SUMO4 and *MAP3K7IP2* Single Nucleotide Polymorphisms and Susceptibility to Rheumatoid Arthritis

JAVIER COSTAS, EVA PEREZ-PAMPIN, ISABEL FERREIROS-VIDAL, MARIA TORRES, CHRISTOPHER PHILLIPS, JOSE LUIS VICARIO, JOSE LUIS PABLOS, ANGEL CARRACEDO, JUAN J. GOMEZ-REINO, and ANTONIO GONZALEZ

ABSTRACT. ***Objective.*** To explore the role of single nuclear polymorphisms (SNP) in 2 candidate genes, *SUMO4* and *MAP3K7IP2*, in susceptibility to rheumatoid arthritis (RA).

Methods. Two cohorts from different Spanish towns totalling 635 patients with RA and 826 controls were studied. Six SNP were genotyped by matrix assisted laser desorption-ionization time-of-flight (MALDI-TOF) with the MassARRAY SNP genotyping system.

Results. We found no association with susceptibility to RA for any of the SNP including a previously described functional variant in the *SUMO4* gene (163A→G). RA susceptibility was independent of the haplotypes defined by the 6 SNP and there was also no association with clinical features of RA.

Conclusion. *SUMO4* and *MAP3K7IP2* SNP did not significantly influence predisposition to and features of RA, in contrast to previous genetic and functional evidence that suggested their involvement. (J Rheumatol 2006;33:1048–51)

Key Indexing Terms:
RHEUMATOID ARTHRITIS | TYPE 1 DIABETES
NUCLEAR FACTOR-κB | GENETIC SUSCEPTIBILITY | IDDM5

Recently, a series of single nucleotide polymorphisms (SNP) in the IDDM5 locus on chromosome 6p25 were found to be strongly associated with susceptibility to type 1 diabetes[1,2]. Among them, a functional SNP in the small ubiquitin-like modifier 4 (*SUMO4*) gene (163A→G or rs237025) was singled out as a susceptibility factor for type 1 diabetes. The IDDM5 locus has also been linked to rheumatoid arthritis (RA) in multicase families[3], indicating shared genetic factors between the 2 autoimmune diseases. Other commonalities between these 2 diseases include their excessive coincidence within families[4], the role of the HLA locus and of the protein tyrosine phosphatase nonreceptor-type 22[5,6], and the characterization of both as Th1 autoimmune diseases. The central role of the transcription factor nuclear factor-κB (NF-κB) in RA inflammation[7] adds to interest in the IDDM5 locus, as the 2 known candidate genes in the locus, *SUMO4* and the mitogen-activated protein kinase kinase kinase 7-interacting protein 2 (*MAP3K7IP2*), regulate this transcription factor. *SUMO4* boosts IκBα-negative regulation of NF-κB transcriptional activity[1] and *MAP3K7IP2* is part of a protein complex that activates NF-κB in response to interleukin 1[8]. These observations all suggest that the IDDM5 locus may have a role in susceptibility to RA.

MATERIALS AND METHODS

Samples. DNA samples were obtained from peripheral blood of Spanish patients with RA and healthy controls. They were from teaching hospitals in 2 widely separated Spanish towns, University Clinical Hospital in Santiago de Compostela and "12 of October" Hospital in Madrid. Patients with RA were classified according to the 1987 American College of Rheumatology criteria[9]. In Santiago, we sought to enroll all patients with RA followed in the Rheumatology Unit, and the first 455 recruited patients (from June 2001 to July 2004) were included. In Madrid, patients were recruited among outpatients of the Rheumatology Unit over an extended period. Control samples, in Santiago, were collected from 643 subjects older than age 55 years undergoing preoperative preparation for elective surgery excluding orthopedics. In Madrid, controls were blood bank donors. Clinical data for patients and controls are provided in Table 1. The Ethical Committee for Clinical Research of Galicia approved this study and all participants gave their written informed consent according to the Declaration of Helsinki.

From Research Laboratory 2, Rheumatology Unit, and National Genotyping Center, Hospital Clinico Universitario de Santiago, Santiago de Compostela; and the Rheumatology Unit, "12 of October" Hospital and the Regional Transfusion Center, Madrid, Spain.

Supported by the Instituto de Salud Carlos III (Spain) with funds from FEDER (European Union) and grants from Xunta de Galicia, Genoma España (Centro Nacional de Genotipado), and Schering-Plough España SA.

J. Costas, PhD; M. Torres, BSc, National Genotyping Center, Hospital Clinico Universitario de Santiago; E. Perez-Pampin, MD; I. Ferreiros-Vidal, BA, MA; A. Gonzalez, MD, PhD, Research Laboratory 2 and Rheumatology Unit, Hospital Clinico Universitario de Santiago; C. Phillips, MSc; A. Carracedo, MD, PhD, National Genotyping Center, Hospital Clinico Universitario de Santiago and the Institute of Legal Medicine, University of Santiago de Compostela; J.L. Vicario, MD, Regional Transfusion Center; J.L. Pablos, MD, PhD, Rheumatology Unit, "12 of October" Hospital; J.J. Gomez-Reino, MD, PhD, Research Laboratory 2, Rheumatology Unit, Hospital Clinico Universitario de Santiago and the Department of Medicine, University of Santiago de Compostela.

Address reprint requests to Dr. A. Gonzalez, Laboratorio de Investigacion 2, Hospital Clinico Universitario de Santiago, 15706 Santiago de Compostela, Spain. E-mail: antonio.gonzalez.martinez.pedrayo@sergas.es

Accepted for publication February 6, 2006.

Table 1. General characteristics of patients with RA and controls from 2 Spanish towns.

	Santiago		Madrid	
	RA, n = 455	Controls, n = 643	RA, n = 180	Controls, n = 183
Female, %	77.2	53.2	72.2	68.3
Age at recruitment*	60 (43–79)	69 (62–76)	—	—
Age at onset*	49 (37–58)		49 (39–57)	
RF, %	58.5		87.1	
Bone erosions, %	67.0		73.1	
DMARD use, %	96.7		—	
Tobacco use, %	20.3		—	

* Median (interquartilic range). The last 4 clinical characteristics from the table were available for more than 75% of the RA patients from Santiago and about 50% of the RA patients from Madrid.

SNP genotyping. Six SNP from Guo, *et al*[1] were analyzed with the MassARRAY SNP genotyping system (Sequenom Inc., San Diego, CA, USA) following the manufacturer's instructions[10] as described[11]. Briefly, the 6 SNP were amplified in a multiplex polymerase chain reaction, followed by minisequencing assays, designed with SpectroDesigner software (Sequenom Inc.) (primer sequences are available at http://bioinformatics.cesga.es/supmat/; accessed March 10, 2006). Products were separated by mass spectrometry analysis with the Bruker Bi-flex matrix assisted laser desorption-ionization time-of-flight (MALDI-TOF) mass spectrometer (Bruker Daltonics, Billerica, MA, USA). Spectral output was analyzed using SpectroTyper–RT v3.1 software (Sequenom Inc.) and by manual review. Samples from the 2 recruiting centers were analyzed simultaneously. As quality control, re-genotyping of 25 samples was undertaken throughout the study. Repeat genotypes (150) were in 100% agreement with original results. The 6 SNP were successfully genotyped in more than 99% of the 1460 individuals studied. Genotype frequencies of each were in accordance with Hardy-Weinberg equilibrium when considered globally or stratified by status, sex, or town of recruitment.

Statistical analysis. Allelic frequencies were compared using chi-squared tests from contingency 2 × 2 tables. Correlation analysis and logistic regression were used to evaluate the relationship between genotypes and clinical features of RA following an additive genetic model with Statistica software (Statsoft, Tulsa, OK, USA). Adjustment by age of onset and sex was done with logistic regression. Accordance with Hardy-Weinberg equilibrium was tested using the Markov chain algorithm[12]. Cocaphase software[13] was used to infer haplotypes by the expectation-maximization algorithm and to compare haplotype frequency distribution by the homogeneity likelihood ratio test. Haplotypes with frequencies below 5% in any of the strata were grouped.

RESULTS

We studied the 163A→G SNP and 5 other flanking SNP that span 150 kb in the IDDM5 locus and include the *MAP3K7IP2* and *SUMO4* genes. They were genotyped in 2 independent case-control collections of DNA samples from Spanish subjects. The first collection included 455 patients with RA and 643 healthy controls from Santiago de Compostela and the second comprised 180 patients with RA and 183 controls from Madrid. Women constituted 77.2% of the RA patients in the Santiago cohort and 72.2% in the Madrid cohort. We observed significant differences between controls from the 2 collections for 2 of the SNP (163A→G and 012Taq; $p < 0.03$ for both). Upon analysis, we realized these differences were due exclusively to differences in the women. Consequently, we analyzed data from the 2 cohorts separately and stratified by sex. There were no significant differences between patients and controls in the allelic frequencies of any of the 6 SNP (Tables 2 and 3) or in the genotypes (available at http://bioinformatics.cesga.es/supmat/). In addition, clinical features of RA were independent of the IDDM5 SNP genotypes, as evaluated by age at disease onset, presence of rheumatoid factor, or radiographic signs of erosion (data not shown).

All 6 SNP were in linkage disequilibrium, but this was much more marked between the 4 central SNP (from 001Msp to 012Taq) that consisted of a haplotype block (mean pairwise $D' = 0.94$) covering the *SUMO4* and *MAP3K7IP2* genes. There were no significant differences in haplotype frequencies between patients and controls from either Santiago or Madrid, either for all 6 SNP (8 haplotypes with estimated frequencies > 5% in at least one of the groups) or for only the 4 SNP of the haplotype block (5 haplotypes with estimated frequencies > 5%). When the haplotypes were also stratified by sex (Table 4) there was a difference between male RA patients and controls from Santiago that did not reach significance ($p = 0.06$); this was not corroborated in male patients and controls from Madrid.

We also analyzed differences in allele frequencies between the 2 collections of samples. There were significant differences between women from the 2 towns; both controls and RA patients from Madrid showed increased allele frequencies. This dissimilarity included the 4 SNP that formed the haplotype block. In contrast, there were no differences between men in any of the subgroups. The differences found in women showed the importance of carefully avoiding population stratification in genetic epidemiology studies to prevent spurious results.

DISCUSSION

The hypothesis implicating these SNP in susceptibility to RA has been supported by the RA linkage to the IDDM5 locus[3,4], in the association of the latter with these SNP in other studies[1,2,14,15], and in the functional consequences of the 163A→G *SUMO4* SNP. This SNP has been shown to affect

Table 2. Lack of difference in IDDM5 SNP allele frequencies between patients with RA and controls from Santiago de Compostela. Values are expressed as percentages.

SNP	Allele	Women			Men		
		Controls, n = 684	RA, n = 702	p	Controls, n = 602	RA, n = 208	p
373 Taq	G	34.5	34.0	NS	33.4	30.6	NS
001Msp	T	42.1	43.7	NS	38.7	37.4	NS
268Hha	T	43.2	45.8	NS	41.9	38.5	NS
163A→G	G	45.2	42.4	NS	48.7	50.5	NS
012Taq	A	45.2	42.1	NS	48.3	49.5	NS
454Msp	G	39.2	41.5	NS	37.0	37.5	NS

n: number of chromosomes; NS: not significant (values of $p < 0.1$ are shown).

Table 3. Lack of difference in IDDM5 SNP allele frequencies between patients with RA and controls from Madrid. Values are expressed as percentages.

SNP	Allele	Women			Men		
		Controls, n = 250	RA, n = 260	p	Controls, n = 116	RA, n = 100	p
373 Taq	G	30.4	32.6	NS	31.9	36.0	NS
001Msp	T	34.4	35.4	NS	41.4	41.0	NS
268Hha	T	36.2	35.8	NS	43.1	42.0	NS
163A→G	G	55.2	56.1	NS	49.1	50.0	NS
012Taq	A	55.2	56.1	NS	49.1	49.0	NS
454Msp	G	32.8	38.7	NS	37.9	45.0	NS

n: number of chromosomes; NS: not significant ($p < 0.1$).

Table 4. Haplotype frequencies of IDDM5 SNP in patients with RA and controls from Santiago and Madrid stratified by sex. Only haplotypes with frequency > 5% in at least one of the groups were considered.

Haplotypes	Santiago				Madrid			
	Women		Men		Women		Men	
	Control	RA	Control	RA	Control	RA	Control	RA
ACCAGA	37.1	32.6	41.7	38.0	47.9	45.1	40.9	36.0
GTTGAG	16.2	16.7	15.7	11.6	14.1	18.9	14.8	18.3
GTTGAA	11.2	9.9	12.6	10.2	8.2	6.0	11.9	12.8
ACCGAG	9.3	10.0	8.3	8.4	6.2	6.7	8.7	9.3
ATTGAG	7.7	7.9	6.2	3.4	6.2	5.0	10.4	5.9
ATTGAA	6.8	9.3	4.4	11.7	5.5	4.7	3.3	2.6
GCCAGA	4.8	5.3	3.3	4.1	3.5	5.2	3.2	2.6
ACCAGG	3.0	3.6	3.2	6.7	2.7	5.1	2.9	9.5

the activity of NF-κB[1], which is a central transcription factor in RA[16]. In addition, a second candidate gene in the type 1 diabetes-associated region, *MAP3K7IP2*[1,2], has also been reported to affect NF-κB function, although no functional polymorphism has been described in this large gene (about 95 kb). Despite this hypothesis, we did not find evidence of an association between RA susceptibility and SNP in these 2 genes. Our sample sizes, the concordant results in the 2 sample collections, and the absence of association with haplotypes or clinical features of RA all support our conclusion that these SNP are not associated with susceptibility to RA. Even the female-specific subpopulation heterogeneity we detected was unlikely to have affected our results because we observed similar results in men and women and because we stratified the study by town and by sex. In addition, our results corroborate an independent report of lack of association with RA susceptibility of the 163A→G *SUMO4* SNP in a UK study[17].

The genetic heterogeneity we found between the 2 cohorts was surprising because it was restricted to women and because all samples were Spanish and this population has low genetic heterogeneity[18]. The differences were found in 2 independent samples consisting of controls and patients with RA, respectively. This duplication excludes population stratification, as cases and controls were selected independently. Also,

systematic errors in genotyping are very unlikely as the samples were studied without previous grouping by sex. Consequently, we believe complex genetic mechanisms including but not restricted to transmission ratio distortion (TRD) could be a possible explanation. There was already evidence of TRD in the 6q25 region[19-21] and TRD among offspring seems widespread in the human genome[22]. However, TRD is not sufficient to explain our results because the differences between sexes were town-dependent and unknown site-specific factors would be required. If indeed complex genetics affect the IDDM5 locus in many populations, they could explain the controversial results of *SUMO4* 163A→G SNP association with type 1 diabetes in studies showing a susceptibility effect for the G allele[1,14,15], or the A allele[2], or lack of association[23,24].

We did not find evidence of an association between *SUMO4* or *MAP3K7IP2* SNP and RA susceptibility or severity. In addition, we did find marked heterogeneity of some SNP in the IDDM5 locus among Spanish women that could interfere with association and linkage studies.

ACKNOWLEDGMENT

We thank sample donors for their collaboration and Drs. Antonio Mera, Manuel Caamaño, Jorge Blanco, and Santos Insua for providing access to their patients. We also thank Yolanda Lopez-Golan and Fina Meijide for their help in recruiting study participants. Marta Picado, Ines Quintela, Cristina Fernandez, and Jorge Amigo provided outstanding technical assistance.

REFERENCES

1. Guo D, Li M, Zhang Y, et al. A functional variant of SUMO4, a new I kappa B alpha modifier, is associated with type 1 diabetes. Nat Genet 2004;36:837-41.
2. Owerbach D, Pina L, Gabbay KH. A 212-kb region on chromosome 6q25 containing the TAB2 gene is associated with susceptibility to type 1 diabetes. Diabetes 2004;53:1890-3.
3. Myerscough A, John S, Barrett JH, Ollier WE, Worthington J. Linkage of rheumatoid arthritis to insulin-dependent diabetes mellitus loci: evidence supporting a hypothesis for the existence of common autoimmune susceptibility loci. Arthritis Rheum 2000;43:2771-5.
4. Lin JP, Cash JM, Doyle SZ, et al. Familial clustering of rheumatoid arthritis with other autoimmune diseases. Hum Genet 1998;103:475-82.
5. Begovich AB, Carlton VE, Honigberg LA, et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 2004;75:330-7.
6. Bottini N, Musumeci L, Alonso A, et al. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nat Genet 2004;36:337-8.
7. Firestein GS. NF-kappaB: Holy Grail for rheumatoid arthritis? Arthritis Rheum 2004;50:2381-6.
8. Takaesu G, Kishida S, Hiyama A, et al. TAB2, a novel adaptor protein, mediates activation of TAK1 MAPKKK by linking TAK1 to TRAF6 in the IL-1 signal transduction pathway. Mol Cell 2000;5:649-58.
9. Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315-24.
10. Costas J, Torres M, Cristobo I, Phillips C, Carracedo A. Relative efficiency of the linkage disequilibrium mapping approach in detecting candidate genes for schizophrenia in different European populations. Genomics 2005;86:280-6.
11. Buetow KH, Edmonson M, MacDonald R, et al. High-throughput development and characterization of a genomewide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Proc Natl Acad Sci USA 2001;98:581-4.
12. Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 1992;48:361-72.
13. Dudbridge F. Pedigree disequilibrium tests for multilocus haplotypes. Genet Epidemiol 2003;25:115-21.
14. Park Y, Park S, Kang J, Yang S, Kim D. Assessing the validity of the association between the SUMO4 M55V variant and risk of type 1 diabetes [letter]. Nat Genet 2005;37:112; author reply 112-3.
15. Wang CY, Yang P, She JX. Reply to "Assessing the validity of the association between the SUMO4 M55V variant and risk of type 1 diabetes." Nat Genet 2005;37:112-3.
16. Firestein GS. NF-kappaB: Holy Grail for rheumatoid arthritis? Arthritis Rheum 2004;50:2381-6.
17. Gibbons LJ, Thomson W, Zeggini E, et al. The type 1 diabetes susceptibility gene SUMO4 at IDDM5 is not associated with susceptibility to rheumatoid arthritis or juvenile idiopathic arthritis. Rheumatology Oxford 2005;44:1390-3.
18. Flores C, Maca-Meyer N, Gonzalez AM, et al. Reduced genetic structure of the Iberian peninsula revealed by Y-chromosome analysis: implications for population demography. Eur J Hum Genet 2004;12:855-63.
19. Naumova AK, Greenwood CM, Morgan K. Imprinting and deviation from Mendelian transmission ratios. Genome 2001;44:311-20.
20. McCann JA, Xu YQ, Frechette R, Guazzarotti L, Polychronakos C. The insulin-like growth factor-II receptor gene is associated with type 1 diabetes: evidence of a maternal effect. J Clin Endocrinol Metab 2004;89:5700-6.
21. Lemire M, Roslin NM, Laprise C, Hudson TJ, Morgan K. Transmission-ratio distortion and allele sharing in affected sib pairs: a new linkage statistic with reduced bias, with application to chromosome 6q25.3. Am J Hum Genet 2004;75:571-86.
22. Zollner S, Wen X, Hanchard NA, Herbert MA, Ober C, Pritchard JK. Evidence for extensive transmission distortion in the human genome. Am J Hum Genet 2004;74:62-72.
23. Smyth DJ, Howson JM, Lowe CE, et al. Assessing the validity of the association between the SUMO4 M55V variant and risk of type 1 diabetes. Nat Genet 2005;37:110-1.
24. Qu H, Bharaj B, Liu XQ, et al. Assessing the validity of the association between the SUMO4 M55V variant and risk of type 1 diabetes. Nat Genet 2005;37:111-2.