2021

Identification of a novel genetic marker for risk of degenerative rotator cuff disease surgery in the UK Biobank

Elizabeth L Yanik
Washington University School of Medicine in St. Louis

Jay D Keener
Washington University School of Medicine in St. Louis

Shiow J Lin
Washington University School of Medicine in St. Louis

Graham A Colditz
Washington University School of Medicine in St. Louis

Rick W Wright
Washington University School of Medicine in St. Louis

See next page for additional authors

Identification of a Novel Genetic Marker for Risk of Degenerative Rotator Cuff Disease Surgery in the UK Biobank

Elizabeth L. Yanik, PhD, ScM, Jay D. Keener, MD, Shiow J. Lin, MS, Graham A. Colditz, DrPH, MD, Rick W. Wright, MD, Bradley A. Evanoff, MD, MPH, Nitin B. Jain, MD, MSPH, and Nancy L. Saccone, PhD

Investigation performed at Washington University School of Medicine, St. Louis, Missouri

Background: While evidence indicates that familial predisposition influences the risk of developing degenerative rotator cuff disease (RCD), knowledge of specific genetic markers is limited. We conducted a genome-wide association study of RCD surgery using the UK Biobank, a prospective cohort of 500,000 people (40 to 69 years of age at enrollment) with genotype data.

Methods: Cases with surgery for degenerative RCD were identified using linked hospital records. The cases were defined as an International Classification of Diseases, Tenth Revision (ICD-10) code of M75.1 determined by a trauma/orthopaedic specialist and surgery consistent with RCD treatment. Cases were excluded if a diagnosis of traumatic injury had been made during the same hospital visit. For each case, up to 5 controls matched by age, sex, and follow-up time were chosen from the UK Biobank. Analyses were limited to European-ancestry individuals who were not third-degree or closer relations. We used logistic regression to test for genetic association of 674,405 typed and >10 million imputed markers, after adjusting for age, sex, population principal components, and follow-up.

Results: We identified 2,917 RCD surgery cases and 14,158 matched controls. We observed 1 genome-wide significant signal ($p < 5 \times 10^{-8}$) for a novel locus tagged by rs2237352 in the *CREB5* gene on chromosome 7 (odds ratio [OR] = 1.17, 95% confidence interval [CI] = 1.11 to 1.24). The single-nucleotide polymorphism (SNP) rs2237352 was imputed with a high degree of confidence (info score = 0.9847) and is common, with a minor allele frequency of 47%. After expanding the control sample to include additional unmatched non-cases, rs2237352 and another SNP in the *CREB5* gene, rs12700903, were genome-wide significant. We did not detect genome-wide significant signals at loci associated with RCD in previous studies.

Conclusions: We identified a novel association between a variant in the *CREB5* gene and RCD surgery. Validation of this finding in studies with imaging data to confirm diagnoses will be an important next step.

Clinical Relevance: Identification of genetic RCD susceptibility markers can guide understanding of biological processes in rotator cuff degeneration and help inform disease risk in the clinical setting.

Level of Evidence: Prognostic Level III. See Instructions for Authors for a complete description of levels of evidence.

Rotator cuff disease (RCD) is the most common cause of shoulder disability[1-4], but studies identifying risk factors for symptomatic RCD have been scarce. Prevention strategies aimed at high-risk groups could dramatically impact

Accumulating evidence indicates that familial predisposition influences the risk of developing degenerative RCD[7-10]. However, knowledge of specific RCD genetic markers is limited. Investigators have evaluated candidate genes and discovered

damage[11-13]. The first genome-wide association study (GWAS) for RCD, of which we are aware, detected 2 associated single-nucleotide polymorphisms (SNPs) involved in apoptosis[14], but this study of <350 patients with RCD had limited statistical power. The authors of a second GWAS were unable to replicate the associations found in the first, and they identified a new SNP associated with RCD[15]. While this study was much larger (8,357 rotator cuff injuries), the definition for RCD was nonspecific, with use of International Classification of Diseases (ICD) codes that might capture shoulder pain from other diseases[15].

Uncertainty still exists about which genetic markers have true associations with degenerative RCD. The UK Biobank population of half a million people with genotype data provided a unique opportunity for a large GWAS with carefully defined RCD cases and controls to identify additional genetic markers and further evaluate the replicability of previous findings.

Materials and Methods

Our study population was derived from the UK Biobank, a population-based prospective cohort of approximately 500,000 U.K. residents[16,17]. Participants 40 to 69 years of age were recruited nationwide from 2006 to 2010 through invitations mailed to people registered with the National Health Service (NHS)[18]. At enrollment, participants gave informed consent and whole blood samples were collected. NHS hospital records were linked to the UK Biobank, providing information on inpatient diagnoses and procedures from 2006 to 2017. Diagnoses were coded using the ICD, Tenth Revision (ICD-10). Procedures were coded using the Office of Population Censuses and Surveys Classification, Fourth Revision (OPCS-4). DNA was extracted from whole blood samples and genotyped using the UK Biobank Axiom Array, which includes 812,428 SNPs and insertion-deletion markers[17,19]. An additional 73 million markers were imputed using a reference haplotype panel to predict genetic markers not directly assayed. We obtained de-identified data from the UK Biobank (Project Number 27034) on 488,292 UK Biobank participants with available genotype results. Of these, 968 participants were excluded because of poor-quality results indicated by either extreme heterozygosity or missingness[17].

We selected cases and controls from the remaining group of 487,324 people. As >94% of participants self-reported European/White ancestry, the case-control population was limited to that group to reduce population stratification. Numerous quality-control measures were undertaken to confirm the data quality of the selected case-control sample[20,21]. People were excluded if there 1 person from each related pair was excluded. When a pair contained a case (as defined below) and a non-case, the non-case was preferentially excluded. Otherwise, 1 person was excluded at random. The remaining subcohort formed the population from which cases and controls were selected.

Cases of degenerative RCD were identified based on a primary or secondary ICD-10 diagnosis code of M75.1. To reduce the risk of misdiagnosis, RCD cases were included only if the diagnosis had been made by a trauma/orthopaedic specialist and the patient underwent a surgical procedure consistent with RCD treatment. As genetic factors are considered important primarily for degenerative RCD, cases were excluded if they occurred concurrently with a traumatic injury diagnosis (e.g., shoulder dislocation).

Up to 5 controls were randomly selected for each case through incidence-density sampling, in which controls were randomly chosen from among individuals without a prior diagnosis of degenerative RCD at the time of the case diagnosis. Incidence-density sampling ensures a representative sample population with comparable follow-up for cases and controls[23,24]. Follow-up was defined as the time since enrollment in the UK Biobank. Controls were matched to cases on the basis of age and reported sex.

The NHS Research Ethics Committee approved the UK Biobank. The Washington University institutional review board determined this study to be exempt from oversight.

We included genetic variants that did not diverge from Hardy-Weinberg equilibrium with a p value of $<1 \times 10^{-6}$ and that had a minor allele frequency (MAF) of >0.004285. The MAF cutoff was based on the formula $25/(2 \times \text{number of cases})$ so ≥25 minor alleles would be expected in cases under the null hypothesis. Only imputed variants with an info score of ≥0.3 were included to remove variants with low-confidence imputed values.

Association tests used logistic regression to model genotype dosage effects on RCD, with covariates for age, sex, follow-up time, and the first 10 population principal components[22]. Quantile-quantile (Q-Q) plots were graphed, and a genomic inflation factor was calculated to check for bias. We required a standard[20,25] genome-wide significance threshold of 5×10^{-8}. For regions harboring GWAS significant signals, we performed an adjusted analysis using the lead SNP as a covariate to detect additional independent signals. We also specifically examined genetic markers identified as significantly associated with RCD in prior literature[11-15,26-30]. For these markers, a Bonferroni-adjusted p value of <0.05 and an odds ratio (OR) indicating an association in the same direction as the original

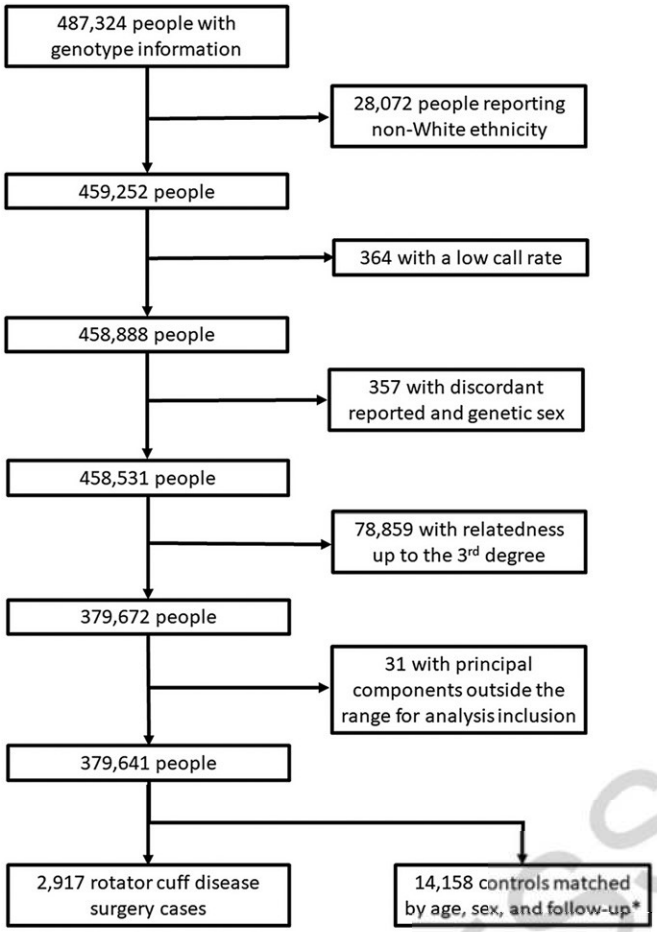


Fig. 1

Flowchart of exclusions made prior to selection of RCD surgery cases and controls from the UK Biobank.

more precisely adjust for matching criteria. Third, associations were estimated in the subgroup of patients ≤60 years old at the time of the RCD surgery (and corresponding controls), as we hypothesized that genetic predisposition could lead to earlier-onset disease.

The research was conducted using the UK Biobank resource (http://www.ukbiobank.ac.uk/). Returned results, including GWAS association results for all typed and imputed variants tests, can be found under application 27034.

were made after a median of 5 years (interquartile range [IQR] = 3 to 6 years, range = 0 to 10 years) of follow-up (i.e., after enrollment in the UK Biobank). The median age at diagnosis was 65 years old (IQR = 59 to 69 years, range = 41 to 78 years), and 48.5% of the cases were women (Table I). Cases were matched with a 1:5 ratio to 14,158 unique controls on the basis of follow-up, age, and sex. For conditional logistic regression to represent incidence-density sampling, some individuals could be selected as controls multiple times, resulting in 14,547 controls.

Initially, >77 million typed and imputed variants were available for analyses. Of these, >66 million were removed because of an MAF of <0.004285 and 50,998 were removed because of a Hardy-Weinberg exact test p value of $<1 \times 10^{-6}$. There remained 674,405 typed and 10,140,917 imputed variants included in the analyses.

The Q-Q plot and genomic inflation factor of 1.02 provided no evidence of bias after accounting for matching factors and the first 10 principal components (Fig. 2).

We observed 1 novel genome-wide significant signal ($p = 4.04 \times 10^{-8}$) at SNP rs2237352 in the *CREB5* gene on chromosome 7 (OR = 1.17, 95% confidence interval [CI] = 1.11 to 1.24; Table II, Fig. 3). SNP rs2237352 was imputed with a high degree of confidence (info score = 0.9847) and is a common variant (MAF = 46.8%). The second strongest signal was for imputed SNP rs12700903 in the *CREB5* gene (OR = 1.17, 95% CI = 1.11 to 1.24, $p = 5.63 \times 10^{-8}$), which is in strong linkage disequilibrium with rs2237352 ($r^2 = 0.98$). Thus, both SNPs represent the same statistical signal. The most significant directly assayed SNP in the *CREB5* gene was rs66539057, but the association was not significant at a genome-wide level (OR = 1.16, 95% CI = 1.09 to 1.23, $p = 1.29 \times 10^{-6}$). Figure 4 shows a detailed view of the associated

TABLE I Characteristics of Population of RCD Surgery Cases and Selected Controls from the UK Biobank

	No (%)*	
	Cases	Controls
Total	2,917	14,158
Age *(yr)*		
At enrollment	61 (55, 65)	61 (55, 65)
At diagnosis	65 (59, 69)	Not applic.
Sex		
Male	1,503 (51.5%)	7,284 (51.4%)

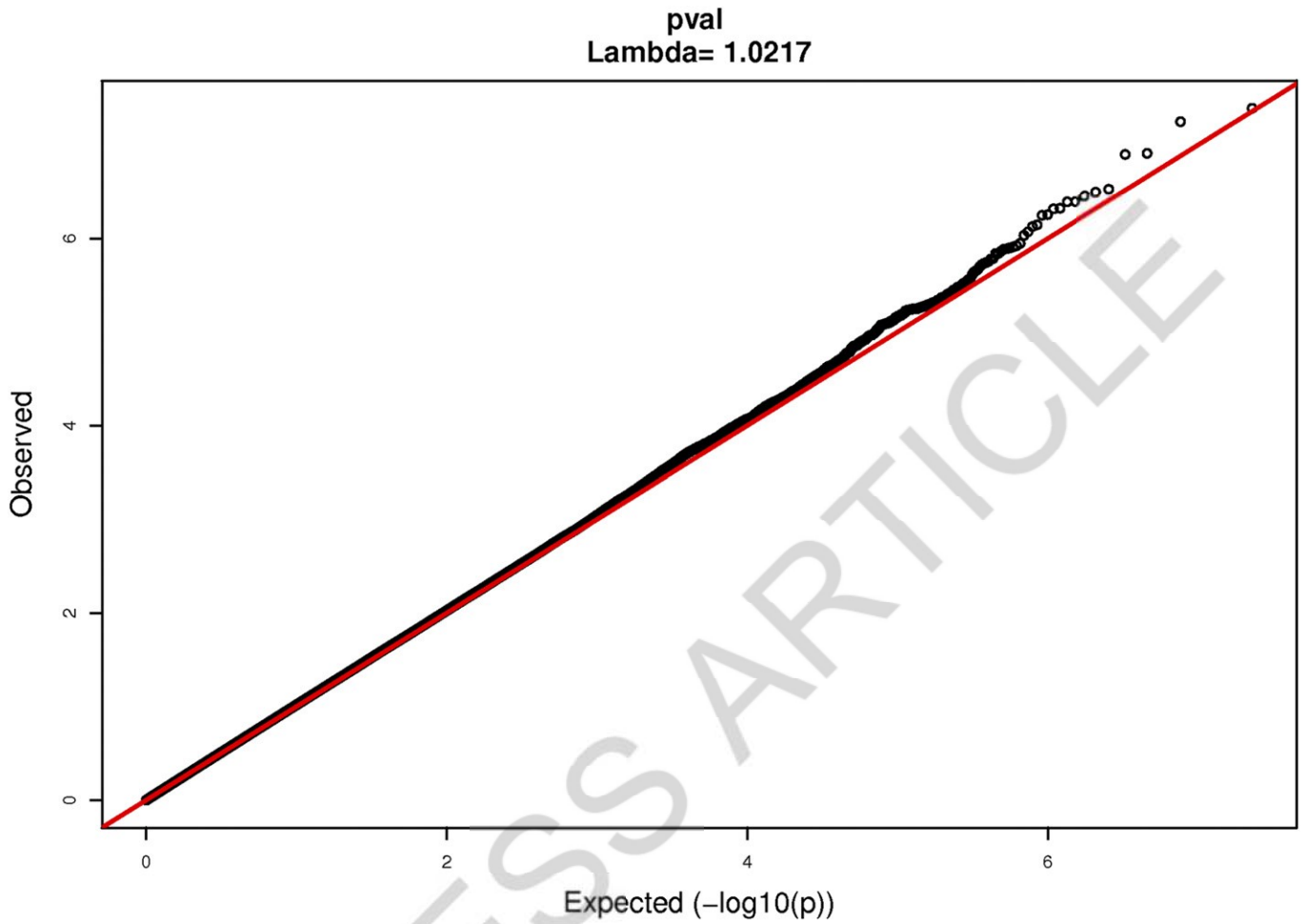


Fig. 2

Q-Q plot comparing the observed p values (pval) with the expected distribution of p values for each association of a genetic variant with RCD surgery in the UK Biobank. Substantial, systematic divergence of the distribution of data points from the red diagonal line would indicate bias. Lambda represents the genomic inflation factor calculated by dividing the median observed test statistics by the median expected test statistic. A genomic inflation factor of 1 indicates no bias.

region, with the lead SNP rs2237352 having strong to moderate linkage disequilibrium with additional SNPs in the region[33]. The strongest signal for an SNP having modest linkage disequilibrium with rs2237352 was for rs4722837 (OR = 0.86, $p = 1.26 \times 10^{-7}$, $r^2 = 0.38$). Results were similar in conditional logistic regression (rs2237352: OR = 1.17, 95% CI = 1.10 to 1.24; rs12700903: OR = 1.17, 95% CI = 1.10 to 1.24) (Table II).

After analyses with adjustment for the rs2237352 genostudy[27]. After Bonferroni adjustment for 29 replication attempts, neither SNP remained significant.

After we expanded the control group to include the larger, unmatched cohort of non-cases (n = 375,560), rs2237352 remained genome-wide significant ($p = 2.29 \times 10^{-8}$; Table II). Additionally, the association of rs12700903 with surgery for degenerative RCD became genome-wide significant ($p = 3.69 \times 10^{-8}$).

There were 735 patients with RCD surgery who were ≤60

TABLE II SNP Associations with RCD Surgery in the UK Biobank*

						Cases and Matched Controls					Cases and Expanded Controls		
							Standard Logistic Regression		Conditional Logistic Regression				
Chromosome	SNP	Gene	A1	A2	Info Score	A1 Frequency	OR (95% CI)	P Value	OR (95% CI)	P Value	A1 Frequency	OR (95% CI)	P Value
7	rs2237352	*CREB5*	C	T	0.9847	0.468	1.17 (1.11, 1.24)	4.04×10^{-8}	1.17 (1.10, 1.24)	8.75×10^{-8}	0.465	1.16 (1.10, 1.22)	2.29×10^{-8}
7	rs12700903	*CREB5*	G	C	0.9744	0.469	1.17 (1.11, 1.24)	5.63×10^{-8}	1.17 (1.10, 1.24)	1.14×10^{-7}	0.466	1.16 (1.10, 1.22)	3.69×10^{-8}
7	rs4722837	*CREB5*	G	A	0.9900	0.454	0.86 (0.81, 0.91)	1.26×10^{-7}	0.86 (0.81, 0.91)	1.72×10^{-7}	0.456	0.87 (0.83, 0.92)	2.72×10^{-7}
7	rs66539057	*CREB5*	T	C	1	0.308	1.16 (1.09, 1.23)	1.29×10^{-6}	1.16 (1.09, 1.23)	1.35×10^{-6}	0.307	1.14 (1.08, 1.20)	5.08×10^{-6}

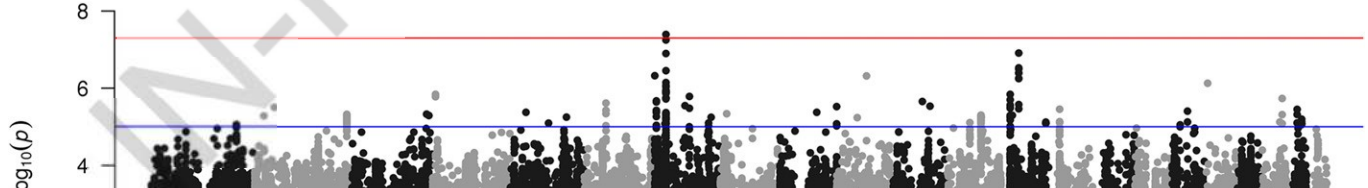
*All regression analyses modeled the effect of genotype dosage on RCD. Standard logistic regression included covariates for age, sex, follow-up time, and the first 10 population principal components (2,917 cases and 14,158 unique controls). Conditional logistic regression incorporated individuals who could be selected multiple times as controls and was conditioned on matched sets of controls with each case to account for all matched covariates (2,917 cases and 14,547 controls). Logistic regression with expanded controls used as controls all non-cases in the UK Biobank that met quality-control standards and included covariates for age, sex, follow-up time, and the first 10 population principal components (2,917 cases and 375,560 unique controls). A1 = coded allele in regression, and A2 = reference allele in regression.

association with a common SNP in the *CREB5* gene. SNP associations reported in prior RCD studies (Table III) were not replicated in our population after multiple test correction. These findings highlight the need for additional large GWASs of degenerative RCD focused on carefully defining case and control status and identifying replicable results.

We identified a novel association between the SNP rs2237352 and surgery for degenerative RCD. The SNP rs12700903 was also associated with surgery for degenerative RCD after expansion of our control group, and it represents the same signal as SNP rs2237352. Both SNPs are located in the *CREB5* gene, which encodes a protein that is part of the cAMP response element-binding protein family[34]. *CREB5* is a transcription factor involved in cell growth, proliferation, and differentiation[35,36]. *CREB5* expression has been associated with plasma interleukin-6 levels and may influence inflammatory response genes[37]. As CREB family proteins influence expression of other genes[38,39], there may be numerous genetic mutations that could influence the same biological pathways. If further research confirms this association, one would expect the genetic risk for RCD to be highly polygenic as is common for most complex traits. Differential *CREB5* expression has also been specifically documented in fibroblasts[40], lending further evidence that mutations in this gene could be of importance for tendon injury and repair.

After adjustment for rs2237352 in our models, no additional signals were detected, which is consistent with this region harboring 1 primary locus associated with degenerative RCD. *CREB5* SNPs in weaker linkage disequilibrium with the top signal did not provide GWAS-significant evidence for another distinct signal in the region. However, this locus could represent an accumulation of weak effects from linked variants that influence degenerative RCD[41]. Notably, as rs2237352 is an intron variant, it may be indicative of an unknown genetic determinant with which it co-segregates.

Most prior RCD genetic epidemiology studies have been candidate gene studies. To our knowledge, 2 GWAS RCD studies in independent populations have been conducted[10,14,15]. Candidate gene studies focus on specific genes with known function potentially related to rotator cuff degeneration, whereas GWAS studies take an agnostic approach to testing association with large portions of the genome. Of the 29 SNPs



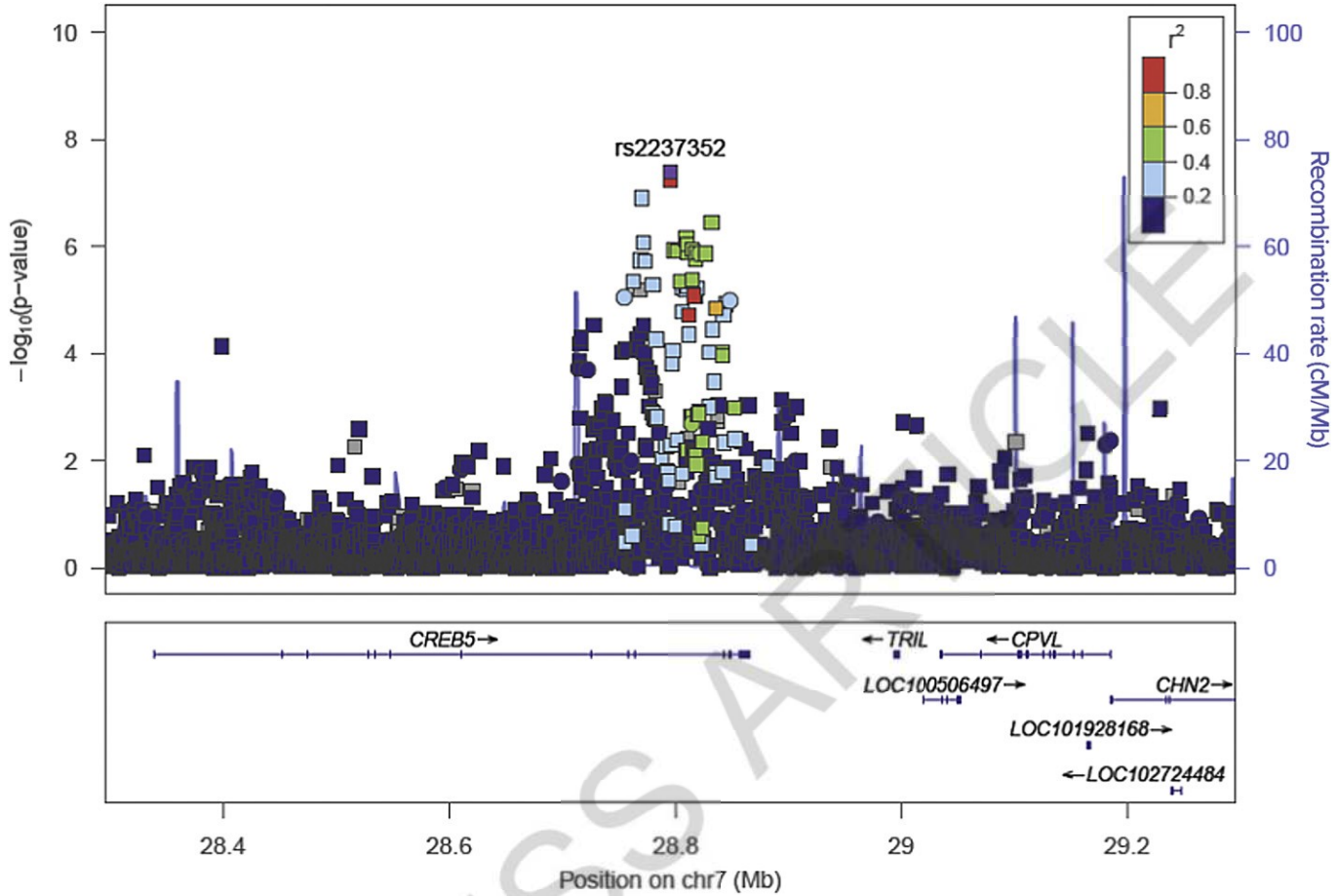


Fig. 4

The association with RCD of SNPs on chromosome 7. This LocusZoom plot shows the association (left y axis; log_{10}-transformed p values) with RCD. Genotyped SNPs are depicted by circles, and imputed SNPs are depicted by squares. Shading of the points represent the linkage disequilibrium (r^2, based on the 1,000 Genomes Project Europeans) between each SNP and the top SNP, indicated by purple shading. Grey points in the plot represent the lack of linkage disequilibrium information between the index SNP (rs2237352) and the plotted SNP.

from prior studies (Table III) that we could evaluate in the UK Biobank, only 1 (rs820218) demonstrated an association in the same direction as it did in the prior study[14], with a nominal uncorrected p value of < 0.05, while none reached genome-wide significance. SNP rs820218 is located in the SAP30BP gene, which encodes a transcriptional regulator protein involved in cell death and apoptosis[14,42]. Numerous studies have shown increased tendon cell apoptosis related to rotator cuff tearing[43-45]. However, as only 1 SNP out of 29 demonstrated a fying large portions of the genetic contributions to other complex diseases and can be more susceptible to publication bias[46-48].

One other genetic RCD study has recently been conducted utilizing the UK Biobank[49], but it used a less specific case definition that appeared to include primary care diagnoses without treatment information. The associated SNPs in our study were not identified as having genome-wide significant signals in that study, which instead revealed 3 other associated

TABLE III Replication Testing in the UK Biobank of SNPs Associated with RCD in the Prior Literature*

			Prior Study Information				Results from UK Biobank	
Chromosome	SNP	Gene	First Author	No. of Cases	Type of Study*	A1	OR (95% CI)	P Value
1	rs4654760	*ALPL*	Peach[26]	22	CG	T	1.00 (0.890, 1.12)	0.958
5	rs3045	*ANKH*	Peach[26]	22	CG	C	1.01 (0.916, 1.10)	0.910
5	rs1011814	*FGF10*	Motta[11]	203	CG	T	1.03 (0.967, 1.09)	0.393
5	rs11750845	*FGF10*	Motta[11]	203	CG	C	1.01 (0.956, 1.07)	0.687
6	rs12527089	*SASH1*	Tashjian[14]	311	GWAS	T	0.959 (0.836, 1.10)	0.545
8	rs13317	*FGFR1*	Motta[11]	203	CG	C	0.961 (0.899, 1.03)	0.237
8	rs1800972	*DEFB1*	Motta[11]	203	CG	C	1.02 (0.956, 1.09)	0.512
9	rs1590	*TGFBR1*	Figueiredo[27]	211	CG	G	1.00 (0.940, 1.07)	0.937
9	rs10759753	*TNC*	Kluger[13]	155	CG	G	0.989 (0.927, 1.05)	0.725
9	rs1138545	*TNC*	Kluger[13,28]	155[13], 120[28]	CG	T	1.06 (0.982, 1.15)	0.136
9	rs2104772	*TNC*	Kluger[28], Figueiredo[27]	120[28], 211[27]	CG	A	1.05 (0.991, 1.11)	0.100
9	rs3789870	*TNC*	Kluger[13]	155	CG	A	0.994 (0.932, 1.06)	0.865
9	rs7021589	*TNC*	Kluger[13]	155	CG	C	1.06 (0.984, 1.15)	0.122
9	rs7035322	*TNC*	Kluger[13]	155	CG	A	1.02 (0.959, 1.09)	0.500
9	rs72758637	*TNC*	Kluger[13]	155	CG	G	1.07 (0.994, 1.16)	0.070
9	rs3196378	*Col5A1*	Figueiredo[27]	211	CG	C	1.02 (0.967, 1.08)	0.424
11	rs12574452	*FGF3*	Motta[11]	203	CG	A	1.04 (0.982, 1.11)	0.174
11	rs1799750	*MMP1*	Assunção[55]	64	CG	†	†	†
11	rs3025058	*MMP3*	Assunção[55]	64	CG	†	†	†
11	rs679620	*MMP3*	Figueiredo[27]	211	CG	C	1.01 (0.957, 1.07)	0.658
14	rs10132091	*ESRRB*	Bonato[12]	49	CG	C	0.991 (0.936, 1.05)	0.760
14	rs1676303	*ESRRB*	Motta[11], Bonato[12]	203[11], 16[12]	CG	C	1.00 (0.915, 1.10)	0.964
14	rs17583842	*ESRRB*	Teerlink[29], Tashjian[30]	175[29], 30[30]	CG	C	0.940 (0.879, 1.01)	0.072
14	rs4903399	*ESRRB*	Motta[11], Bonato[12]	203[11], 49[12]	CG	T	1.03 (0.959, 1.11)	0.418
16	rs2285053	*MMP2*	Figueiredo[27]	211	CG	T	1.08 (0.985, 1.19)	0.101
16	rs71404070		Roos[15]	8,357	GWAS	A	0.993 (0.866, 1.14)	0.919
17	rs820218	*SAP30BP*	Tashjian[14]	311	GWAS	A	0.934 (0.879, 0.991)	0.025
17	rs2277698	*TIMP2*	Figueiredo[27]	211	CG	T	0.870 (0.795, 0.952)	0.002
19	rs1800470	*TFGB1*	Figueiredo[27]	211	CG	G	1.02 (0.961, 1.08)	0.532
19	rs1800469	*TFGB1*	Figueiredo[27]	211	CG	A	1.01 (0.951, 1.08)	0.710
20	rs17576	*MMP9*	Figueiredo[27]	211	CG	G	1.01 (0.949, 1.07)	0.821

*CG = candidate gene study, and GWAS = genome-wide association study. †Not measured or did not meet quality-control filtering criteria in the UK Biobank.

There are important limitations to the current study. First, outcomes and thus could not evaluate how genetics influenced

than those in other large RCD GWASs[15,49], and we did not appear to be substantially undercapturing RCD surgery cases based on rates in the literature[52]. A large number of genetic markers were available for examination, including typed markers and markers imputed with a high degree of confidence. This strength of the UK Biobank will improve in the future as plans are in place for whole genome sequencing of the population[53].

A more comprehensive understanding of genetic susceptibility to degenerative RCD could aid treatment and prevention in several ways. First, someone with a genetic predisposition for RCD could derive greater benefits from changing modifiable risk factors such as smoking or occupational burdens. Second, a predisposition for cuff degeneration may also indicate an impaired ability of the cuff to heal following surgical repair, which could influence cuff-repair indications. Third, genetic susceptibility markers may point to key biological pathways in cuff degeneration that could direct future basic-science research, leading to novel therapeutics.

We identified a novel SNP in the *CREB5* gene associated with surgery for degenerative RCD in a general population sample of the U.K. Replication of this finding will be important in the future. Future examination of the genetic determinants of other chronic tendon disorders, including investigation of commonalities across such disorders, would be useful. The extensive information available in the UK Biobank could allow future evaluation of risk models incorporating genetics, non-genetic characteristics, and gene-environment interactions[54]. Identification of potentially important genetic markers in our study and others can allow a more focused study of these markers in smaller cohorts with more detailed clinical information, including investigations of how genetic factors may influence RCD progression and outcomes after surgical treatment. ■

Elizabeth L. Yanik, PhD, ScM[1]
Jay D. Keener, MD[1]
Shiow J. Lin, MS[1]
Graham A. Colditz, DrPH, MD[1]
Rick W. Wright, MD[1]
Bradley A. Evanoff, MD, MPH[1]
Nitin B. Jain, MD, MSPH[2]
Nancy L. Saccone, PhD[1]

[1]Departments of Orthopaedic Surgery (E.L.Y., J.D.K., and R.W.W.), Surgery (E.L.Y. and G.A.C.), Genetics (S.J.L. and N.L.S.), and General Medical Studies (B.A.E.), Washington University School of Medicine, St. Louis, Missouri

[2]Department of Physical Medicine and Rehabilitation, University of Texas Southwestern, Dallas, Texas

Email address for E.L. Yanik: yanike@wustl.edu

ORCID iD for E.L. Yanik: 0000-0002-5835-0201
ORCID iD for J.D. Keener: 0000-0002-1665-4346
ORCID iD for S.J. Lin: 0000-0002-4292-1751
ORCID iD for G.A. Colditz: 0000-0002-7307-0291
ORCID iD for R.W. Wright: 0000-0003-4018-7132
ORCID iD for B.A. Evanoff: 0000-0003-0085-333X
ORCID iD for N.B. Jain: 0000-0002-4362-8522
ORCID iD for N.L. Saccone: 0000-0002-4710-3926

References

1. Yamamoto A, Takagishi K, Osawa T, Yanagawa T, Nakajima D, Shitara H, Kobayashi T. Prevalence and risk factors of a rotator cuff tear in the general population. J Shoulder Elbow Surg. 2010 Jan;19(1):116-20.
2. Liem D, Buschmann VE, Schmidt C, Gosheger G, Vogler T, Schulte TL, Balke M. The prevalence of rotator cuff tears: is the contralateral shoulder at risk? Am J Sports Med. 2014 Apr;42(4):826-30. Epub 2014 Feb 5.
3. Sher JS, Uribe JW, Posada A, Murphy BJ, Zlatkin MB. Abnormal findings on magnetic resonance images of asymptomatic shoulders. J Bone Joint Surg Am. 1995 Jan;77(1):10-5.
4. Tempelhof S, Rupp S, Seil R. Age-related prevalence of rotator cuff tears in asymptomatic shoulders. J Shoulder Elbow Surg. 1999 Jul-Aug;8(4):296-9.
5. Mather RC 3rd, Koenig L, Acevedo D, Dall TM, Gallo P, Romeo A, Tongue J, Williams G Jr. The societal and economic value of rotator cuff repair. J Bone Joint Surg Am. 2013 Nov 20;95(22):1993-2000.
6. Colvin AC, Egorova N, Harrison AK, Moskowitz A, Flatow EL. National trends in rotator cuff repair. J Bone Joint Surg Am. 2012 Feb 1;94(3):227-33.
7. Tashjian RZ. Epidemiology, natural history, and indications for treatment of
12. Bonato LL, Quinelato V, Pinheiro AdaR, Amaral MV, de Souza FN, Lobo JC, Aguiar DP, Augusto LM, Vieira AR, Salles JI, Cossich VR, Guimarães JA, de Gouvêa CV, Granjeiro JM, Casado PL. ESRRB polymorphisms are associated with comorbidity of temporomandibular disorders and rotator cuff disease. Int J Oral Maxillofac Surg. 2016 Mar;45(3):323-31. Epub 2015 Nov 14.
13. Kluger R, Burgstaller J, Vogl C, Brem G, Skultety M, Mueller S. Candidate gene approach identifies six SNPs in tenascin-C (TNC) associated with degenerative rotator cuff tears. J Orthop Res. 2017 Apr;35(4):894-901. Epub 2016 Jun 19.
14. Tashjian RZ, Granger EK, Farnham JM, Cannon-Albright LA, Teerlink CC. Genome-wide association study for rotator cuff tears identifies two significant single-nucleotide polymorphisms. J Shoulder Elbow Surg. 2016 Feb;25(2):174-9. Epub 2015 Sep 5.
15. Roos TR, Roos AK, Avins AL, Ahmed MA, Kleimeyer JP, Fredericson M, Ioannidis JPA, Dragoo JL, Kim SK. Genome-wide association study identifies a locus associated with rotator cuff injury. PLoS One. 2017 Dec 11;12(12):e0189317.
16. Collins R. What makes UK Biobank special? Lancet. 2012 Mar 31;379(9822):1173-4.
17. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D,

21. Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, Derks EM. A tutorial on conducting genome-wide association studies: quality control and statistical analysis. Int J Methods Psychiatr Res. 2018 Jun;27(2):e1608. Epub 2018 Feb 27.

22. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006 Aug;38(8):904-9. Epub 2006 Jul 23.

23. Lubin JH, Gail MH. Biased selection of controls for case-control analyses of cohort studies. Biometrics. 1984 Mar;40(1):63-75.

24. Richardson DB. An incidence density sampling program for nested case-control analyses. Occup Environ Med. 2004 Dec;61(12):e59.

25. Barsh GS, Copenhaver GP, Gibson G, Williams SM. Guidelines for genome-wide association studies. PLoS Genet. 2012 Jul;8(7):e1002812. Epub 2012 Jul 5.

26. Peach CA, Zhang Y, Dunford JE, Brown MA, Carr AJ. Cuff tear arthropathy: evidence of functional variation in pyrophosphate metabolism genes. Clin Orthop Relat Res. 2007 Sep;462:67-72.

27. Figueiredo EA, Loyola LC, Belangero PS, Campos Ribeiro-Dos-Santos ÂK, Emanuel Batista Santos S, Cohen C, Wajnsztejn A, Martins de Oliveira A, Smith MC, Pochini AC, Andreoli CV, Ejnisman B, Cohen M, Leal MF. Rotator cuff tear susceptibility is associated with variants in genes involved in tendon extracellular matrix homeostasis. J Orthop Res. 2020 Jan;38(1):192-201. Epub 2019 Oct 6.

28. Kluger R, Huber KR, Seely PG, Berger CE, Frommlet F. Novel tenascin-C haplotype modifies the risk for a failure to heal after rotator cuff repair. Am J Sports Med. 2017 Nov;45(13):2955-64. Epub 2017 Sep 27.

29. Teerlink CC, Cannon-Albright LA, Tashjian RZ. Significant association of full-thickness rotator cuff tears and estrogen-related receptor-β (ESRRB). J Shoulder Elbow Surg. 2015 Feb;24(2):e31-5. Epub 2014 Sep 11.

30. Tashjian RZ, Granger EK, Zhang Y, Teerlink CC, Cannon-Albright LA. Identification of a genetic variant associated with rotator cuff repair healing. J Shoulder Elbow Surg. 2016 Jun;25(6):865-72. Epub 2016 Apr 8.

31. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 Sep;81(3):559-75. Epub 2007 Jul 25.

32. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015 Feb 25;4:7.

33. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010 Sep 15;26(18):2336-7. Epub 2010 Jul 15.

34. National Center for Biotechnology Information. CREB5 cAMP responsive element binding protein 5 [Homo sapiens (human)]. National Center for Biotechnology Information. 2020. Accessed 2021 Mar 18. https://www.ncbi.nlm.nih.gov/gene?cmd=Retrieve&dopt=full_report&list_uids=9586

35. Yu M, Du G, Xu Q, Huang Z, Huang X, Qin Y, Han L, Fan Y, Zhang Y, Han X, Jiang Z, Xia Y, Wang X, Lu C. Integrated analysis of DNA methylome and transcriptome identified CREB5 as a novel risk gene contributing to recurrent pregnancy loss. EBioMedicine. 2018 Sep;35:334-44. Epub 2018 Aug 10.

36. He S, Deng Y, Liao Y, Li X, Liu J, Yao S. CREB5 promotes tumor cell invasion and correlates with poor prognosis in epithelial ovarian cancer. Oncol Lett. 2017 Dec;14(6):8156-61. Epub 2017 Oct 20.

37. Nevalainen T, Kananen L, Marttila S, Jylhä M, Hervonen A, Hurme M, Jylhävä J. Transcriptomic and epigenetic analyses reveal a gender difference in aging-associated inflammation: the Vitality 90+ study. Age (Dordr). 2015 Aug;37(4):9814. Epub 2015 Jul 19.

38. Bourtchuladze R, Frenguelli B, Blendy J, Cioffi D, Schutz G, Silva AJ. Deficient long-term memory in mice with a targeted mutation of the cAMP-responsive element-binding protein. Cell. 1994 Oct 7;79(1):59-68.

39. Montminy MR, Bilezikjian LM. Binding of a nuclear protein to the cyclic-AMP response element of the somatostatin gene. Nature. 1987 Jul 9-15;328(6126):175-8.

40. The Broad Institute of MIT and Harvard. GTeX Portal. CREB5. 2019. Accessed 2020 Mar 18. https://www.gtexportal.org/home/

41. Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. Cell. 2017 Jun 15;169(7):1177-86.

42. Li JF, Liu LD, Ma SH, Che YC, Wang LC, Dong CH, Zhao HL, Liao Y, Li QH. HTRP—an immediate-early gene product induced by HSV1 infection in human embryo fibroblasts, is involved in cellular co-repressors. J Biochem. 2004 Aug;136(2):169-76.

43. Benson RT, McDonnell SM, Knowles HJ, Rees JL, Carr AJ, Hulley PA. Tendinopathy and tears of the rotator cuff are associated with hypoxia and apoptosis. J Bone Joint Surg Br. 2010 Mar;92(3):448-53.

44. Lee HJ, Kim YS, Ok JH, Song HJ. Apoptosis occurs throughout the diseased rotator cuff. Am J Sports Med. 2013 Oct;41(10):2249-55. Epub 2013 Jun 20.

45. Osti L, Buda M, Del Buono A, Osti R, Massari L, Maffulli N. Apoptosis and rotator cuff tears: scientific evidence from basic science to clinical findings. Br Med Bull. 2017 Jun 1;122(1):123-33.

46. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. Nat Genet. 2001 Nov;29(3):306-9.

47. Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. Nat Rev Genet. 2002 May;3(5):391-7.

48. Border R, Johnson EC, Evans LM, Smolen A, Berley N, Sullivan PF, Keller MC. No support for historical candidate gene or candidate gene-by-interaction hypotheses for major depression across multiple large samples. Am J Psychiatry. 2019 May 1;176(5):376-87. Epub 2019 Mar 8.

49. Tashjian RZ, Kim SK, Roche MD, Jones KB, Teerlink CC. Genetic variants associated with rotator cuff tearing utilizing multiple population-based genetic resources. J Shoulder Elbow Surg. 2021 Mar;30(3):520-31. Epub 2020 Jul 12.

50. Urwin M, Symmons D, Allison T, Brammah T, Busby H, Roxby M, Simmons A, Williams G. Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation. Ann Rheum Dis. 1998 Nov;57(11):649-55.

51. Ostör AJ, Richards CA, Prevost AT, Speed CA, Hazleman BL. Diagnosis and relation to general health of shoulder disorders presenting to primary care. Rheumatology (Oxford). 2005 Jun;44(6):800-5. Epub 2005 Mar 15.

52. Judge A, Murphy RJ, Maxwell R, Arden NK, Carr AJ. Temporal trends and geographical variation in the use of subacromial decompression and rotator cuff repair of the shoulder in England. Bone Joint J. 2014 Jan;96-B(1):70-4.

53. UK Biobank. UK Biobank leads the way in genetics research to tackle chronic diseases. 2019. Accessed 2020 May 13. https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/news/uk-biobank-leads-the-way-in-genetics-research-to-tackle-chronic-diseases-1

54. Yanik EL, Colditz GA, Wright RW, Saccone NL, Evanoff BA, Jain NB, Dale AM, Keener JD. Risk factors for surgery due to rotator cuff disease in a population-based cohort. Bone Joint J. 2020 Mar;102-B(3):352-9.

55. Assunção JH, Godoy-Santos AL, Dos Santos MCLG, Malavolta EA, Gracitelli MEC, Ferreira Neto AA. Matrix metalloproteases 1 and 3 promoter gene polymorphism is associated with rotator cuff tear. Clin Orthop Relat Res. 2017 Jul; 475(7):1904-10.