Gene-Gene and Gene-Environment Interactions Involving HLA-DRB1, *PTPN22,* and Smoking in Two Subsets of Rheumatoid Arthritis

Henrik Källberg, Leonid Padyukov, Robert M. Plenge, Johan Rönnelid, Peter K. Gregersen, Annette H. M. van der Helm-van Mil, Rene E. M. Toes, Tom W. Huizinga, Lars Klareskog, and Lars Alfredsson, for the Epidemiological Investigation of Rheumatoid Arthritis (EIRA) Study Group

Gene-gene and gene-environment interactions are key features in the development of rheumatoid arthritis (RA) and other complex diseases. The aim of this study was to use and compare three different definitions of interaction between the two major genetic risk factors of RA—the HLA-DRB1 shared epitope (SE) alleles and the *PTPN22* R620W allele—in three large case-control studies: the Swedish Epidemiological Investigation of Rheumatoid Arthritis (EIRA) study, the North American RA Consortium (NARAC) study, and the Dutch Leiden Early Arthritis Clinic study (in total, 1,977 cases and 2,405 controls). The EIRA study was also used to analyze interactions between smoking and the two genes. "Interaction" was defined either as a departure from additivity, as interaction in a multiplicative model, or in terms of linkage disequilibrium—for example, deviation from independence of penetrance of two unlinked loci. Consistent interaction, defined as departure from additivity, between HLA-DRB1 SE alleles and the A allele of *PTPN22* R620W was seen in all three studies regarding anti-CCP–positive RA. Testing for multiplicative interactions demonstrated an interaction between the two genes only when the three studies were pooled. The linkage disequilibrium approach indicated a gene-gene interaction in EIRA and NARAC, as well as in the pooled analysis. No interaction was seen between smoking and *PTPN22* R620W. A new pattern of interactions is described between the two major known genetic risk factors and the major environmental risk factor concerning the risk of developing anti-CCP–positive RA. The data extend the basis for a pathogenetic hypothesis for RA involving genetic and environmental factors. The study also raises and illustrates principal questions concerning ways to define interactions in complex diseases.

Rheumatoid arthritis (RA [MIM 180300]) is a prototype of an autoimmune disease with complex etiology that is assumed to involve several genetic as well as environmental factors. The clinical hallmark of RA is symmetrical inflammatory arthritis. The disease is more common among women than men. The major risk factors that have so far been reproducibly identified are genetic variations in the major histocompatibility complex, class II, DR beta 1 (HLA-DRB1 [MIM 142857]) and protein tyrosine phosphatase (*PTPN22* [MIM 600716])[1–4] genes and one environmental risk factor, smoking.[5–8] With regard to all three risk factors, the major effects have been seen in one subset of RA, characterized by the presence of antibodies to citrullinated proteins (anti-CCP), but not in the subset of RA in which these antibodies are not detected.[2,3,7,8] Recently, a pronounced gene-environment interaction was identified between smoking and HLA-DRB1 shared epitope (SE) alleles.[6–8] The demonstration of this gene-environment interaction, together with immunological studies in animal models of arthritis, led us to form a new etiologic hypothesis suggesting that smoking contributes to citrullination and to triggering of anti-citrulline immunity, which is restricted by HLA-DRB1 SE alleles.[7,9] Since the *PTPN22* gene codes for a tyrosine phosphatase, with a potential function in the regulation of T-cell and B-cell activation, it is of obvious interest for the study of the etiology of RA to know how this more recently described risk gene interacts with the classical HLA-DRB1 SE genes, as well as with smoking.

Studies of interaction among risk factors in the epidemiological literature have classically been performed using a departure from the additivity model originally described by Rothman, where a term—the attributable proportion due to interaction (AP)—is used to quantify the contribution of interaction to a disease risk, as compared with the contribution of each of the two risk factors added to each other.[10–12] This method can also be used to quantify gene-gene interactions for unlinked loci. An alternative common method for quantifying gene-gene interactions is based on the calculation of the two risk factors' product

From the Institute of Environmental Medicine, Karolinska Institutet (H.K.; L.A.), Rheumatology Unit, Department of Medicine, Karolinska Institutet/Karolinska Hospital (L.P.; L.K.), and Stockholm Center for Public Health, Stockholm County Council (L.A.), Stockholm; Broad Institute of MIT and Harvard, Cambridge, MA (R.M.P.); Division of Rheumatology, Brigham and Women's Hospital, Boston (R.M.P.); Unit of Clinical Immunology, Uppsala University, Uppsala, Sweden (J.R.); Robert S. Boas Center for Genomics and Human Genetics, Feinstein Institute for Medical Research, North Shore Long Island Jewish Health System, Manhasset, NY (P.K.G.); and Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands (A.H.M.v.d.H.-v.M.; R.E.M.T.; T.W.H.)

Received January 16, 2007; accepted for publication February 13, 2007; electronically published April 2, 2007.

Address for correspondence and reprints: Dr. Henrik Källberg, Institute for Environmental Medicine, Box 210, Karolinska Institutet, 171 77 Stockholm, Sweden. E-mail: henrik.kallberg@ki.se

Am. J. Hum. Genet. 2007;80:867–875. 0002-9297/2007/8005-0006$15.00
DOI: 10.1086/516736

term in a logistic-regression model (multiplicative or statistical interaction). Recently, another method for detecting gene-gene interaction, based on deviation from independence of penetrance in two unlinked loci, was proposed.[13]

As reported in this article, we used these three methods to calculate interactions between the HLA-DRB1 SE alleles and *PTPN22* risk allele (R620W) in three different major case-control studies of RA: the Swedish Epidemiological Investigation of Rheumatoid Arthritis (EIRA) study, the North American RA Consortium (NARAC) study, and a Dutch case-control study based on the Leiden early arthritis cohort.[5,14–17] We used the largest one of these, which is also the one in which smoking information is the most detailed (the EIRA study[5–7,14]), to calculate interaction between the *PTPN22* risk allele and smoking, with regard to the risk of anti-CCP–positive RA, using the departure-from-additivity model. A significant interaction between HLA-DRB1 SE alleles and the *PTPN22* R620W allele was seen with all methods when the three studies were pooled, but the departure-from-additivity model was able to identify interaction in each one of the studies analyzed separately, whereas the multiplicative model identified this interaction only when the three studies were pooled. No interaction was seen between smoking and *PTPN22*. This new description of the interaction patterns among the three major known risk factors of RA, which all work in one but not the other major subset of the disease, provides a basis for further etiologic investigations of this disease, with an increased emphasis on T-cell activation and specificity. In addition, the data illustrate the need for consideration and comparison of different methods for calculating interactions in complex diseases.

Material and Methods

This report is based on data from three different studies: (1) the EIRA, (2) the NARAC, and (3) the Leiden Early Arthritis Clinic (Leiden EAC). In each study, ethical permission was obtained from the relevant ethical committees, and all the participants consented to contribute to the study.

In the EIRA study, we recruited cases and controls aged 18–70 years during May 1996 to December 2003 from a geographically defined area in the south and central regions of Sweden. A case was defined as a person in the study base who was given a first-time diagnosis of RA, according to the American College of Rheumatology (ACR) criteria of 1987,[18] by rheumatologists at private as well as general health care units in the participating area. Controls were randomly selected from a continuously updated national population register, with consideration given to age, sex, and living area. If the selected control was not traceable, reported having RA, or refused to participate, a new control was selected using the same procedure. More details about the study are described in other reports of the EIRA study.[5–7,14,19]

The second study used in this report, the NARAC, consists of affected sibling pairs who fulfilled the 1987 ACR criteria[18] and matched, unrelated controls recruited from the New York Cancer Project. Details regarding NARAC are provided elsewhere.[15,16]

The third study used, the Leiden EAC, consists of patients with confirmed RA with early onset and hospital-based controls with and without deep venous thrombosis. Details on the Leiden EAC and the controls used are given elsewhere.[17,20]

Biological Parameters and Genotyping

Analysis of anti-CCP antibodies for samples from the EIRA study, as well as for samples from the Leiden EAC, was made with the Immunoscan-RA Mark2 ELISA, as described elsewhere.[2,21] Cases with serum antibody levels >25 U/ml were regarded as anti-CCP positive. Rheumatoid factor status for cases was determined using standard procedures. Antibodies toward CCP in NARAC cases were measured using a second-generation commercial anti-CCP ELISA (Inova Diagnostics). Cases with antibody levels >20 U/ml were regarded as anti-CCP positive.

Genotyping procedure of the *PTPN22* and HLA-DRB1 genes is described in previous articles.[2,4,7–9,22] Allocation to SE was based on local guidelines as described elsewhere, with the HLA-DRB1 alleles DRB1*01, DRB1*04, and DRB1*10 defined as SE alleles. For simplicity of evaluation, we assumed a dominant A allele model for the *PTPN22* R620W polymorphism and a dominant SE allele model. Subjects with SE alleles were classified as having single or double SE alleles, for a separate evaluation of a potential gene-dosage effect.

Smoking Information

Smoking was considered only for the EIRA study. On the basis of self-reported information from the EIRA questionnaire, we divided participants into "ever smokers" and "never smokers." An ever smoker was defined as a person who had ever smoked cigarettes. A never smoker was defined as a person who had never smoked cigarettes, pipes, cigars, etc.

Statistical Analysis

We calculated odds ratios (ORs) and 95% CIs of developing anti-CCP–positive RA and anti-CCP–negative RA associated with each of the factors—*PTPN22* R620W alleles, HLA-DRB1 SE alleles, and smoking—by means of logistic-regression models. We estimated their separate effects, as well as the effect of their simultaneous presence. All analyses based on the EIRA study were adjusted for the matching variables: age (10 categories), living area (11 categories), and sex, together with smoking and presence of HLA-DRB1 SE alleles when appropriate. Analyses based on the NARAC study or the Leiden EAC were adjusted for sex and age (using the same categories as in EIRA).

Interaction effects between the presence of HLA-DRB1 SE and *PTPN22* R620W alleles were evaluated using three different approaches. First, interaction between presence of HLA-DRB1 and *PTPN22* R620W alleles was evaluated using departure from additivity of effects as the interaction criteria (biological interaction), as suggested by Rothman.[10,11] To quantify the amount of interaction in terms of departure from additivity of effects, the AP was calculated together with the 95% CI.[12] The AP is the proportion of the incidence among persons exposed to two interacting factors that is attributable to the interaction per se (i.e., reflecting their joint effect beyond the sum of their independent effects). A more detailed description regarding explanation, definition, and programs used to calculate AP with CIs is given elsewhere.[10,12,23] We assessed the AP for each study as well as the pooled estimate when all studies were combined (with adjustment for study). To investigate potential mechanistic pathways further, we also calculated interaction effects between the pres-

Table 1. ORs and 95% CIs for Developing Anti-CCP–Positive RA Associated with Presence of *PTPN22* and HLA-DRB1 Alleles and with Smoking

Study and Risk Factor	No. of Cases/ Controls	OR	95% CI	OR[a]	95% CI
EIRA:					
PTPN22	230/179	1.6[b]	1.3–2.1	1.7[b]	1.3–2.2
HLA-DRB1 SE	615/414	5.6[b]	4.4–7.2	5.8[b]	4.5–7.6
Smoking	520/479	1.8[b]	1.4–2.2	1.9[b]	1.5–2.4
NARAC:					
PTPN22	113/116	2.8[c]	2.0–3.8	2.6[c]	1.8–3.7
HLA-DRB1 SE	311/314	11.3[c]	7.7–16.5	12.0[c]	8.0–17.9
Smoking	197/357	1.3[c]	1.0–1.7	1.3[c]	1.0–1.8
Leiden EAC:					
PTPN22	58/154	1.8[c]	1.3–2.6	1.8[c]	1.2–2.7
HLA-DRB1 SE	202/308	5.6[c]	3.9–7.9	6.1[c]	4.1–8.9

NOTE.—Results are displayed for each study.
[a] All risk factor variables in the same logistic-regression model.
[b] OR adjusted for sex, age, and living area.
[c] OR adjusted for sex and age.

ence of *PTPN22* R620W alleles and smoking and between smoking and the presence of HLA-DRB1 SE alleles, using the departure-from-additivity–based method.

Second, interaction between presence of HLA-DRB1 and *PTPN22* R620W alleles was evaluated using the cross-product of the risk factors in a logistic-regression model as interaction criteria (multiplicative or statistical interaction).[10,11] We also pooled the three studies, to increase power to detect multiplicative interaction. In the analysis of multiplicative interaction, we adjusted for age, living area (only EIRA), and sex, as well as for study in the pooled analysis.

Third, we used a recently developed method to assess gene-gene interaction, which is based on deviation from independence of penetrance of two unlinked loci. This is referred to as the "linkage disequilibrium (LD) statistic."[13] We also calculated χ^2 statistics to compare allele distribution in the different studies, for cases and controls separately. We used SAS software for Windows, version 9.1 (SAS Institute), to analyze the data.

Results

This report is based on a total of 1,977 cases and 2,405 controls, distributed across the different studies as follows: EIRA, 1,183 cases and 793 controls; NARAC, 430 cases and 731 controls; and Leiden EAC, 364 cases and 881 controls.

The EIRA study mainly involved individuals born in Sweden. When those born outside Sweden were also taken into consideration, 97% were of white origin. Of the cases, ~61% had anti-CCP–positive RA. The corresponding figures for presence of anti-CCP in the NARAC and the Leiden EAC studies were 81% and 56%, respectively. As published elsewhere,[2,4] the A allele of *PTPN22* R620W and the HLA-DRB1 SE alleles were both associated with increased risks of developing anti-CCP–positive RA, separately as well as when analyzed simultaneously in the same model (table 1). In general, ORs calculated from the NARAC study were higher than those calculated from the EIRA or the Leiden EAC study. In NARAC, the A allele of *PTPN22* R620W and the HLA-DRB1 SE alleles were also associated with increased ORs for anti-CCP–negative RA, although not at the magnitude as those for anti-CCP–positive RA (table 2).

Interaction between PTPN22 *Alleles and HLA-DRB1 SE, Regarding Risk of Anti-CCP–Positive and Anti-CCP–Negative RA*

The cell frequencies for different combinations of HLA-DRB1 and *PTPN22* alleles for cases and controls in the three studies are displayed in table 3. There was a high resemblance in allele distribution for HLA-DRB1 and *PTPN22* in the three patient materials. The controls, however, differed in allele distribution both for HLA-DRB1 and *PTPN22,* with higher prevalence regarding both HLA-DRB1 SE alleles and the A allele of *PTPN22* R620W in the EIRA as compared with the NARAC ($P < .001$) and the Leiden controls ($P < .001$).

In the EIRA study, the OR of developing anti-CCP–positive RA for subjects having at least one copy of the A allele of *PTPN22* R620W and at least one HLA-DRB1 SE allele was 9.9 (95% CI 6.8–14.3), compared with subjects with the nonsusceptible GG genotype of *PTPN22* R620W and without HLA-DRB1 SE alleles (table 4). In the NARAC study, ORs were higher than corresponding ORs based on the EIRA material. The overall results from Leiden EAC were similar to the results from the EIRA study. No significant risk for anti-CCP–positive RA was conferred by the R620W A allele in individuals lacking the HLA-DRB1 SE allele, although the pooled estimate for the three studies combined was associated with a slightly increased risk (OR 1.4 [95% CI 1.0–2.1]). Since the results were similar for women and men, we display results for both sexes combined only.

Quantification of the interaction between HLA-DRB1 SE and the *PTPN22* R620W A allele, by means of calculating the AP, showed that the interaction was statistically significant in all three studies (EIRA AP 0.5 [95% CI 0.3–0.7], NARAC AP 0.7 [95% CI 0.5–0.9], Leiden EAC AP 0.4 [95% CI 0.1–0.8], and pooled material AP 0.5 [95% CI 0.4–0.6]).

Table 2. ORs and 95% CIs for Developing Anti-CCP–Negative RA Associated with Presence of *PTPN22* and HLA-DRB1 Alleles and with Smoking for the EIRA Study

Study and Risk Factor	No. of Cases/ Controls	OR	95% CI	OR[a]	95% CI
EIRA:					
PTPN22	114/179	1.2[b]	.9–1.5	1.2[b]	.9–1.6
HLA-DRB1 SE	261/414	1.2[b]	.9–1.5	1.2[b]	.9–1.5
Smoking	264/479	.9[b]	.7–1.2	.9[b]	.7–1.2
NARAC:					
PTPN22	22/116	1.9[c]	1.1–3.3	2.0[c]	1.1–3.4
HLA-DRB1 SE	58/314	3.1[c]	1.9–5.1	3.1[c]	1.9–5.1
Leiden EAC:					
PTPN22	26/152	.9[c]	.6–1.4	.9[c]	.6–1.4
HLA-DRB1 SE	83/385	1.4[c]	1.0–2.0	1.4[c]	1.0–2.0

NOTE.—Results are displayed for each study.
[a] All risk factor variables in the same logistic-regression model.
[b] OR adjusted for sex, age, and living area.
[c] OR adjusted for sex and age.

Table 3. Presence or Absence of R620W *PTPN22* Susceptible Allele and HLA-DRB1 SE Alleles for Anti-CCP–Positive and Anti-CCP–Negative Cases and Controls, by Study

Study Group, Population, and HLA-DRB1 Status	No. (%) of Individuals with	
	No R620W *PTPN22* Susceptible Allele	Any R620W *PTPN22* Susceptible Allele
EIRA:		
Anti-CCP–positive cases:		
No SE	75 (10)	31 (4)
Any SE	416 (58)	199 (28)
Anti-CCP–negative cases:		
No SE	142 (31)	59 (13)
Any SE	206 (45)	55 (12)
Controls:		
No SE	284 (36)	95 (12)
Any SE	330 (42)	84 (10)
NARAC:		
Anti-CCP–positive cases:		
No SE	28 (8)	8 (2.3)
Any SE	206 (59.3)	105 (30.3)
Anti-CCP–negative cases:		
No SE	20 (24)	5 (6)
Any SE	41 (49)	17 (21)
Controls:		
No SE	348 (48)	69 (9)
Any SE	267 (37)	47 (6)
Leiden EAC:		
Anti-CCP–positive cases:		
No SE	28 (14)	9 (4)
Any SE	122 (59)	48 (23)
Anti-CCP–negative cases:		
No SE	61 (39)	13 (8)
Any SE	70 (45)	13 (8)
Controls:		
No SE	413 (47)	83 (9)
Any SE	316 (36)	69 (8)

We also investigated interaction between the HLA-DRB1 SE and the *PTPN22* R620W A alleles, using formulas describing multiplicative interactions (table 4). We did not find evidence of a significant multiplicative interaction between the susceptibility alleles in any of the studies analyzed separately. However, when we tested the pooled material for evidence of a positive interaction term, we found evidence of multiplicative interaction ($P = .025$) (table 4). We also found evidence of significant gene-gene interaction by using the LD statistic, which is a measure of deviance from independent penetrance of two unlinked loci. We obtained significant P values for the pooled material ($P = .027$) (table 4) and the EIRA ($P = .02$) and NARAC ($P = .035$) studies, but not for the Leiden EAC study ($P = .76$).

The OR for the presence of the R620W A allele and two HLA-DRB1 SE alleles was 25.7 (95% CI 17.0–38.9) when compared with those lacking both HLA-DRB1 SE alleles and the *PTPN22* R620W A allele (all studies combined) (table 5); the corresponding AP was 0.5 (CI 0.2–0.7). We did not find any statistically significant increased ORs of anti-CCP–negative RA regarding the *PTPN22* R620W allele or HLA-DRB1 SE alleles in the EIRA or the Leiden EAC studies.

Interaction between PTPN22 *and Smoking, Regarding Risk of Anti-CCP–Positive and Anti-CCP–Negative RA*

A significant interaction (significantly increased AP) has previously been demonstrated between HLA-DRB1 SE and smoking, concerning the risk to develop anti-CCP–positive RA.[7] When we performed the same analysis for interaction between the presence of different allelic forms of the *PTPN22* gene and smoking, no interaction between these two susceptibility factors was observed (data not shown) for either anti-CCP–positive or anti-CCP–negative RA.

Interaction among PTPN22*, HLA-DRB1 SE, and Smoking, Regarding Risk of Anti-CCP–Positive and Anti-CCP–Negative RA*

Finally, all three factors were added into an analysis of their combined effects as susceptibility factors for anti-CCP–positive RA and anti-CCP–negative RA in the EIRA study. As shown in table 6 and figure 1, a complex pattern of interactions was observed, with very high ORs for the A allele of *PTPN22* R620W in combination with HLA-DRB1 SE alleles and very high ORs also for smoking in combination with HLA-DRB1 SE alleles. Taken together, ORs between 20 and 25 were observed for combinations of HLA-DRB1 SE alleles, smoking, and the A allele of *PTPN22* R620W. No increased ORs for any of the risk factors or any of the combinations were seen for anti-CCP–negative RA (table 6 and fig. 1*B*). Sex-specific subanalysis was performed in all the different studies reported. Stratification by sex, however, demonstrated only marginal differences in observed effects (data not shown).

Discussion

The major finding reported in this article is the presence of an interaction between the two major genetic polymorphisms, HLA-DRB1 SE and *PTPN22* R620W, for developing anti-CCP–positive RA. An additional finding is the absence of an interaction between smoking and the *PTPN22* risk allele. A point for discussion raised in our report and illustrated by our findings is that different methods for quantifying interactions may yield quite different results and that different methods should be considered when analyzing interactions in complex diseases.

The current study uses data from three major case-control studies of different white populations—that is, the Swedish EIRA study, the North American NARAC study, and the Dutch Leiden EAC. In all three studies, HLA-DRB1 and *PTPN22* genetic polymorphisms were previously shown to be associated with an increased risk for anti-CCP–positive, but not anti-CCP–negative RA.[2,3,7,8] On the basis of these findings, we calculated the interaction between the genes in the two subsets of RA.

When the departure-from-additivity model was used, clear evidence of interaction was seen in all three studies analyzed separately, as well as when analysis was done on pooled data from all the studies. When the multiplicative

Table 4. ORs for Developing Anti-CCP–Positive RA According to Presence or Absence of Minor R620W *PTPN22* and HLA-DRB1 SE Alleles

Presence or Absence of Allele(s), by Study		No. of Cases/ Controls	OR	95% CI	AP (95% CI)	Deviation from		
						Additivity	Multiplicity	Independence of Penetrance
EIRA[a]								
R620W *PTPN22*	HLA-DRB1 SE				.5 (.3–0.7)	$P < .001$	$P = .06$	$P = .022$
None	None	75/284	1.0	...				
None	Any	416/330	5.0	3.7–6.7				
Any	None	31/95	1.2	.8–2.0				
Any	Any	199/84	9.9	6.8–14.3				
NARAC[b]								
R620W *PTPN22*	HLA-DRB1				.7 (.5–.9)	$P < .001$	$P = .05$	$P = .035$
None	None	28/348	1.0	...				
None	Any	206/267	9.3	6.0–14.3				
Any	None	8/69	1.5	.6–3.4				
Any	Any	105/47	30.2	17.6–51.9				
Leiden EAC[b]								
R620W *PTPN22*	HLA-DRB1 SE				.4 (.1–.7)	$P = .0016$	$P = .29$	$P = .76$
None	None	28/413	1.0	...				
None	Any	122/316	5.7	3.7–8.9				
Any	None	9/83	1.5	.7–3.3				
Any	Any	48/69	11.0	6.4–19.0				
All studies[c]								
R620W *PTPN22*	HLA-DRB1 SE				.5 (.4–.6)	$P < .001$	$P = .025$	$P = .027$
None	None	131/1,045	1.0	...				
None	Any	744/913	6.1	4.9–7.5				
Any	None	48/247	1.4	1.0–2.1				
Any	Any	352/200	13.2	10.2–17.2				

NOTE.—Tests for interaction between these susceptible alleles according to three different methods. Results are displayed for the EIRA, NARAC, and Leiden EAC studies separately, as well as pooled.

[a] OR adjusted for age, sex, and living area (only EIRA).

[b] OR adjusted for age and sex.

[c] OR adjusted for age, sex, and study.

model was used, evidence of a gene-gene interaction was seen only when the three materials were pooled but not when they were analyzed separately, suggesting a lack of power to detect epistasis in each study separately. When the LD method was used, the pooled analysis as well as the separate analyses on EIRA and NARAC gave evidence of interaction. Taken together, these data thus provide strong evidence of the existence of a gene-gene interaction between HLA-DRB1 SE alleles and the *PTPN22* R620W allele in increasing the risk for the development of anti-citrulline antibody–positive RA.

The concept of interaction was discussed by Rothman in connection to the introduction of the component cause model.[11] Referring to the concept of this model, intended to explain causality, two risk factors can be independent if no pathway to disease requires the involvement of both risk factors. Alternatively, a biological interaction is considered to be present if there is at least one pathway toward disease that requires the involvement of both risk factors. Interaction in this sense is detected by comparing the effect among those exposed to both factors with the sum of the effects attributed to each of the two factors per se. Thus, "interaction," as defined by a deviation from additivity, has a biological interpretation. When two contributory causes are interacting, there exists a pathway to disease requiring both factors. Thus, such interactions are relevant from a mechanistic perspective. Application of the previously described concept of additive interaction to the results from the present study would imply that having antigen-presenting cells with MHC class II molecules containing SE alleles (component cause one) and T-cells with dysfunctional down-regulation associated with the *PTPN22* R620W allele (component cause two) would increase disease manifestation more than the expected sum of each separate cause. This interaction is seen entirely among the anti-CCP–positive cases. The observed interaction thus indicates the existence of a disease mechanism for anti-CCP positive RA that requires the simultaneous presence of the HLA-DRB1 SE alleles and the *PTPN22* R620W allele.

An alternative common method for quantifying interactions is based on the calculation of the two risk factors' product term in a logistic-regression model (multiplicative

Table 5. ORs for Developing Anti-CCP–Positive RA According to Presence or Absence of Minor R620W *PTPN22* Allele and HLA-DRB1 SE Alleles

Presence or Absence of Allele(s), by Study		No. of Cases/ Controls	OR	95% CI	AP (95% CI)	
					Single	Double
EIRA[a]						
R620W *PTPN22*	HLA-DRB1 SE				.5 (.3–.7)	.4 (.0–.8)
None	None	75/284	1.0	...		
None	Single	249/269	3.6	2.7–5.0		
None	Double	167/61	11.3	7.6–16.9		
Any	None	31/95	1.2	.8–2.0		
Any	Single	126/67	7.9	5.3–11.8		
Any	Double	73/17	18.1	9.9–33.0		
NARAC[b]						
R620W *PTPN22*	HLA-DRB1 SE				.7 (.5–.9)	.6 (.3–.9)
None	None	28/348	1.0	...		
None	Single	110/224	5.8	3.6–9.1		
None	Double	96/43	28.3	16.4–49.1		
Any	None	8/69	1.5	.6–3.4		
Any	Single	57/38	20.0	11.1–36.2		
Any	Double	48/9	73.7	31.3–173.5		
Leiden EAC[b]						
R620W *PTPN22*	HLA-DRB1 SE				.4 (.1–.7)	.5 (−.04–1.0)
None	None	28/413	1.0	...		
None	Single	92/263	5.1	3.2–8.0		
None	Double	30/53	8.7	4.7–15.9		
Any	None	9/83	1.5	.7–3.3		
Any	Single	37/57	9.8	5.5–17.5		
Any	Double	11/12	17.4	6.7–44.0		
All studies[c]						
R620W *PTPN22*	HLA-DRB1 SE				.5 (.4–.7)	.5 (.2–.7)
None	None	131/1,045	1.0	...		
None	Single	451/756	4.5	3.6–5.6		
None	Double	293/157	13.7	10.4–18.0		
Any	None	48/247	1.4	1.0–2.1		
Any	Single	220/162	10.3	7.8–13.6		
Any	Double	132/38	25.7	17.0–38.9		

NOTE.—Tests for interaction between these susceptible alleles according to three different methods. Results are displayed for the EIRA, NARAC, and Leiden EAC studies separately, as well as pooled.

[a] OR adjusted for sex, age, and living area.

[b] OR adjusted for sex and age.

[c] OR adjusted for sex, age, and study.

interaction). In relation to the notion of interaction given above—that is, the existence of at least one pathway toward disease that requires the involvement of both risk factors—interaction may be detected by means of a departure from the multiplicative model as well, though this detection will have a lower sensitivity and higher specificity in relation to the corresponding detection by means of departure from additivity.

Both the additive and multiplicative models offer the possibility to have additional terms in the model to adjust for other factors (e.g., environmental factors). The ability to adjust for additional factors may thus make these models more usable than the currently used LD method.

These analyses have been based on an analysis of three different patient collections; therefore, it is important to consider the differences among them, particularly with regard to the appropriateness of combining the data for some analyses. The case-control study (EIRA) that provides one basis for this and previous[5–7,14,19] investigations of risk factors of RA was designed to capture all incident cases of RA in a certain geographical region of Sweden and to use age- and sex-matched controls randomly selected from the population in the same area. The high recruitment rate of both cases and controls[14] suggests that selection bias in both cases and controls is small and, thus, that results well represent true values for a Swedish population with early RA. The cases with RA in the NARAC study are prevalent cases who were recruited from various U.S. hospitals after

Table 6. ORs for Developing Anti-CCP–Positive or Anti-CCP–Negative RA According to Presence or Absence of Minor R620W *PTPN22* Allele, Presence or Absence of HLA-DRB1 SE Alleles, and Being an Ever or Never Smoker (Based on the EIRA Study)

Presence or Absence of Allele(s) and Smoking Status, by Anti-CCP Status			No. of Cases/ Controls	OR[a]	95% CI
Anti-CCP Positive					
R620W *PTPN22*	HLA-DRB1 SE	Smoking			
None	None	Never	25/106	1.0	...
Any	None	Never	10/31	1.3	.6–3.2
None	Single	Never	68/111	2.7	1.6–4.7
Any	Single	Never	37/27	6.5	3.3–12.8
None	Double	Never	38/33	5.1	2.7–9.8
Any	Double	Never	23/6	18.1	6.5–50.1
None	None	Ever	50/178	1.3	.7–2.2
Any	None	Ever	21/64	1.5	.8–2.9
None	Single	Ever	181/158	5.4	3.3–8.9
Any	Single	Ever	89/40	11.3	6.2–20.5
None	Double	Ever	129/28	23.6	12.8–43.8
Any	Double	Ever	50/11	23.4	10.4–52.4
Anti-CCP Negative					
R620W *PTPN22*	HLA-DRB1 SE	Smoking			
None	None	Never	65/109	1.0	...
Any	None	Never	21/31	1.0	.5–1.9
None	Single	Never	71/111	1.0	.6–1.5
Any	Single	Never	16/27	.9	.5–1.9
None	Double	Never	20/33	.9	.5–1.8
Any	Double	Never	5/6	1.4	.4–4.8
None	None	Ever	77/178	.7	.5–1.1
Any	None	Ever	38/64	1.0	.6–1.6
None	Single	Ever	96/158	1.0	.7–1.5
Any	Single	Ever	27/40	1.3	.7–2.3
None	Double	Ever	19/28	1.1	.6–2.2
Any	Double	Ever	7/11	1.1	.4–3.2

[a] OR adjusted for sex, age, and living area.

their diagnosis had been verified by a rheumatology specialist. Cases, who were of white origin, were selected on the basis of the severity of their disease, which is associated with presence of HLA-DRB1 SE alleles, which could explain why we observe higher ORs in the NARAC study. The controls were recruited from a healthy population in New York during a cancer surveillance program, and individuals were selected to match RA cases for age, sex, and ethnicity. The RA cases in the EAC study are incident cases recruited from the Leiden area, as described elsewhere,[17] and the diagnosis was confirmed by a rheumatologist. The controls were recruited from a matched hospitalized group that included a high proportion of subjects with venous thrombosis.

The study design in EIRA allows us to interpret the ORs as relative risks, which is not the case for NARAC or Leiden EAC. The formula for calculating the AP contains relative risk; hence, the estimated APs for the NARAC and the Leiden EAC might be less accurate than those for the EIRA.

Concerning the interactions between smoking and the HLA-DRB1 and *PTPN22* genes, we restricted our analysis to the EIRA material, since, in NARAC, the smoking information was collected differently among cases (prevalent cases with a long history of disease) and controls (from another study). Also, since smoking was associated with deep venous thrombosis in the Leiden EAC and a substantial part of the Leiden controls suffered from it, inclusion of this sample would have led to biased estimates.

In contrast to the previous demonstration of an interaction between smoking and HLA-DRB1 SE in EIRA,[6,7] we found no interaction (neither with the additive nor multiplicative models) between smoking and presence of the *PTPN22* R620W A allele. Further, none of these three susceptibility factors and neither of their combinations provided any significantly increased risk for anti-CCP–negative RA in the EIRA material.

The two genes discussed in the present report do both have functions related to activation of the adaptive immune system. The main function of HLA-DR antigens is to present antigens to T lymphocytes, whereas the protein tyrosine phosphatase coded for by the *PTPN22* gene appears to have important functions in setting the threshold for T-cell activation,[24] although other functions have also been implicated for this enzyme.[25] The fact that both these genes are associated almost exclusively with the anti-CCP–positive subset of RA, and the fact that the *PTPN22* R620W A allele appears to contribute a risk mainly when HLA-

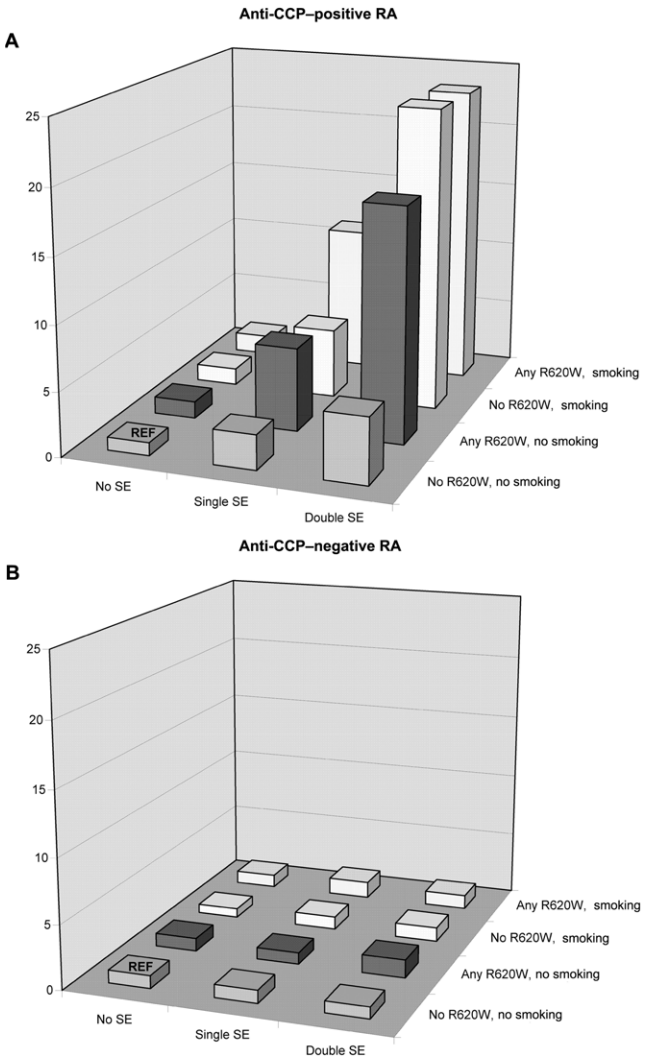

Figure 1. Histograms on ORs for developing anti-CCP–positive (*A*) and anti-CCP–negative (*B*) RA for different combinations of absence or presence of R620W *PTPN22,* single or double HLA-DRB1 SE alleles, and smoking (based on the EIRA study).

DRB1 SE alleles are present, thus provide strong additional evidence that MHC class II–dependent T-cell activation is of central pathogenetic importance for the subset of RA characterized by presence of anti-CCP antibodies. An etiologic hypothesis has recently been proposed that may explain the strong interaction between smoking and HLA-DRB1 SE alleles—the proposition being that smoking may contribute to citrullination of proteins in the lung and that immune activation against such posttranslationally modified proteins may occur preferentially in individuals carrying HLA-DRB1 SE alleles, since citrullination may specifically increase the affinity between a protein and an SE-containing HLA-DR β chain.[26] This hypothesis was made even more attractive by the recent demonstrations that citrullination of self-antigens may make them more immunogenic and arthritogenic[9] and that transfer of antibodies to citrullinated fibrinogen enhances the development of antibody-transferred arthritis in mice.[27] The absence of interaction between smoking and presence of the *PTPN22* R620W A allele adds to the significance of the previous finding of SE-smoking interaction, since it is now shown that interactions with smoking is specific to HLA-DRB1 SE and is not seen for other genes related to immune activation, at least not for *PTPN22*.

In a more general sense, the present and previous studies of gene-gene and gene-environment interactions in RA show how a given genetic polymorphism—such as the R620W allele of *PTPN22,* which alone provides a moderate OR thought to be typical for many risk genes in complex diseases—can exert much more significant effects in combination with other risk factors—here, the HLA-DRB1 SE alleles. The data also emphasize how an environmental factor can interact with susceptibility genes in different subsets of a disease. Our analysis also illustrates how interaction between genetic variations, as well as environmental risk factors, needs to be investigated in the context of genetic and environmental risk factors described elsewhere, to understand the complexity of complex diseases.

Acknowledgments

The EIRA study was supported by grants from the Swedish Medical Research Council, the Swedish Council for Working Life and Social Research, King Gustaf V's 80-year foundation, the Swedish Rheumatic Foundation, the Stockholm County Council, the insurance company Arbetsmarknadens Försäkringsaktiebolag, and the County of Sörmland Research and Development Center. We thank all participating controls and patients with RA and the following people who recruited patients for EIRA: Ingeli Andréasson, Landvetter; Eva Baecklund, Akademiska Hospital; Ann Bengtsson and Thomas Skogh, Linköping Hospital; Birgitta Nordmark, Johan Bratt, and Ingiäld Hafström, Karolinska University Hospital; Kjell Huddénius, Rheumatology Clinic (Stockholm); Shirani Jayawardene, Bollnäs Hospital; Ann Knight, Hudiksvall Hospital and Uppsala University Hospital; Ido Leden, Kristianstad Hospital; Göran Lindahl, Danderyd Hospital; Bengt Lindell, Kalmar Hospital; Christin Lindström and Gun Sandahl, Sophiahemmet; Björn Löfström, Katrineholm Hospital; Ingmar Petersson, Spenshult Hospital; Christoffer Schaufelberger, Sahlgrenska University Hospital; Patrik Stolt, Västerås Hospital; Berit Sverdrup, Eskilstuna Hospital; Olle Svernell, Västervik Hospital; and Tomas Weitoft, Gävle Hospital. For excellent data collection, we thank Marie-Louise Serra and Lena Nise, who provided invaluable contributions to the collection of data and maintenance of the database. The NARAC is supported by National Institutes of Health grants RO1 AR44422 and NO1-AR-2–2263 and by the National Arthritis Foundation. In addition to the original recruiting investigators, we thank Marlena Kern, Annette Lee, Robert Lundsten, Lindsey Criswell, Michael Seldin, Christopher Amos, Elaine Remmers, and Daniel Kastner, for their major ongoing contributions to the data management and analysis of this collection. The NARAC is indebted to the patients enrolled in this study, for their generosity and continuing participation. The Leiden EAC is supported by The Dutch National Arthritis Foundation (Nationaal Reumafonds) and the AutoCure organization. We also would

like to thank Prof. Dr. F. R. Rosendaal, Department of Epidemiology, Leiden University Medical Centre, for providing controls.

Web Resource

The URL for data presented herein is as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for RA, HLA-DRB1, and *PTPN22*)

References

1. Gregersen PK, Silver J, Winchester RJ (1987) The shared epitope hypothesis. Arthritis Rheum 30:1205–1213
2. Huizinga TWJ, Amos CI, van der Helm-van Mil AHM, Chen W, van Gaalen FA, Jawaheer D, Schreuder GMT, Wener M, Breedvald FC, Ahmad N, et al (2005) Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum 52:3433–3438
3. Plenge RM, Padyukov L, Remmers EF, Purcell S, Lee AT, Karlson EW, Wolfe F, Kastner DL, Alfredsson L, Altshuler D, et al (2005) Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with *PTPN22, CTLA4,* and *PADI4.* Am J Hum Genet 77:1044–1060
4. Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC, Ardlie KG, Huang Q, Smith AM, Spoerke JM, et al (2004) A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (*PTPN22*) is associated with rheumatoid arthritis. Am J Hum Genet 75:330–337
5. Stolt P, Bengtsson C, Nordmark B, Lindblad S, Lundberg I, Klareskog L, Alfredsson L (2003) Quantification of the influence of cigarette smoking on rheumatoid arthritis: results from a population based case-control study, using incident cases. Ann Rheum Dis 62:835–841
6. Padyukov L, Silva C, Stolt P, Alfredsson L, Klareskog L (2004) A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. Arthritis Rheum 50:3085–3092
7. Klareskog L, Stolt P, Lundberg K, Källberg H, Bengtsson C, Grunewald J, Ronnelid J, Harris HE, Ulfgren AK, Rantapaa-Dahlqvist S, et al (2006) A new model for an etiology of RA: smoking may trigger HLA-DR (SE)-restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum 54:38–46
8. Linn-Rasker SP, van der Helm-van Mil AH, Van Gaalen FA, Kloppenburg M, de Vries R, le Cessie S, Breedveld FC, Toes RE, Huizinga TW (2006) Smoking is a risk factor for anti-CCP antibodies only in rheumatoid arthritis patients that carry HLA-DRB1 shared epitope alleles. Ann Rheum Dis 65:366–371
9. Lundberg K, Nijenhuis S, Vossenaar E, Palmblad K, van Venrooij WJ, Klareskog L, Zendman AJ, Harris HE (2005) Citrullinated proteins have increased immunogenicity and arthritogenicity and their presence in arthritic joints correlates with disease severity. Arthritis Res Ther 7:R458–R467
10. Ahlbom A, Alfredsson L (2005) Interaction: a word with two meanings creates confusion. Eur J Epidemiol 20:563–564
11. Rothman KJ (2002) Epidemiology: an introduction. Oxford University Press, New York
12. Andersson T, Alfredsson L, Kallberg H, Zdravkovic S, Ahlbom A (2005) Calculating measures of biological interaction. Eur J Epidemiol 20:575–579
13. Zhao J, Jin L, Xiong M (2006) Test for interaction between two unlinked loci. Am J Hum Genet 79:831–845
14. Bengtsson C, Nordmark B, Klareskog L, Lundberg I, Alfredsson L, EIRA Study Group (2005) Socioeconomic status and the risk of developing rheumatoid arthritis: results from the Swedish EIRA study. Ann Rheum Dis 64:1588–1594
15. Jawaheer D, Seldin MF, Amos CI, Chen WV, Shigeta R, Monteiro J, Kern M, Criswell LA, Albani S, Nelson JL, et al (2001) A genomewide screen in multiplex rheumatoid arthritis families suggests genetic overlap with other autoimmune diseases. Am J Hum Genet 68:927–936
16. Jawaheer D, Lum RF, Amos CI, Gregersen PK, Criswell LA (2004) Clustering of disease features within 512 multicase rheumatoid arthritis families. Arthritis Rheum 50:736–741
17. Van Aken J, van Bilsen JH, Allaart CF, Huizinga TW, Breedveld FC (2003) The Leiden Early Arthritis Clinic. Clin Exp Rheumatol 21:S100–S105
18. Arnett FC, Eworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, Healey LA, Kaplan SR, Liang MH, Luthra HS, et al (1988) The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 31:315–324
19. Stolt P, Kallberg H, Lundberg I, Sjogren B, Klareskog L, Alfredsson L (2005) Silica exposure is associated with increased risk of developing rheumatoid arthritis: results from the Swedish EIRA-study. Ann Rheum Dis 64:582–586
20. van der Meer FJ, Koster T, Vandenbroucke JP, Briet E (1997) The Leiden Thrombophilia Study (LETS). Thromb Haemost 78:631–635
21. Rönnelid J, Wick MC, Lampa J, Lindblad S, Nordmark B, Klareskog L, van Vollenhoven RF (2005) Longitudinal analysis of anti-citrullinated antibodies (anti-CP) during 5 years follow-up in early rheumatoid arthritis: anti-CP status is a stable phenotype that predicts worse disease activity and greater radiological progression. Ann Rheum Dis 64:1744–1749
22. Olerup O, Zetterquist H (1992) HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: an alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation. Tissue Antigens 39:225–235
23. Hosmer DW, Lemeshow S (1992) Confidence interval estimation of interaction. Epidemiology 3:452–456
24. Lee AT, Li W, Liew A, Bombardier C, Weisman M, Massarotti EM, Kent J, Wolfe F, Begovich AB, Gregersen PK (2005) The PTPN22 R620W polymorphism associates with RF+ rheumatoid arthritis in a dose-dependent manner but not with HLA-SE status. Genes Immun 6:129–133
25. Gregersen PK, Batliwalla F (2005) *PTPN22* and rheumatoid arthritis: gratifying replication. Arthritis Rheum 52:1952–1955
26. Hill JA, Southwood S, Sette A, Jevnikar AM, Bell DA, Cairns E (2003) Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis associated HLA-DRB1*0401 MHC class II molecule. J Immunol 171:538–541
27. Hill J, Wehrli B, Jevnikar AM, Bell DA, Cairns E (2003) Citrullinated fibrinogen induces arthritis in HLA–DRB1*0401 transgenic mice. Arthritis Rheum 48 Suppl 9:S348