Candidate Gene Approach Identifies Six SNPs in Tenascin-C (TNC) Associated With Degenerative Rotator Cuff Tears

Rainer Kluger,[1] Joerg Burgstaller,[2,3] Claus Vogl,[2] Gottfried Brem,[2,3] Michal Skultety,[1,2] Simone Mueller[2]

[1]Orthopedic Department of SMZOst Donauspital, Langobardenstrasse 122, A-1220 Vienna, Austria, [2]Institute of Animal Breeding and Genetics, University of Veterinary Medicine Vienna, Veterinärplatz1, A-1210 Vienna, Austria, [3]Institute of Biotechnology in Animal Production, Department IFA-Tulln, University of Natural Resources and Life Sciences, A-3430 Tulln, Austria

Received 15 September 2015; accepted 25 May 2016
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/jor.23321

ABSTRACT: Evidence for a heritable predisposition to rotator cuff tears (RCTs) is growing. Unrelated Caucasian individuals with surgically diagnosed full thickness RCTs (cases) and elderly Caucasian controls with intact rotator cuffs were screened for differences at the candidate genes: TNC, Col5A1, TIMP-1, MMP-1, MMP-2, MMP-3, MMP-9, and MMP-13. A first cohort (59 cases; 32 controls) was genotyped with the Sequenom MassARRAY iPLEX system. Of 142 SNPs within about 67–kbp of the TNC gene, 30 were tested for differences in proportions between cases and controls. A second, matched cohort (96 patients; 44 controls) was also genotyped for the same 30 SNPs, but with the KASP^{TM} genotyping technology. Combining the two cohorts and after Bonferroni correction, six SNPs were significantly associated with RCT. Compared to controls, RCT patients showed a significantly higher rate of homozygosity at rs72758637, rs7021589, and rs1138545; a significantly higher rate of heterozygosity at rs10759753, rs3789870, and rs7035322 and a higher minor allele frequency at rs3789870. Rs1138545, a missense SNP in exon10 might be of biological significance because it varies the amino acid sequence close to the TNC-FNIII5 domain. The FNIII5 domain binds multiple growth factors and co-ligates with integrins during tendon healing. © 2016 Orthopaedic Research Society. Published by Wiley Periodicals, Inc. J Orthop Res

Keywords: tenascin C; genetic variant; rotator cuff tear; single nucleotid polymorphism; FNIII5

The prevalence of rotator cuff tears (RCTs) in the general population is approximately 20%, increasing from 10% in the age group of 50–59 years to 50% in people aged over 80 years.[1] In addition to aging, factors such as trauma, overuse, mechanical subacromial bony conflict, and cigarette smoking contribute to the rotator cuff disease.[2,3] A potential heritable predisposition to RCTs has recently been suggested after excess familial clustering was found in 3,091 patients in a Utah Population Database.[4]

Genetic risk factors for RCT have been described in three studies. Peach et al. investigated patients with cuff tear arthropathy and found an association with variants in the ANKH and TNAP gene, encoding for proteins that regulate the extracellular concentration of inorganic pyrophosphate.[5] Motta et al. allocated single nucleotide polymorphisms (SNPs) as risk variants for RCT in genes encoding for FGFs, DEFB1, an immune modulating protein, and ESRRB, a protein involved in hypoxia related tissue response.[6] Teerlink et al. confirmed the association of ESRRB variants and RCTs.[7]

We selected candidate genes on the basis of preexisting association analyses for Achilles tendon ruptures (TNC, Col5A1, and MMP-3),[8–11] tendinopathies of the elbow (Col5A1),[12] ruptures of the posterior tibial tendon (MMP-1),[13] and included the matrix metalloproteinase genes MMP-1, MMP-2, MMP-3, MMP-9, MMP-13, TIMP-1 because they are specifically expressed in torn rotator cuff tendons.[14–16] For example, TNC dinucleotide repeat polymorphisms[8] as well as TNC SNPs[9] have been associated with Achilles tendinopathies and ruptures in South African, and Australien cohorts with European Caucasian ancestry. TNC is a modular ECM glycoprotein composed of a series of epidermal growth factor like repeats, fibronectin type III-like repeats and a C-terminal fibrinogen-like globular domain. TNC is specifically expressed upon tissue damage, potentially as a result of mechanical strain, loss of ECM integrity or oxidative stress. Expression is maintained throughout subsequent inflammation and repair by the action of numerous cytokines, and growth factors.[17–19] Achilles tendinopathies are also associated with putative risk variants in the Col5A1 gene[10] and/or in the MMP-3[11] gene. In addition, an association of Col5A1-SNPs with chronic degenerative tendon changes at the elbow (tennis elbow) has recently been reported.[12] The Col5A1 gene encodes for the pro-α1(V) chain of the heterotrimer type V collagen. Although type V collagen is a minor fibrillar collagen, it intercalates with type I collagen to form heterotypic fibrils and is therefore a critical structural component in fibrillogenesis. Matrix metalloproteinases MMP-2, MMP-9, and metallopeptidase inhibitor TIMP-1 are upregulated in human rotator cuff tendons following physical exercise.[14] In contrast MMP-1, MMP-3, MMP-9,[15,16] and MMP-13[16] are expressed in torn rotator cuff tendons. More specifically, MMP-1 and MMP-9 expression was found to be significantly higher in "intrinsic factor related" articular side tears than in "wear related" bursal side tears.[15] The objective of this study was to investigate selected SNPs in MMP-1, MMP-2, MMP-3, MMP-9, MMP-13, TIMP-1, TNC, and Col5A1 genes in patients with full thickness RCTs as well as healthy

Conflicts of interest: None.
Grant sponsor: Department of Orthopedics SMZOst Donauspital.
Correspondence to: Rainer Kluger (T: + 0043-1-28802743511; F: 0043-1-288023580; E-mail: Rainer.kluger@wienkav.at)

controls. All study participants were non smokers and none of them had a history of shoulder trauma. Of note, despite age being the major threat to rotator cuff integrity, a significant part of the population remains unaffected by this disease until their senium.[20] In an attempt to gain more insight into this phenomenon, we chose healthy control individuals who were on average two decades older than patients in the RCT group. Moreover RCTs of the herein investigated patients were either large or massive in size and affected mostly both shoulders.

METHODS

Study Design: Analytical Observational Study

Level of Evidence: 3 (Case-Control Study)

Study Population. All participants of the study completed medical history questionaire forms including medication use, personal history of systemic disease and inflammatory disease. The clinical orthopedic assessment covered ROM of joints, history of tendon pathologies (tendinopathies, tendon ruptures), and sports participation. Prior to participation in this study all participants gave informed written consent. This study was approved by the ethics committee of the Medical University Vienna (reference number EK: 06/017-VK).

RCT patients ($n = 155$) had either a large to massive rotator cuff tear in one shoulder ($n = 25$) or large to massive tears in both shoulders ($n = 130$) (tear size classification according to Cofield). Diagnosis was made during arthroscopic surgery at our institution. Patients were excluded if the cuff tear was combined with a history of calcifying tendinitis, trauma or systemic disease/inflammatory condition. Altogether 76 control participants were recruited from two different Viennese institutions for retired people. All control individuals had intact supraspinatus-, infraspinatus, and subscapularis tendons in both shoulders. The diagnoses were made through ultrasonography assessment, performed by an experienced orthopedic physician (R.K.). Exclusion criteria for control participants were prior operations of either shoulder, a history of a humeral fracture or an infiltration or conservative shoulder treatment in the last 24 months or a systemic disease/inflammatory condition.

The initial cohort consisted of 59 consecutive patients (cases), who were all nonsmoking from the Vienna area and European Caucasian ancestry (mean age 71 ± 7.1 years, 46% male) and 32 consecutive controls participants, who were all nonsmoking from the Vienna area and European Caucasian ancestry (mean age 87 ± 5.7 years, 27% male). The body mass index was 25.8 kg/cm^2 (range 16.0–36.2 kg/cm^2) in patients and 24.8 kg/cm^2 (15.4–36.2 kg/cm^2) in control individuals. A subsequent independent replication cohort consisted of 96 consecutive cases, who were all nonsmoking from the Vienna area and European Caucasian ancestry (mean age 65 ± 8.2 years, 40% male), and 44 consecutive controls who were all nonsmoking from the Vienna area and European Caucasian ancestry (mean age 87 ± 6.5 years, 25% male). In this cohort, the body mass index was 25.2 kg/cm^2 (range 18.2 –37.3 kg/cm^2) in patients and 22.2 kg/cm^2 (16.5–35.2 kg/cm^2) in controls. All cases and controls were matched for confounding variables for RCT: BMI, arm dominance, heavy manual labor, shoulder trauma, and smoking status. Arterial hypertension was not assessed but there was a trend that elderly participants used antihypertensive drugs more often than younger participants. We did not match the age of the cohorts for reasons explained below.

SNP Selection. Initially, 79 SNPs covering the exons and the exon-intron border regions of the eight candidate genes TNC, COL5A1, TIMP-1, MMP-1, MMP-2, MMP-3, MMP-9, and MMP-13 were selected after being classified as markers for putative coding mutations, non-synonymous (missense mutations), STOP-gained, STOP-lost, splice-site or frameshift variants.[21] The NCBI database[21] was used for annotation and positioning of the SNPs.

DNA Extraction. Genomic DNA was extracted from white blood cells using GenElute Mammalian genomic DNA Miniprep kit (Sigma–Aldrich, St.Louis, MO) according to the manufacturer's instructions.

Genotyping. For the first cohort (59 patients; 32 controls) genotyping was performed with the Sequenom MassARRAY iPLEX system (Sequenom, Hamburg, Germany) at the Department for Agrobiotechnology, IFA Tulln (Austria). A section of DNA containing the variant position was amplified from each individual by PCR, before a high-fidelity single-base primer extension reaction over the SNP being assayed was undertaken, using nucleotides of modified mass. The different alleles therefore produce oligonucleotides with mass differences that can be detected using highly accurate Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight mass spectrometry.[22] The SNP multiplex assays were designed using the Sequenom Assay Design 3.1 software, PCR primers, and extension primers. SNP genotyping was performed using the iPLEX® GOLD Complete Genotyping kit with SpectroCHIPs® II in the 384 format (Sequenom, Hamburg, Germany). We followed the manufacturer's protocol with a single modification: To reduce unspecific primer extension, 5 ng sheared salmon sperm DNA (Invitrogen, Lofer, Austria) per reaction was added to the PCR mastermix. Results were analysed with the Sequenom Typer 4.0 software (Sequenom, Hamburg, Germany).

Sequence Analysis. Selected regions were analysed by traditional Sanger sequencing. For amplification and sequencing we used the following primers: TNC1 F1 CCCAAAGGTCCCTAGATCCAA, TNC1 R1 GAGTAACCATGGCCTCAGGTTC, TNC2 F1 TCACCCAGCCTCTCAGTGGTAT, TNC2 R1 AATCCGGAAGCTCTCCACTTG, TNC3 F1 CTTAGGAAGTCATT GTGGGTCCA, TNC3 R1 TCTTCTCCTCCCCTGC TTTG. PCR products were sequenced directly using the ABI PRISM Ready Reaction DyeDeoxy Terminator Cycle Sequencing Kit (Applied Biosystems, Life Technologies, Carlsbad, CA) and an ABI 3,730 XL platform (Applied Biosystems, Life Technologies). The resulting sequences were aligned using the NCBI Basic Local Alignment Search Tool (BLAST, http://blast.ncbi.nlm.nih.gov/ Blast.cgi) and CodonCode Aligner, version 1.2.1, (CodonCode Corporation, Centerville, MA). For reference NCBI GenBank sequence NG_029637.1 was used. In the replication study (96 patients; 44 controls) genotyping was performed using KASP™ genotyping technology at the LGC Genomics Laboratories (Queens Road, Teddington, Middlesex, TW11 0LY, UK) (for details see: http://www.kbioscience.co.uk/reagents/KASP.html).

Statistical Analysis. Power calculation was performed with the OPENEPI software (http://www.openepi.com/Power/PowerCC.htm). For the final statistical analysis the two cohorts where combined and analysed jointly since they were sampled within 3 years from the same population. Genotypes of patients and controls were tested separately and jointly for deviation from Hardy–Weinberg proportions with a chi-square test. With all chi-square tests, p-values were estimated using simulations. Deviations between the patient and control groups in genotype and allele frequencies were also tested with chi-square tests. A Bonferroni correction changed the nominal significance level from 0.05 to $0.05/30 = 0.00167$, because 30 association tests of genotypes with phenotypes were performed in this test family. As a measure of linkage disequilibrium (LD), the correlation coefficient between pairs of biallelic loci (SNPs) was used. In the LD analyses, only SNPs were included, where all three genotypes were present in at least one of the two groups and the frequency of the minor allele in the sample was at least five. An r^2-cutoff of 0.8 was used to cluster the SNPs into linkage groups. Within linkage groups, we inferred haplotypes using an EM algorithm (function "em.haplo" in the R-package "haplo.stat", Sinnwell JP, and Schaid DJ (2015). haplo.stats: Statistical Analysis of Haplotypes with Traits and Covariates when Linkage Phase is Ambiguous. R package version 1.7.1. http://CRAN.R-project.org/package = haplo.stats).

RESULTS

Candidate Gene Search

Of the 79 SNPs in the eight candidate genes TNC, TIMP-1, Col5A1, MMP-1, MMP-2, MMP-3, MMP-9 and MMP-13, 21 were found to be polymorphic, and 17 of these SNPs were included in a first statistical analysis. (Suppl. Table S1) Tests for deviation from Hardy–Weinberg equilibrium showed only one SNP in MMP 9 to deviate from equilibrium in the cases. Although we conclude from statistical reasoning and the tests that we used, that this result is expected by chance, we plan to investigate MMP9 loci in a future study. Differences in genotype and allele frequencies between patient and control groups did not differ significantly for SNPs in MMP-1, MMP-2, MMP-3, MMP-9, MMP-13, and Col5A1 genes, while the two coding SNPs rs13321 and rs1138545 in the TNC gene on chromosome 9q31-34 differed significantly between patients and controls.

In Depth Analysis of TNC

We therefore focused on the TNC gene and checked a total of 142 SNPs (30 HAPMAP tagging and 112 additionals) spanning from map position 115 087 373 in exon 3–115 020 084 in the 3′-UTR region (=67 010 bp) of the TNC gene. Of these, 52 SNPs turned out to be polymorphic in our sample. To exclude that our initial data originated from paralogous sequences, to confirm alignment to the human reference TNC gene, as well as to double check our Sequenom MassARRAY findings, we resequenced 4,016 bps of the TNC gene by classical Sanger sequencing whereby 77 SNPs (26 from HAPMAP tagging and 51 additionals) were reassessed. Sixteen SNPs were non-synonymous, four SNPs synonymous, two splice donor variants and the remaining 55 SNPs intronic. Eventually 30 polymorphic SNPs met the inclusion criteria for an in depth statistical analysis. We genotyped all samples for these SNPs. Some SNPs showed an excess, others a scarcity of heterozygotes compared to the expected HW proportions. Altogether, eight of the 30 SNPs showed deviations from HWE in either groups (Table 1).

Gene Structure of Assessed TNC Search Area

We inferred pairwise linkage disequilibria (LD) in the region (Suppl. Fig. S1; Suppl. Table S3). Interestingly the assessed gene segment contains groups of SNPs characterized by a complete or nearly complete LD and similar allele frequencies. Using an r^2-cutoff of 0.8, five such groups were inferred for grouping (Fig. 1). Moreover the positions of the SNPs belonging to different LD groups interdigitate in this segment of the TNC gene. Among some SNPs of different LD groups pairwise associations were found (e.g., SNPs in the groups LG1, LG2, LG5 (Fig. 1)); others showed no association. With so many interdigitating, relatively independent LD groups, confident inference of haplotypes from the unphased patient data for the whole region seemed useless. Instead we inferred haplotypes within some LD groups using an EM algorithm mentioned above. In LD group 2 (Fig. 1-LG2), the major alleles of all four SNPs are nearly perfectly associated as are the minor alleles, such that the two haplotypes containing the major and minor alleles, respectively, make up an expected proportion of 0.98 (Suppl. Table S4B). The associations in LD group 1 (Fig. 1-LG1) are less tight, but nevertheless the two haplotypes corresponding again to the major and minor alleles combine to a proportion of 0.82 (Suppl. Table S4A). We note again that especially these two groups are interdigitating (Fig. 1). We note that we Sanger-sequenced over boundaries of LD groups and thus confirm the correct alignment of this part of the gene.

Comparison of TNC Gene Region With HAPMAP

This sequence of positions of SNPs is also found in the HAPMAP database and in other hominids, specifically, in the alignment of *H. sapiens* and *P. abelii* DNA (conservation 96%), and the RNAs of *Homo sapiens* and *Pongo abelii* (conservation 98%), *Gorilla gorilla* (conservation 99%), *Pan troglodytes* (conservation 99%), and *Pan pansicus* (conservation 99%).

Association of Six SNPs With the Degenerative RCT Phenotype

With all 155 patients and 76 controls, 15 SNPs were significantly associated with the phenotype; after Bonferroni correction six variants remained significant (Table 1 and Table 2). Note that since the genes cluster into LD groups with similar allele frequencies the Bonferroni correction, which assumes independence, is overly conservative.

Table 1. SNPs on the *Homo sapiens* Tenascin C (TNC) (gi|342837707|ref|NG_029637.1|), RefSeqGene on Chromosome 9 (Genome Reference Consortium Human Build 38 Patch Release 2)

				Genotype Frequ.				Genotype Frequ.					
	Alleles			Controls				Cases				Controls vs. Cases	
SNP	maj/min	Map pos.	HWE. all	*n*	hom1	hom2	HWE	*n*	hom1	hom2	HWE	Genotype Frequ.	Allele Frequ.
rs1757095	[G>A]	115086115	1.000	68	0.912	0.000	1.000	151	0.861	0.007	1.000	0.574	0.297
rs1061494	[A>G]	115084301	0.012	73	0.233	0.164	0.107	142	0.317	0.127	0.051	0.391	0.257
rs2225330	[C>T]	115074121	0.830	72	0.681	0.056	0.234	158	0.652	0.032	0.805	0.551	1.000
rs2210108	[C>T]	115073909	0.875	74	0.635	0.068	0.307	156	0.391	0.109	0.298	0.002	0.003
rs1138545	[G>A]	115073620	0.808	75	0.573	0.000	0.033	158	0.759	0.044	0.016	**0.001***	0.062
rs3789870	[C>T]	115072997	0.653	71	0.662	0.070	0.140	156	0.372	0.135	0.495	**0.000***	**0.001***
rs3789871	[T>C]	115072975	0.021	71	0.845	0.028	0.078	152	0.882	0.013	0.137	0.692	0.441
rs17240680	[A>T]	115047838	0.774	72	0.750	0.014	0.592	158	0.753	0.013	0.470	1.000	1.000
rs2104772	[T>A]	115046506	0.893	75	0.267	0.240	1.000	157	0.389	0.140	1.000	0.077	0.025
rs11793430	[C>G]	115045841	0.726	74	0.446	0.054	0.095	156	0.622	0.058	0.483	0.029	0.051
rs10817704	[C>T]	115045801	0.011	74	0.919	0.014	0.124	157	0.911	0.013	0.054	1.000	1.000
rs10122770	[G>A]	115045229	1.000	76	0.895	0.000	1.000	158	0.905	0.006	0.347	0.870	1.000
rs953288	[T>G]	115043440	0.691	71	0.451	0.169	0.190	156	0.244	0.231	0.528	0.008	0.007
rs10759753	[T>C]	115043015	0.766	72	0.639	0.069	0.276	158	0.380	0.108	0.179	**0.001***	0.002
rs72758637	[G>C]	115042922	1.000	75	0.547	0.000	0.018	151	0.762	0.040	0.049	**0.000***	0.024
rs7021589	[A>G]	115042388	1.000	76	0.566	0.000	0.019	158	0.778	0.038	0.023	**0.000***	0.021
rs1061495	[A>G]	115042265	1.000	73	0.630	0.055	0.718	155	0.645	0.032	0.799	0.756	0.702
rs1411456	[C>T]	115041603	1.000	73	0.644	0.068	0.279	157	0.395	0.108	0.382	0.002	0.003
rs10759752	[T>C]	115041441	0.788	73	0.425	0.178	0.209	158	0.247	0.222	0.442	0.007	0.006
rs10817703	[C>A]	115040706	0.049	72	0.931	0.000	1.000	144	0.924	0.014	0.021	0.728	0.799
rs10982503	[G>A]	115036558	0.849	73	0.562	0.068	1.000	146	0.616	0.034	0.457	0.434	0.332
rs12347433	[A>G]	115035318	0.712	75	0.560	0.080	0.556	157	0.611	0.032	0.463	0.282	0.238
rs1330361	[C>T]	115035115	0.113	72	0.708	0.042	0.674	153	0.752	0.033	0.173	0.744	0.481
rs16932078	[G>A]	115034707	0.654	70	0.543	0.057	0.772	158	0.380	0.133	0.604	0.039	0.012
rs3789875	[G>T]	115033009	1.000	73	0.589	0.082	0.352	156	0.609	0.032	0.474	0.273	0.463
rs72758634	[C>A]	115031134	0.795	71	0.606	0.000	0.055	158	0.766	0.038	0.039	0.002	0.123
rs13321	[G>C]	115030304	0.885	74	0.581	0.054	1.000	158	0.380	0.133	0.617	0.009	0.003
rs7035322	[G>T]	115024894	0.528	71	0.662	0.070	0.132	150	0.407	0.080	0.068	**0.001***	0.005
rs10982497	[G>A]	115022769	0.853	67	0.597	0.075	0.514	152	0.612	0.033	0.478	0.427	0.532
rs10982496	[C>G]	115020203	0.139	73	0.425	0.094	0.081	152	0.283	0.016	0.630	0.111	0.091

The first column corresponds to the SNP-code; the second to the mutation (major /minor); the third to the position according to GRCh38.p2; the fourth to the *p*-value of the test of Hardy–Weinberg equilibrium in all individuals (HWE.all); the next four columns refer to the control group, with the number, the proportion of homozygous genotypes of the major (hom1) and minor (hom2) alleles, and the test for HWE (HWE.controls); the next four columns refer to the same information for the patient group; the last two columns refer to chi-square tests of differences between patients and controls with respect to genotypes and alleles. A boldface *p*-value with an asterisk "*" indicates significance after Bonferroni. Note that while the nominal significance level alpha is 0.05, the Bonferroni-corrected alpha level is 0.00167 as explained above. The MAF frequencies and the gene annotation of all 30 SNPs are shown in Suppl. Table S2. Details of genotype and minor allele frequency distributions of the six (after Bonferroni correction) significantly different TNC SNPs are presented in Table 2.

For SNPs in LD group1 (Fig. 1-LG1) healthy controls were significantly more often homozygous than cases, for the two intronic SNPs rs10759753 and rs3789870 also after Bonferroni correction. The intronic rs7035322 is also linked to LD group 1 (maximal pairwise $r^2 > 0.7$) and is also significantly associated with the phenotype after Bonferroni correction. Moreover at rs3789870 the minor allele (T) was significantly more frequent in RCT patients, again even after Bonferroni correction (Table 2).

For SNPs in LD group 2 (Fig. 1-LG2), healthy controls were significantly more often heterozygous than cases, for the three SNPs rs72758637, rs7021589, and rs1138545 even after Bonferroni correction (Table 2).

Potential Biological Significance

Of the six SNPs that are associated with the RCT phenotype rs7021589 and rs1138545 might have biological significance. Intronic rs7021589 is a protein quantitative trait locus (PQTL) with a cis-regulatory function on TNC protein plasma levels. Rs1138545 (G > A) is a missense variant located in exon10. The triplet C(G)T is transcribed into an arginine (R), the variant C(A)T into a histidine (H) (Fig. 2). Rs1138545

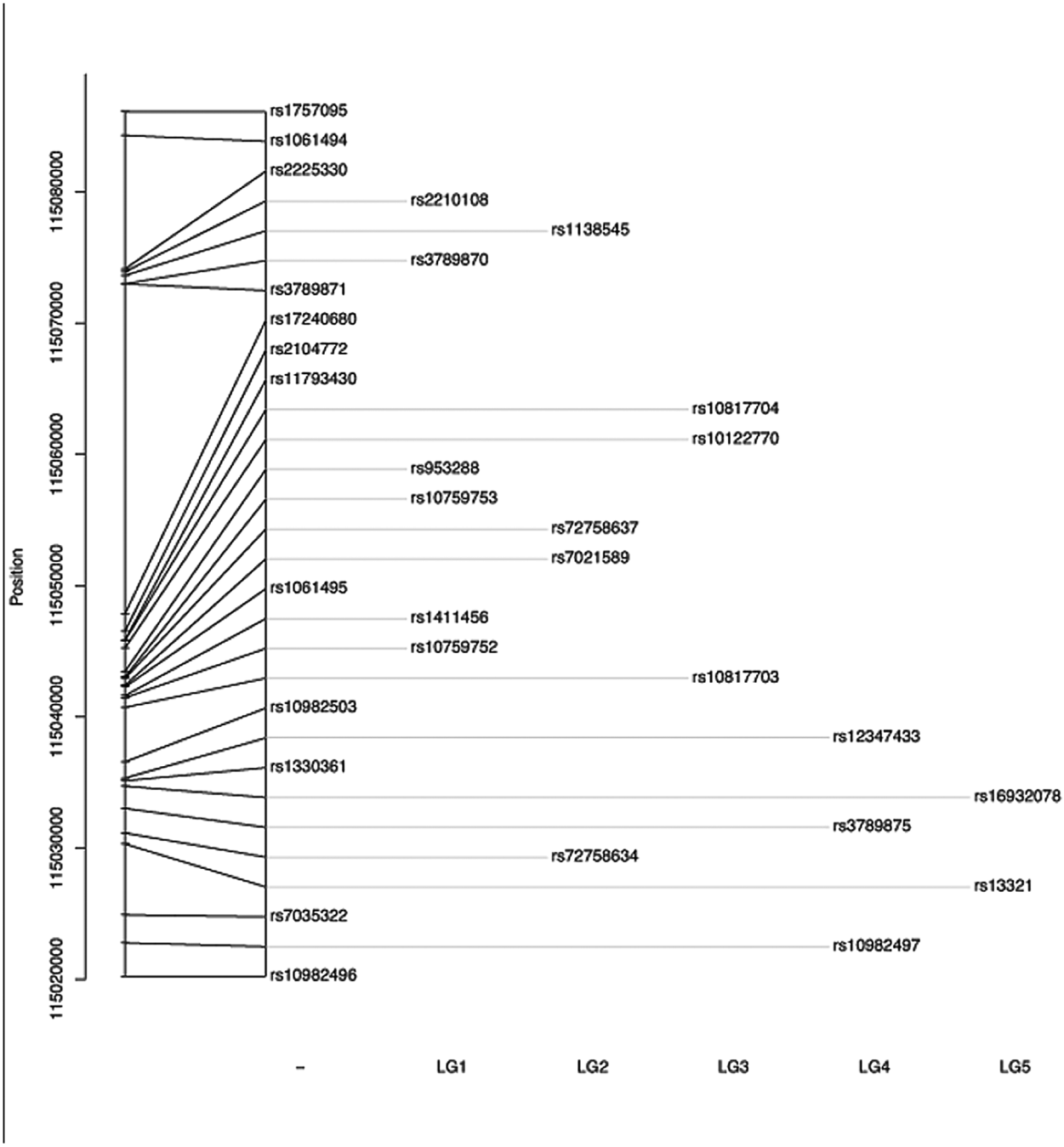

Figure 1. Schematic depiction of a 66 kbp TNC gene segment and the position of the 30 polymorphic SNPs used for an in depth statistical analysis. The gene segment contains groups of SNPs characterized by a complete or nearly complete LD and similar allele frequencies. Using an r^2-cutoff of 0.8, five such groups were inferred for grouping. Note that the six (after Bonferroni correction) significant SNPs are found in two overlapping LD groups: LG1: rs3789870 (intron 10; map.pos. 115072997) and rs10759753 (intron 17; map.pos. 115043015); Note that rs7035322 (intron 26; map.pos. 115024894) is associated with members of LG1 at $r^2 > 0.7$. LG2: rs1138545 (exon10; map.pos. 115073620), rs72758637 (intron 17; map.pos. 115042922), and rs7021589 (intron 17; map.pos. 115042388).

transcribes to the amino acid [R/H] of the amino acid sequence [R/H]VKASTEQAPEL which connects the FNIII5 domain with the FNIIIA1 domain of the TNC protein.

DISCUSSION

This study identified 15 SNPs within a region spanning from introns 9 to 26 (48 kbp) of the TNC gene, which are significantly associated with susceptibility for degenerative RCT. After Bonferroni correction, six SNPs remained significant. Compared to controls, RCT patients showed a significantly higher rate of homozygosity at rs72758637, rs7021589, and rs1138545, which are members of a linkage group with a pairwise linkage disequilibrium of $r^2 > 0.7$. RCT patients also showed a significantly higher rate of heterozygosity at

Table 2. Genotype and Minor Allele Frequency Distributions of the Six Sequence Variants Within the TNC Gene That Where (After Bonferroni Correction) Significantly Different Between Healthy Controls and Degenerative RCT Cases

TNC	Controls	Cases	*p*-Values
rs1138545			
n	75	155	
G/G	0.573	0.759	**0.001***
G/A	0.427	0.197	
A/A	0.000	0.044	
minor A	0.213	0.142	0.062
HWE	0.033	0.016	
rs3789870			
n	71	155	
C/C	0.662	0.372	**0.000***
C/T	0.268	0.493	
T/T	0.070	0.135	
minor T	0.204	0.381	**0.001***
HWE	0.140	0.495	
rs10759753			
n	72	155	
T/T	0.639	0.380	**0.001***
T/C	0.292	0.512	
C/C	0.069	0.108	
minor C	0.215	0.364	0.002
HWE	0.276	0.179	
rs72758637			
n	75	151	
G/G	0.547	0.762	**0.000***
G/C	0.453	0.198	
C/C	0.000	0.040	
minor C	0.227	0.139	0.024
HWE	0.018	0.049	
rs7021589			
n	76	155	
A/A	0.566	0.778	**0.000***
A/G	0.434	0.184	
G/G	0.000	0.038	
minor G	0.217	0.130	0.021
HWE	0.019	0.023	
rs7035322			
n	71	150	
G/G	0.662	0.407	**0.001***
G/T	0.268	0.513	
T/T	0.070	0.080	
minor T	0.204	0.337	0.005
HWE	0.132	0.068	

Deviations between the patient and control groups in genotype and allele frequencies were tested with chi-square tests. A boldface *p*-value with an asterisk "*" indicates significance after Bonferroni. Note that while the nominal significance level alpha is 0.05, the Bonferroni-corrected alpha level is 0.00167.

rs10759753, rs3789870, and rs7035322 and a higher *T*-allele frequency at rs3789870, which are members of an overlapping linkage group with a pairwise linkage disequilibrium of $r^2 > 0.85$. It is likely that further causal variants are located in these two LD groups as well as in other linkage groups that were significantly (but not after Bonferroni correction) associated with the RCT phenotype. The interdigitating pattern of the linkage groups is in accordance with HAPMAP. We have experimentally verified a sequence containing a boundary of two LD groups by Sanger sequencing, and confirmed the correctness of the alignment with DNA and RNA data of other hominids.

The probably most exciting finding of this study is the pairwise association of rs1138545 (exon10) and pQTlocus rs7021589 (intron 17) with the RCT phenotype. Missense variant rs1138545 transcribes to the amino acid [R/H] of the bridging sequence [R/H] VKASTEQAPEL which connects the FNIII5 domain with the FNIIIA1 domain of TNC. Altogether exon10 encodes for the FNIII5 domain as well as the [R/H] VKAST part of the bridge whereas exon11 encodes for the EQAPEL part of the bridge as well as the FNIIIA1 domain . The fifth TCN-FNIII domain not only binds to heparin, glypican, contactin, neurocan, perlican, lecticans, and periostin but also has a high affinity toward a large number of growth factors: Members of the platelet-derived growth factor (PDGF) family including the vascular endothelial growth factor (VEGF), fibroblast growth factor (FGF), transforming growth factor-b (TGF-b), and neuro-tropin families, the insulin-like growth factor-binding proteins (IGF-BPs), and others, with nanomolar binding affinities.[18,19,23] Therefore the fifth FNIII domain plays a pivotal role in effective ECM synthesis and ECM organization in injured tendon tissue.[19] The missense variant rs1138545 is located just one amino acid away from the FNIII5 domain, such that the ability of the FNIII5 domain to bind and sequester growth factors locally might be modified by the variant bridging sequence. A further step in clarifying the role of mutant and wild type TNC-FNIII 5/FNIIIA1 bridging sequence would therefore be to perform binding assays with various GFs.

Rs1138545 was in strong pairwise LD with the intronic pQTlocus rs7021589. In a recent study of genes that encode for altogether 778 proteins including signaling proteins, cytokines, growth factors, and kinases, rs7021589 was the most significant of 776 864 genetic markers associated with the abundance of proteins in plasma in a genome-wide association.[24] In particular the minor allele (G) variant of rs7021589 had a significant effect on TNC protein levels in plasma likely through its cis-regulatory function.[24] Cis SNPs like rs7021589 have gained growing attention as potential candidate genes for human diseases or common traits.

Of note our healthy controls showed a significantly higher minor allele (G) count at rs7021589 than RCT cases; the MAF of controls was 0.217 that of cases 0.130 ($p < 0.021$) (Suppl. Table S2).

Interestingly, some genetic variants that pose a risk for Achilles tendon tendinopathy/rupture in South African individuals with European Caucasian ancestry[8,9] locate in close vicinity to RCT-associated

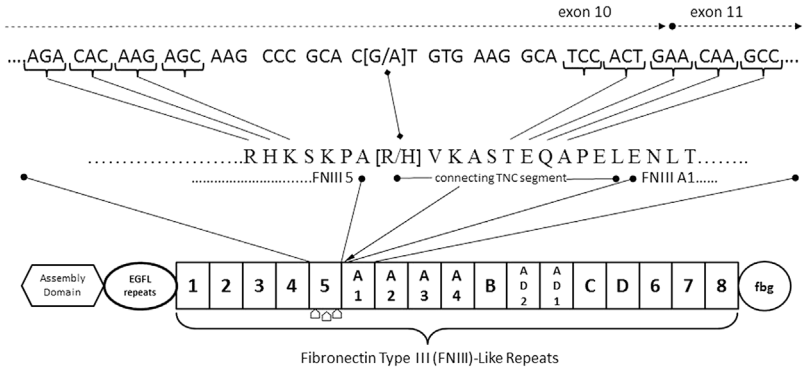

Figure 2. ⌂) FNIII5 promiscuously binds growth factors (23) PDGF-AA, PDGF-AB, PDGF-BB, PDGF-DD, VEGF-A165, VEGF-B, VEGF-C, PlGF-2, PlGF-3, β-FGF, FGF-4, FGF-6, FGF-7, FGF-8, FGF-10, FGF-17, FGF-18, TGF-β1, TGF-β2, (BMP)-2 IGF-BP3, IGF-BP5, neurotrophin-3 (NT-3), brain-derived growth factor (BDNF) as well as heparin and various integrins.

SNPs of our study. While no statistical association between rs145995427[8] a guanine-thymine dinucleotide repeat polymorphism (map.pos. 115 044 385) and RCT was detected in our samples ($\chi^2 = 8.9152$; $p = 0.2622$), markers rs2104772[9] (exon17) and rs13321[9] (exon24) were both significantly associated with RCT (Table 1).

The MAF frequency at rs2104772 was significantly higher in our controls (MAF 0.487) compared to RCT cases (MAF 0.376) ($p < 0.025$). Similarily in the Achilles tendinopathy study[9] the MAF at rs2104772 was significantly higher in South African controls (MAF 0.46) than in South African cases (MAF 0.37) ($p = 0.017$). Another coding SNP is rs13321 within the fibrinogen domain at the most distal part of the TNC molecule. This variant exchanges the uncharged amide derivative glutamine (major G-allele) to glutamic acid (minor C-allele), which may affect the folding of the protein and its interaction with other proteins.[19] Of note, rs13321 has recently been highlighted as part of a haplotype, which contains four SNPs in the COL27A1 gene and two SNPs in the TNC gene and associates with Achilles tendinopathy.[9] Two of our 155 RCT patients reported an Achilles tendon rupture in their medical history and two of our 76 control individuals reported episodes of Achilles tendinopathies.

We also looked at joint risk effects that combinations of risk genes might have. Combined risks have for instance been found for SNPs in the Col5A1 and MMP-3 gene (rs679620) in association with Achilles tendon injury,[11] and for genes of the inflammatory pathway that associate with RCT.[6] We tested SNPs in the Col5A1 and MMP-3 gene, but were unable to calculate any combined risk from our data. In addition, we found no significant difference between the allelic and genotypic distributions in the SNPs of MMP-1, MMP-2, MMP-9, and MMP-13 genes in our samples. Limitations: Since only a few polymorphisms in each gene have been assessed in this study, it is quite feasible that genetic variations at these genes could still be involved in susceptibility to RCT. Secondly, the sensitivity to detect full thickness RCTs with ultrasonography is considered to be investigator-specific. However, it has been verified by the first author as well as generally at his institution that ultrasonography and magnetic resonance tomography are equally powerful in detecting full thickness RCTs.[25,26] Thirdly, we have matched patients and control individuals for most RCT risk factors, but deliberately chose rather young patients with massive degenerative tears to compare with individuals in their senium, who still had intact rotator cuff tendons, to maximize the phenotypic contrast between cases and controls.

CONCLUSION

We found 15 SNPs in the TNC gene significantly associated with degenerative rotator cuff tendon tears. Some of these SNPs have been associated with Achilles tendon pathologies in former studies. Other phenotype related TNC SNPs open a new research track for the role of TNC growth factor binding domains in ECM synthesis and ECM organization in injured tendon tissue.

AUTHORS' CONTRIBUTIONS

RK: research concept, sample collection, patient evaluation, manuscript; JB: genotyping (Sequenom MassARRAY iPLEX system); CV: analysis and interpretation of data; GB: revising paper critically; MS: laboratory assistance; SM: sequencing, PCR. All authors have read and approved the final submitted manuscript.

REFERENCES

1. Yamamoto A, Takagishi K, Osawa T, et al. 2010. Prevalence and risk factors of a rotator cuff tear in the general population. J Shoulder Elbow Surg 19:116–120.
2. McFarland EG, Maffulli N, Buono AD, et al. 2013. Impingement is not impingement: the case for calling it "Rotator Cuff Disease". Muscles Ligaments Tendons J 11:196–200.
3. Baumgarten KM, Gerlach D, Galatz LM, et al. 2010. Cigarette smoking increases the risk for rotator cuff tears. Clin Orthop Relat Res 468:1534–1541.
4. Tashjian RZ, Farnham JM, Albright FS, et al. 2009. Evidence for an inherited predisposition contributing to the risk for rotator cuff disease. J Bone Joint Surg Am 91:1136–1142.
5. Peach CA, Zhang Y, Dunford JE, et al. 2007. Cuff tear arthropathy: evidence of functional variation in pyrophosphate metabolism genes. Clin Orthop Relat Res 462:67–72.

6. Motta Gda R, Amaral MV, Rezende E, et al. 2014. Evidence of genetic variations associated with rotator cuff disease. J Shoulder Elbow Surg 23:227–235.
7. Teerlink CC, Cannon-Albright LA, Tashjian RZ. 2015. Significant association of full-thickness rotator cuff tears and estrogen-related receptor-β (ESRRB). J Shoulder Elbow Surg 24:e31–e35.
8. Mokone GG, Gajjar M, September AV, et al. 2005. The guanine-thymine dinucleotide repeat polymorphism within the tenascin-C gene is associated with achilles tendon injuries. Am J Sports Med 33:1016–1021.
9. Saunders CJ, van der Merwe L, Posthumus M, et al. 2013. Investigation of variants within the COL27A1 and TNC genes and Achilles tendinopathy in two populations. J Orthop Res 31:632–637.
10. September AV, Cook J, Handley CJ, et al. 2009. Variants within the COL5A1 gene are associated with Achilles tendinopathy in two populations. Br J Sports Med 43:357–365.
11. Raleigh SM, van der Merwe L, Ribbans WJ, et al. 2009. Variants within the MMP3 gene are associated with Achilles tendinopathy: possible interaction with the COL5A1 gene. Br J Sports Med 43:514–520.
12. Altinisik J, Meric G, Erduran M, et al. 2015. The BstUI and DpnII variants of the COL5A1 gene are associated with tennis elbow. Am J Sports Med 43:1784–1789.
13. Godoy-Santos A, Cunha MV, Ortiz RT, et al. 2013. MMP-1 promoter polymorphism is associated with primary tendinopathy of the posterior tibial tendon. J Orthop Res 31:1103–1107.
14. Jelinsky SA, Rodeo SA, Li J, et al. 2011. Regulation of gene expression in human tendinopathy. BMC Musculoskeletal Disord 12:86.
15. Lakemeier S, Braun J, Efe T, et al. 2011. Expression of matrix metalloproteinases 1, 3, and 9 in differing extents of tendon retraction in the torn rotator cuff. Knee Surg Sports Traumatol Arthrosc 19:1760–1765.
16. Lo IK, Marchuk LL, Hollinshead R, et al. 2004. Matrix metalloproteinase and tissue inhibitor of matrix metalloproteinase mRNA levels are specifically altered in torn rotator cuff tendons. Am J Sports Med 32:1223–1229.
17. Flück M, Mund SI, Schittny JC, et al. 2008. Mechano-regulated tenascin-C orchestrates muscle repair. Proc Natl Acad Sci USA 105:13662–13667.
18. Giblin SP, Midwood KS. 2015. Tenascin-C: form versus function. Cell Adh Migr 9:48–82.
19. Udalova IA, Ruhmann M, Thomson SJ, et al. 2011. Expression and immune function of tenascin-C. Crit Rev Immunol 31:115–145.
20. Longo UG, Franceschi F, Ruzzini L, et al. 2008. Histopathology of the supraspinatus tendon in rotator cuff tears. Am J Sports Med 36:533–538.
21. NCBI: National Center for Biotechn. Inform., US National Library of Medicine 8600 Rockville Pike, Bethesda MD, 20894 USA, http://www.ncbi.nlm.nih.gov/
22. Buggs RJ, Chamala S, Wu W, et al. 2010. Characterization of duplicate gene evolution in the recent natural allopolyploid tragopogon miscellus by next-generation sequencing and sequenom iPLEX MassARRAY genotyping. Mol Ecol 1:132–146.
23. De Laporte L, Rice JJ, Tortelli F, et al. 2013. Tenascin C promiscuously binds growth factors via its fifth fibronectin type III-like domain. PLoS ONE 8:e62076.
24. Lourdusamy A, Newhouse S, Lunnon K, et al. 2012. Identification of cis-regulatory variation influencing protein abundance levels in human plasma. Hum Mol Genet 21: 3719–3726.
25. Kluger R, Bock P, Mittlböck M, et al. 2011. Long-term survivorship of rotator cuff repairs using ultrasound and magnetic resonance imaging analysis. Am J Sports Med 39:2071–2081.
26. Kluger R, Mayrhofer R, Kröner A, et al. 2003. Sonographic versus magnetic resonance arthrographic evaluation of full-thickness rotator cuff tears in millimeters. J Shoulder Elbow Surg 12:110–116.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article.