified in a cohort of 85 subjects with chronic Achilles tendinopathy (Posthumus et al. 2008). This sequence variant seems to be the first specific genetic element to be included in multifactorial models developed to understand the etiology and risk factors for ACL rupture (Posthumus et al. 2009). Large combined study including two populations (Swedish and South African) confirm that the TT genotype of the *COL1A1* Sp1 binding site polymorphism is underrepresented in acute soft tissue ruptures, in particular of the CL. The clinical relevance of this finding is that the COL1A1 TT genotype protects from an acute soft tissue rupture and should be included in future risk models for acute soft tissue ruptures (Collins et al. 2010).

ACL rupture is a particularly unfortunate injury for athletes. It has been described as one of the most severe injuries sustained in a sporting population (Brooks et al. 2005). Regular participation in organized sports may place an individual at up to 10 times greater risk of ACL rupture (Parkkari et al. 2008). Sports where sudden deceleration or change in direction is necessary, risk of ACL rupture is higher (Marshall et al. 2007). Skiers are one of the groups of athletes with the highest frequency of ACL rupture (Habelt et al. 2011). ACL injuries represent 17.2% of all ski injuries (Kim et al. 2012).

Stępień-Słodkowska et al. (2013) examined 138 male recreational skiers with surgically diagnosed primary ACL ruptures and 183 apparently healthy male skiers with a comparable level of exposure to ACL injury as a control group. DNA samples were genotyped for the *COL1A1 +1245G/T* polymorphisms. There was a significant difference in the genotype distribution between skiers and controls but there was no statistical difference in allele distribution.

This research shows that the risk of ACL ruptures was around 1.43 times lower in carriers of a minor allele G as compared to carriers of the allele T.

On the other hand, studies conducted by Collins et al. (2012) who examined *COL1A1 +1245G/T* polymorphism in ACL rupture group consisted of 112 skiers and snowboarders, compared to a control group of 190 which did not show any association between genotype and risk of ACL rupture.

In other studies, Stępień-Słodkowska et al. (in press) investigated *+1997G/T* polymorphisms in the *COL1A1* gene in two groups: 180 male and female recreational skiers with surgically diagnosed primary ACL ruptures and 245 apparently healthy male and female skiers with a comparable level of exposure to ACL injury, none of whom had any self-reported history of ligament or tendon injury. Results indicated a higher frequency of the *COL1A1* G allele and GG genotype in patients with ACL injury. The frequency of the G allele was higher in the cases (90.6%), and also statistically significant when compared with controls (84.7%). In conclusion, the *+1997G/T COL1A1* gene is one of the genetic markers to be taken into consideration in the identification of the risk of ACL injury.

ACL rupture is one of the most severe musculoskeletal soft tissue injuries in professional sport (Brooks et al. 2005). It is also a frequent injury among soccer players. Interesting studies including two polymorphisms in the gene *COL1A1* was conducted by Ficek et al. (2013). The association of *COL1A1 −1997G/T* (rs1107946) and *+1245G/T* (rs1800012) polymorphisms was examined. Recent studies comprising a transcription analysis that included *COL1A1 −1997G/T* and *COL1A1* Sp1 *+1245G/T* polymorphisms in the 5' flank of *COL1A1* revealed that the levels of transcription are influenced by the haplotype rather than genotype (Jin et al. 2009). This prompted the scientists to test the hypothesis that interaction between two or more polymorphisms within the *COL1A1* gene may influence the predisposition for ACL injury. Subjects were 91 male professional soccer players with surgically diagnosed primary anterior cruciate ligament ruptures. The control group consisted of 143 apparently healthy male professional soccer players, who were without any self-reported history of ligament or tendon injury. Both groups were from the same soccer teams, of the same ethnicity, of similar age and had a comparable level of exposure to the risk of ACL injury. Studies showed that the *COL1A1* G-T haplotype was associated with a reduced risk of ACL injury in a group of professional soccer players. Consequently, carriers of two copies of this haplotype may have a reduced risk of ACL injury. There is a suggestion that identifying the genetic profile associated with ACL ruptures via haplotype analysis have become a worthy alternative to single-locus analysis.

Collagen V (*COL5A1*)

Collagen V is a widely distributed quantitatively minor fibrillar collagen forming between 1–3% of total collagen content of tendon extracellular matrix (Chanut-Delalande et al. 2004). Type V collagen is co-expressed with type I collagen (Birk 2001). It intercalates with the type I collagen molecules to form heterotypic fibrils in non-cartilaginous connective tissues, where it modulates fibrillogenesis (Collins and Posthumus 2011). The predominant isoform of type V collagen is a heterotrimer consisting of two α1(V) and one α2(V) chains, which are encoded for by the *COL5A1* (9q) and *COL5A2* (2q) genes, respectively. Some isoforms contain an α3(V) chain encoded for by the *COL5A3* gene (19q) (Collins and Raleigh 2009). Collagen V plays an essential role in fibril assembly and lateral fibril growth within connective tissues (including tendon) (Riley 2004). Seven single nucleotide polymorphic sites are annotated within the 3'-UTR. The C to T single nucleotide polymorphism (SNP rs12722 or *Bst*UI RFLP) was commonly associated with ''exercise-related'' phenotypes (Collins and Posthumus 2011). Recently, rs12722 polimorphism in the 3'-untranslated region (3'-UTR) of *COL5A1* was shown to be associated with chronic Achilles

tendinopathy (Mokone et al. 2006; September et al. 2009) and Ehlers–Danlos syndrome (Wenstrup et al. 2006; Malfait et al. 2010).

Interestingly, *COL5A1* rs12722 polymorphism also seems to be associated with ACL rupture. Posthumus et al. (2009a) studied two groups: 129 white participants (38 women) with surgically diagnosed anterior cruciate ligament ruptures and 216 physically active control participants (84 women) without any history of ACL injury. All participants were genotyped for the *COL5A1 Bst*UI and *Dpn*II RFLPs. Reserch showed a significant difference in the *Bst*UI RFLP genotype frequency between the anterior cruciate ligament rupture and physically active control groups among the female participants, but not the male participants. The CC genotype in the female participants was significantly underrepresented in the anterior cruciate ligament rupture group compared with the controls. There were no differences in the *Dpn*II RFLP genotype distributions between the anterior cruciate ligament rupture and physically active control groups. This study proved that *COL5A1* rs12722 polymorphism is associated with the risk of anterior cruciate ligament ruptures in female athletes.

Case-control studies by Stępień-Słodkowska et al. (In press) confirmed there is an association with *COL5A1* rs12722 polymorphism and ACL injury in male population. However, *COL5A1 Bst*UI RFLP C/T and DpnII RFLP C/T were found to be in linkage disequilibrium. 138 male recreational skiers with surgically diagnosed primary ACL ruptures were tested. The control group consisted of 183 apparently healthy male recreational skiers, who were without any self-reported history of ligament or tendon injury. Results showed there was a tendency to underrepresentation of C-T haplotype in ACL rupture group compared to controls ($p = 0.091$). Higher frequency of the *COL5A1* C-T (*Bst*UI RFLP C/T and *Dpn*II RFLP C/T polymorphisms) haplotype suggests an association with reduced risk of anterior cruciate ligament injury in a group of apparently healthy male recreational skiers.

Collagen XII and XIV (*COL12A1* and *COL14A1*)

XII and XIV collagen are associated with the surface of the collagen fibril and are members of the Fibril Associated Collagens with Interrupted Triple helices (FACITs) sub-family (Shaw and Olsen 1991). Type XII and XIV collagens are homotrimers consisting of 3 α1(XII) chains (Dublet 1991). Collagen XII is encoded by a single gene, *COL12A1*, mapped to chromosome 6q12-q13 (Gordon et al. 1987). The *COL14A1* has been mapped on chromosome 8q23 and encodes the α1(XIV) chain (Schnittger et al. 1995). Similar to type V collagen, type XII collagen may be involved in fibrillogenesis (Birk et al. 1990; Young et al. 2002).

Previous studies examined the association of polymorphisms in the *COL12A1* and *COL14A1* genes with the risk of Achilles tendinopathy or Achilles ligament rupture (September et al. 2008). No spastically significant differences were identified in the genotype, allele or haplotype distributions between the subjects and control group.

Very few studies concern the relationship of *COL12A1* and risk of an anterior cruciate ligament rupture. Posthumus et al. (2010) investigated whether the sequence variants within *COL12A1* are associated with ACL ruptures. One hundred and twenty-nine subjects, including 38 females with ACL ruptures and 216 (83 females) physically active controls were included in this study. All participants were genotyped for the *Alu*I and BsrI restriction fragment length polymorphisms (RFLPs) within *COL12A1*. Studies showed that The *COL12A1 Alu*I RFLP is associated with ACL rupture. This relationship was significant only for female subjects. The results suggested that females with an AA genotype are at increased risk of ACL ruptures.

Ficek et al. (in press) conducted similar studies focusing on a group of Polish football players. *COL12A1 Alu*I RFLP (A9285G) polymorphism was tested as the one that may be associated with the risk of ACL rupture

(Posthumus et al. 2010). The genotype distribution in the study subjects were not different from those in controls and led to the conclusion opposite to Posthumus et al. (2010), that *COL12A1 Alu*I RFLP polymorphism is not associated with a predisposition for ACL injury.

Conclusions

Anterior cruciate ligament rupture is a multifactorial injury caused by several extrinsic and intrinsic risk factors (Griffin et al. 2006). Among the genetic risk factors, DNA sequence variants within genes that code for specific collagen proteins (collagen I and V) and tendon extracellular matrix proteins like collagen XII have recently been shown to be significantly associated with ACL rupture (Khoschnau et al. 2008; Posthumus et al. 2008, 2009, 2009a, 2010).

Genes that are also likely to be tied to the risk of the ACL rupture are the matrix metalloproteinase genes. MMPs are the main physiological mediators of extracellular matrix (ECM) degradation and remodeling (Birkedal-Hansen et al. 1993; Somerville et al. 2003). *MMP3* gene has recently been shown to associate with the risk of chronic Achilles tendinopathy (Raleigh et al. 2009). Moreover, it was also shown that the *MMP3* rs679620 variant and the *COL5A1* rs12722 variant interact to modify the risk of tendinopathy (Raleigh et al. 2009). The latest studies of Posthumus et al. (2012) have proved that four functional variants within four MMP genes clustered together on chromosome 11q22 are also associated with the risk of ACL ruptures.

The results of the above studies provide some practical implications. We already know the genes which contribute to higher ACL injury risk. In addition, studies suggest that ACL rupture is a polygenic trait (Ficek et al. 2013) and identifying the genetic profile associated with ACL ruptures via haplotype could be very beneficial for sport performance. Clarifying the complex relationship between gene variants and ACL rupture may be useful for coaches to optimize training for a particular athlete, and to reduce the risk of ACL rupture. However, further studies need to be done to fully explain the impact of gene variants on ACL rupture.

References

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

Birk D.E., Fitch J.M., Babiarz J.P., Doane K.J., Linsenmayer T.F. Collagen fibrillogenesis in vitro: interaction of types I and V collagen regulates fibril diameter. J. Cell Sci. 1990; 95: 649–657.

Birkedal-Hansen H., Moore W.G., Bodden M.K., Windsor L.J., Birkedal-Hansen B., DeCarlo A., Engler J.A. Matrix metalloproteinases: a review. Crit. Rev. Oral. Biol. Med. 1993; 4: 197–250.

Brooks J.H.M., Fuller C.W., Kemp S.P.T., Reddin D.B. Epidemiology of injuries in English professional rugby union: part 1 match injuries. Br. J. Sports Med. 2005; 39: 757–766.

Chanut-Delalande H., Bonod-Bidaud C., Cogne S., Malbouyres M., Ramirez F., Fichard A., Ruggiero F. Development of a functional skin matrix requires deposition of collagen V heterotrimers. Mol. Cell Biol. 2004; 24: 6049–6057.

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: 12888–12895.

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

Collins M. and Raleigh S.M. Genetic Risk Factors for Musculoskeletal Soft Tissue Injuries. Med. Sport Sci. 2009; 54: 136–149.

Collins M., O'Connell K., Stępień-Słodkowska M., Buryta M. Polymorphism +1245G/T within the *COL1A1* gene with the risk of ACL injuries in skiers and snowboarders. Zeszyty naukowe Uniwersytetu Szczecińskiego. 2012; 771 (28): 215–221.

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