Genetic polymorphisms in PTPN22, PADI-4, and CTLA-4 and risk for rheumatoid arthritis in two longitudinal cohort studies: evidence of gene-environment interactions with heavy cigarette smoking

Citation
Costenbader, Karen H., Shun-Chiao Chang, Immaculata De Vivo, Robert Plenge, and Elizabeth W. Karlson. 2008. Genetic polymorphisms in PTPN22, PADI-4, and CTLA-4 and risk for rheumatoid arthritis in two longitudinal cohort studies: evidence of gene-environment interactions with heavy cigarette smoking. Arthritis Research & Therapy 10(3): R52.

Published Version
doi:10.1186/ar2421

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:4891670

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
Research article

Genetic polymorphisms in PTPN22, PADI-4, and CTLA-4 and risk for rheumatoid arthritis in two longitudinal cohort studies: evidence of gene-environment interactions with heavy cigarette smoking

Karen H Costenbader¹, Shun-Chiao Chang²,³, Immaculata De Vivo²,³, Robert Plenge¹ and Elizabeth W Karlson¹

¹Division of Rheumatology, Immunology, and Allergy, Department of Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA
²Channing Laboratory, Department of Medicine, Brigham and Women’s Hospital, Harvard Medical School, 75 Francis Street Boston, Massachusetts 02115, USA
³Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, Massachusetts 02115, USA

Corresponding author: Karen H Costenbader, kcostenbader@partners.org

Received: 17 Sep 2007 Revisions requested: 14 Nov 2007 Revisions received: 27 Mar 2008 Accepted: 7 May 2008 Published: 7 May 2008

Arthritis Research & Therapy 2008, 10:R52 (doi:10.1186/ar2421)

Abstract

Introduction PTPN22, PADI-4, and CTLA-4 have been associated with risk for rheumatoid arthritis (RA). We investigated whether polymorphisms in these genes were associated with RA in Caucasian women included in two large prospective cohorts, adjusting for confounding factors and testing for interactions with smoking.

Methods We studied RA risk associated with PTPN22 (rs2476601), PADI-4 (rs2240340), and CTLA-4 (rs3087243) in the Nurses’ Health Study (NHS) and NHSII. Participants in NHS were aged 30 to 55 years at entry in 1976; those in NHSII were aged 25 to 42 years at entry in 1989. We confirmed incident RA cases through to 2002 in NHS and to 2003 in NHSII by questionnaire and medical record review. We excluded reports not confirmed as RA. In a nested case-control design involving participants for whom there were samples for genetic analyses (45% of NHS and 25% of NHSII), each incident RA case was matched to a participant without RA by year of birth, menopausal status, and postmenopausal hormone use. Genotyping was performed using Taqman single nucleotide polymorphism allelic discrimination on the ABI 7900 HT (Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA 94404 USA) with published primers. Human leukocyte antigen shared epitope (HLA-SE) genotyping was performed at high resolution. We employed conditional logistic regression analyses, adjusting for smoking and reproductive factors. We tested for additive and multiplicative interactions between each genotype and smoking.

Results A total of 437 incident RA cases were matched to healthy female control individuals. Mean (± standard deviation) age at RA diagnosis was 55 (± 10), 57% of RA cases were rheumatoid factor (RF) positive, and 31% had radiographic erosions at diagnosis. PTPN22 was associated with increased RA risk (pooled odds ratio in multivariable dominant model = 1.46, 95% confidence interval [CI] = 1.02 to 2.08). The risk was stronger for RF-positive than for RF-negative RA. A significant multiplicative interaction between PTPN22 and smoking for more than 10 pack-years was observed (P = 0.04). CTLA-4 and PADI-4 genotypes were not associated with RA risk in the pooled results (pooled odds ratios in multivariable dominant models: 1.27 [95% CI = 0.88 to 1.84] for CTLA-4 and 1.04 [95% CI = 0.77 to 1.40] for PADI-4). No gene-gene interaction was observed between PTPN22 and HLA-SE.

Conclusion After adjusting for smoking and reproductive factors, PTPN22 was associated with RA risk among Caucasian women in these cohorts. We found both additive and multiplicative interactions between PTPN22 and heavy cigarette smoking.

ACR = American College of Rheumatology; CCP = cyclic citrullinated peptide; CI = confidence interval; HLA = human leukocyte antigen; NHS = Nurses’ Health Study; OR = odds ratio; RA = rheumatoid arthritis; RF = rheumatoid factor; SE = shared epitope; SNP = single nucleotide polymorphism.
Introduction
Rheumatoid arthritis (RA) is a systemic autoimmune disease that is characterized by chronic, destructive, and debilitating arthritis. The most common inflammatory arthritis, it affects approximately 1% of the population [1]. The etiology of RA is unknown, but it is presumed that environmental factors trigger its development in the genetically predisposed. The strongest genetic risk factor, which is responsible for approximately 30% of the genetic contribution to the development of RA, is human leukocyte antigen (HLA) type [2,3]. The HLA-shared epitope (SE) is the strongest known genetic risk factor for RA, with two copies associated with a relative risk of 5 to 6 in Caucasians [4-6]. However, RA, like many complex human diseases, is polygenic in origin. It is likely that many other genes, both inside and outside the HLA region, contribute to disease predisposition [5,7-10]. RA genetic studies have reported polymorphisms PTPN22 (R620W, rs2476601) [9,11-16], CTLA-4 (CT60, rs3087243) [17,18], and PADI-4 (PADI4_94, rs2240340) [8] to be associated with increased risk for RA. More recently, polymorphisms in or near the genes encoding STA74 [19], Traf1-C5 [20], TNF-AIP2 [16], and IL2RA [16] have been reported in whole-genome scans. The non-HLA genetic polymorphism that has been most strongly replicated across multiple independent studies is PTPN22 [9,11-14,16,21-24]. This missense allele (C→T) has been associated in past studies with an odds ratio (OR) for RA of approximately 1.8 [9], and it appears to carry greater risk for autoantibody-positive RA [12,24,25]. Homozygotes for the variant T allele are at greatest risk for RA (OR = 4.6) [12]. This polymorphism has also been associated with increased risk for type 1 diabetes [26,27] and systemic lupus erythematosus [11,13]. The associations of CTLA-4 and PADI-4 polymorphisms with RA risk have been less well replicated [17,18,28].

In a large pooled replication study, CTLA4 CT60 polymorphism was associated with anti-cyclic citrullinated peptide antibody (anti-CCP)-positive RA [25].

Twin studies conducted in the UK and Finland have estimated that 50% to 60% of the variation in RA susceptibility is accounted for by genetic factors [29], leaving 40% to 50% probably due to environmental exposures. Cigarette smoking is the best established environmental risk factor for RA, with risk increasing in proportion to duration and intensity of exposure [30-35]. Case-control studies conducted in Sweden, Holland, and North America have identified an interaction between presence of the HLA-SE alleles and cigarette smoking in determining RA risk, in particular that of anti-CCP-positive RA [36-38]. Female reproductive factors such as early age at menarche, irregular menses, and use of postmenopausal hormones have also been related to increased RA risk, and prolonged duration of breast-feeding was found to be protective against development of RA in the Nurses’ Health Study (NHS) [39-41].

We aimed to validate previous findings of increased risk for RA associated with polymorphisms in the PTPN22, PADI-4 and CTLA-4 genes, and to assess whether of behavioral and reproductive factors that are known to be associated with RA risk influence these findings. We also investigated potential additive and multiplicative interactions between each of these polymorphisms and the presence of the HLA-SE. To do this, we conducted a case-control study nested within the NHS and NHSII; those studies include two large cohorts of women, who were followed closely over many years for behavioral and reproductive factors before the onset of disease.

Materials and methods
Study population
The NHS includes a prospective cohort of 121,700 female nurses, aged 30 to 55 years in 1976 when the study began. The NHSII was established in 1989, when 116,608 female nurses aged 25 to 42 years completed a baseline questionnaire about their medical histories and lifestyles. Ninety-four per cent of the NHS participants from 1976 to 2002, and 95% of NHSII participants from 1989 to 2003 have remained in active follow up (5% to 6% no longer respond to questionnaires and have not been confirmed as dead). All aspects of this study were approved by the Partners’ HealthCare Institutional Review Board.

Identification of rheumatoid arthritis
As previously described [35], we employed a two-stage procedure in which all nurses who self-reported any connective tissue disease received a screening questionnaire for connective tissue disease symptoms [42], and – if positive – a detailed medical record review for American College of Rheumatology (ACR) classification criteria for RA [43], in order to identify and validate incident cases of RA. The presence or absence of rheumatoid factor (RF) and other features of RA was based on medical record review. Those in whom four of the seven ACR criteria were documented in the medical record were considered to have definite RA. For this nested case-control study, we also included a small number of women (n = 14) with three documented ACR criteria for RA, a diagnosis of RA by their physician, and agreement by two rheumatologists on the diagnosis of RA.

Population for analysis
We excluded prevalent RA cases diagnosed before the cohort was assembled, nonresponders, and women who reported any connective tissue disease that was not subsequently confirmed to be RA by medical record review. Women were censored when they failed to respond to any subsequent biennial questionnaires. Among the women in each cohort who had provided a sample for genetic analyses, each participant with confirmed incident RA was matched by year of birth, menopausal status, and postmenopausal hormone use to a healthy woman in the same cohort without RA. To minimize population stratification, and given that most cohort participants are Cau-
casian, we limited the analyses to Caucasian matched pairs of women. In 1992 (NHS) and 1989 (NHSII), all participants were asked to provide data concerning their own racial backgrounds in more detailed categories. Of the Caucasian women in NHS and NHSII included in this analysis, 2% reported pure Scandinavian heritage, 15% reported pure Southern European, and 83% reported other or mixed Caucasian backgrounds. There were no significant differences in the distributions of these ethnicities between cases and controls ($\chi^2$ with two degrees of freedom, $P = 0.30$).

**Blood sampling**

From 1989 to 1990, 32,826 (27%) NHS participants aged 43 to 70 years agreed to provide blood samples for future NHS studies. Between 1996 and 1999, 29,613 (25%) of the women included in the NHSII cohort (aged 32 to 52 years at that time) also agreed to have their blood drawn for future investigations. All samples were collected in heparinized tubes and sent to us by overnight courier in chilled containers. On receipt, the blood samples were centrifuged, aliquoted, and stored in liquid nitrogen freezers at -70°F (-57°C). The demographic and exposure characteristics of the NHS and NHSII participants who provided blood samples were found to be very similar to those of the overall cohorts [44,45].

**DNA extraction from blood**

DNA was extracted from buffy coats from 96 samples in 3 to 4 hours. A volume of 50 µl of buffy coat was diluted with 150 µl phosphate-buffered saline and processed using the QiAmp™ (QIAGEN Inc., Chatsworth, CA, USA) 96-spin blood kit protocol. The protocol entails adding protease, the sample, and lysis buffer to 96-well plates. The plates are then mixed and incubated at 158°F (70°C), before adding ethanol and transferring the samples to columned plates. The columned plates are then centrifuged and washed with buffer. Adding elution buffer and centrifuging elutes the DNA. The average yield from 50 µl of buffy coat (based on 1,000 samples) is 5.5 µg with a standard deviation of 2.2 (range 2.0 to 16.4). These methods are semiautomated using a Qiagen 8000 robot to increase throughput and decrease manual pipetting errors.

**Buccal cell collection method and DNA extraction in NHS**

Forty thousand women in NHS who did not give blood in 1989 to 1990 were asked to give a buccal cell sample in 2002. To date we have collected an additional 21,733 buccal cell samples (18% of the NHS cohort). A collection kit was sent to participants, consisting of instructions for the buccal cell collection and the necessary supplies (a small bottle of mouthwash, a plastic cup with a screwtop cap, a ziplock plastic bag and absorbent sheet, and a stamped, self-addressed bubble envelope), as well as an informed consent form. Participants were instructed to fill the cup with mouthwash, swish the mouthwash in their mouth vigorously, and then spit back into the cup. Returned samples were processed using ReturPure-Gene DNA Isolation Kit (Gentra Systems, Minneapolis, MN, USA) to extract genomic DNA from human cheek cells. The extracted DNA was archived in liquid nitrogen freezers using specific tracking software. The average DNA recovery from these specimens measured using PicoGreen was 59 ng/µl.

**Whole-genome amplification**

For all genomic DNA samples, an aliquot was put through a whole-genome amplification protocol using the GenomPhi DNA amplification kit (GE Healthcare, Piscataway, NJ) to yield high-quality DNA sufficient for single nucleotide polymorphism (SNP) genotyping.

**Single nucleotide polymorphism genotyping**

DNA was genotyped using Taqman SNP allelic discrimination on the ABI 7900 HT (Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA 94404 USA) using published primers [8,9,18,46]. We studied only the CT4-4 CT60 (rs3087243) allele. We chose the PADI4_94 allele (rs2240340) of the haplotype first described by Suzuki and coworkers [8], because it had the strongest association in a Japanese population and was replicated in a large meta-analysis [25]. Using the same methods, we also genotyped the lactase gene (rs4988235), which is known to exhibit substantial variation in allele frequency from Northern to Southern Europe, in order to test for population stratification in this nested case-control study [47,48].

**HLA-DRB1 shared epitope determination**

Low-resolution HLA-DRB1 genotyping was performed by polymerase chain reaction with sequence specific primers using OLERUP SSP kits (QIAGEN, West Chester, PA, USA). We used primers to amplify DNA samples that contained sequences for HLA-DRB1*04, *01, *10 and *14, along with consensus primers and appropriate positive and negative control samples. For samples with positive two-digit HLA signals, sequence-specific primers were used for high resolution four-digit shared epitope allele detection of DRB1*0401, *0404, *0405, *0101, *0102, *1402, and *1001. OLERUP SSP computer software (QIAGEN) was used to determine four-digit HLA types. Quality control split samples were included, randomly interspersed with study samples.

**Covariate information**

Information was collected from the women in both cohorts via biennial questionnaires regarding diseases, lifestyle, and health practices. Age was updated in each cycle. Reproducible covariates were chosen based on our past findings of associations between reproductive factors and risk for developing RA in this cohort [41]. Data on parity, total duration of breast-feeding, menopausal status, and postmenopausal hormone use were selected from the questionnaire cycle before the date of RA diagnosis (or index date in controls). Self-reported menopausal status and age at menopause are highly reproducible in our cohorts; in a validation study of a subsample of NHS participants, 82% of naturally postmenopausal
women reported the same age at menopause (within 1 year) on two questionnaires mailed 2 years apart [49].

Participants in both cohorts were asked at baseline whether they were a current smoker or had ever smoked in the past and the age at which they began to smoke. Current smokers were asked for the number of cigarettes typically smoked per day and former smokers reported the age at which they stopped smoking and the number of cigarettes smoked per day before quitting. On each subsequent questionnaire, participants reported whether they currently smoked and the number of cigarettes smoked per day. From these reports, we calculated pack years of smoking (product of years of smoking and packs of cigarettes per day).

Other potential confounders examined included, body mass index, which was computed for each 2-year time interval using the most recent weight (in kilograms) divided by height (in meters squared), as reported at baseline. Alcohol intake was reported at least every 4 years and categorized in grams per day. Husband’s educational level was assessed in 1992 in NHS and 1999 in NHSII, and was included as a proxy for socioeconomic level.

Statistical analyses
We verified Hardy-Weinberg equilibrium for each of the genotypes among controls in each of the datasets (NHS blood, NHS cheek cells, and NHSII blood). We employed conditional logistic regression analyses, conditioned on matching factors, and adjusted for potential confounders, including cigarette smoking and reproductive factors assessed before diagnosis of RA. All analyses were first conducted separately in each cohort and then on data pooled from the two cohorts. Because the $P$ value for heterogeneity was significant for the $CTLA-4$ genotype, we also meta-analytically pooled results from the two cohorts using a DerSimonian and Laird random effects model [50]. In analyses stratified by the presence of RF among the RA cases, we employed unconditional logistic regression analyses, adjusting for each of the matching factors, in addition to the covariates above. For analyses of $PTPN22$, we employed a dominant model because the minor allele frequencies were low (9% in controls and 14% in cases). In analyses involving $CTLA-4$ and $PADI-4$, we assessed the risk for RA in dominant, additive, and recessive models.

Gene-environment and gene-gene interactions
We conducted assessments for gene-environment interactions by testing for both additive and multiplicative interactions. For additive interactions, we calculated the attributable proportion due to interaction using a $2 \times 2$ factorial design to analyze the data [51-53]. (There is evidence of interaction when the attributable proportion is not equal to 0.) Ninety-five per cent confidence intervals (CIs) were calculated using the delta method as described by Hosmer and Lemeshow [54].

We tested for multiplicative interaction using an interaction variable (for example, gene × smoking) in the conditional logistic regression models. The significance of the interaction was determined using the Wald $\chi^2$ test of the interaction variable. In the combined NHS-NHSII nested case-control study dataset, we assessed for interactions between the presence of each polymorphism and cigarette smoking categorized both as ever/never, and then dichotomized as $\leq 10$ or $\geq 10$ pack-years of smoking, because this is the threshold we previously identified to be associated with increased risk for RA [35]. Using similar methods, we tested for gene-gene interaction between $PTPN22$ and $HLA-SE$ in influencing RA susceptibility in analyses limited to NHS and NHSII blood samples. SAS version 9.1 (SAS Institute, Cary, NC, USA) was used for all analyses.

Results
A total of 437 pairs of Caucasian women, each containing one woman with incident RA and her matched control, were included in these analyses, after removing 18 women because of missing data for all genotypes examined. The characteristics of the RA cases at diagnosis in each of the two cohorts are shown in Table 1. The cases in the NHS had a mean (± standard deviation) age of 57 years (± 9), as compared with 43 (± 5) in the younger NHSII cohort, because of the different ages targeted for enrollment in the two cohorts. Otherwise, the cases were similar in terms of the prevalence of RF, erosions, nodules, and proportion diagnosed by a member of the ACR. All cases and controls in these analyses were Caucasian, and the mean (± standard deviation) number of ACR criteria for the classification of RA was 5 (± 1) [43].

Table 2 shows the characteristics of the RA cases and matched controls at the time of RA diagnosis (or index date for the controls). A higher proportion of RA cases and controls were postmenopausal at RA diagnosis in the NHS than in the NHSII cohort, but the proportions of premenopausal and postmenopausal women among cases and controls were similar in each of the cohorts, as were the proportions currently receiving postmenopausal hormones. In NHSII a slightly higher percentage of women with RA were parous as compared with their matched controls (94% and 86%), but this was not true in the NHS cohort (91% of RA cases and 95% of controls). Among women in the NHSII with RA, a higher proportion had husbands who were college educated as compared with their matched controls (39% compared to 18%), but this was not true in the NHS cohort (20% in each group). No significant differences in allele frequencies of the lactase gene (rs4988235) in cases compared with controls were found. This argues strongly against any significant population stratification in our samples.

The genotype and allele frequencies of the RA cases and controls for the three candidate genotypes are shown in Table 3. None of the $PTPN22$, $CTLA-4$, or $PADI-4$ genotype distribu-
tions deviated from Hardy-Weinberg equilibrium, either in each cohort or in the combined dataset. Overall, genotyping call rates were 97.5% for \textit{PTPN22}, 96.4% for \textit{CTLA-4}, 97.9% for \textit{PADI-4}, and 98.7% for \textit{HLA-SE}. The frequency of the T allele of the \textit{PTPN22} polymorphism was significantly higher among RA cases than among controls ($\chi^2$ with one degree of freedom, $P = 0.001$ for pooled NHS and NHSII cohorts). The mutant alleles were not statistically associated with RA case status for the other two genotypes, namely \textit{PADI-4} and \textit{CTLA-4}. As expected, \textit{HLA-SE} alleles were highly significantly associated with risk for RA. (A slighter higher frequency of NHS cheek cell DNA samples could not be HLA genotyped: 3% of cases and controls, as compared with 0% to 2% of NHS and NHSII case and control DNA samples from blood.)

Table 4 includes the results of conditional logistic regression analyses of risk for RA associated with each of the genotypes, performed separately in each cohort, and then on pooled data. The final multivariable model includes pack-years of cigarette smoking, age at menarche, regularity of menses, parity, and total duration of breast-feeding. Further adjustment for body mass index, alcohol intake, husband's educational level, and

**Table 1**

| Characteristic                        | NHS ($n = 388$) | NHSII ($n = 49$) |
|---------------------------------------|----------------|-----------------|
| Age at diagnosis (mean ± SD)          | 57 ± 9         | 43 ± 5          |
| Rheumatoid factor positive ($n$ [%])  | 223 (57%)      | 27 (55%)        |
| Rheumatoid nodules ($n$ [%])          | 51 (13%)       | 6 (12%)         |
| Radiographic changes ($n$ [%])        | 117 (30%)      | 17 (35%)        |
| Diagnosed by a member of ACR ($n$ [%])| 320 (84%)      | 47 (96%)        |
| Mean number of ACR criteria for RA [44], (SD) | 5 ± 1          | 5 ± 1           |

ACR, American College of Rheumatology; NHS, Nurses' Health Study; RA, rheumatoid arthritis; SD, standard deviation.

**Table 2**

| Characteristics | NHS (388 matched pairs) | NHSII (49 matched pairs) |
|-----------------|--------------------------|--------------------------|
| Matching factors| RA cases | Controls | RA cases | Controls |
| Age (years; mean ± SD) | 57 ± 9 | 57 ± 9 | 43 ± 5 | 43 ± 5 |
| Postmenopausal ($n$ [%]) | 263 (68%) | 255 (66%) | 14 (29%) | 14 (29%) |
| Current PMH use ($n$ [%]) | 111 (36%) | 101 (34%) | 11 (69%) | 12 (75%) |
| Other characteristics | RA cases | Controls | RA cases | Controls |
| Ever cigarette smokers ($n$ [%]) | 246 (63%) | 220 (57%) | 19 (39%) | 22 (45%) |
| Pack-years among smokers, mean (SD) | 24 ± 17 | 24 ± 20 | 16 ± 9 | 11 ± 6 |
| Parous ($n$ [%]) | 352 (91%) | 367 (95%) | 46 (94%) | 42 (86%) |
| Breastfed babies ≥ 12 months total, among parous ($n$ [%]) | 44 (13%) | 65 (18%) | 18 (43%) | 18 (45%) |
| Age at menarche <12 years ($n$ [%]) | 87 (22%) | 77 (20%) | 16 (33%) | 15 (31%) |
| Irregular menstrual cycle ($n$ [%]) | 55 (14%) | 49 (13%) | 9 (18%) | 11 (22%) |
| Husband education > college graduate ($n$ [%]) | 77 (20%) | 77 (20%) | 19 (39%) | 9 (18%) |
| Physical activity (hours/day mean ± SD) | 2 ± 3 | 2 ± 5 | 2 ± 3 | 3 ± 3 |
| Alcohol intake (g/day; mean ± SD) | 6 ± 10 | 6 ± 9 | 2 ± 1 | 2 ± 2 |
| Body mass index (kg/m²; mean ± SD) | 25 ± 4 | 26 ± 5 | 27 ± 7 | 26 ± 5 |

*Percentage is calculated among postmenopausal women or parous women, with unknown/missing group excluded. For the rest of variables, percentage was calculated with missing category included as a separate category. NHS, Nurses' Health Study; PMH, postmenopausal hormone; RA, rheumatoid arthritis; SD, standard deviation.*
Table 3

**Genotype and allele frequencies of the RA cases and controls for the three candidate genotypes (PTPN22, CTLA-4, and PADI-4) and genotype frequencies of HLA-SE**

| Gene       | Genotype/allele | NHS blood (219 matched pairs) | NHS cheek cell (169 matched pairs) | NHSII blood (49 matched pairs) |
|------------|-----------------|-------------------------------|-----------------------------------|--------------------------------|
|            |                 | Cases                        | Controls                          | Cases                        | Controls                          |
| PTPN22a    | Genotype        | Cases                        | Controls                          | Cases                        | Controls                          |
| CC         |                 | 160 (74%)                    | 178 (82%)                         | 125 (77%)                    | 131 (80%)                         | 37 (79%)                         | 44 (94%)                         |
| CT         |                 | 50 (23%)                     | 38 (18%)                          | 30 (18%)                     | 31 (19%)                          | 7 (15%)                          | 3 (6%)                           |
| TT         |                 | 5 (2%)                       | 1 (0.5%)                          | 8 (5%)                       | 1 (1%)                            | 3 (6%)                           | 0 (0%)                           |
| Missing    |                 | 4                            | 2                                | 6                            | 6                                | 2                                | 2                                |
| °P         |                 | 0.08                         | 0.07                             | 0.06                         |                                   |                                   |                                   |
| Allele     |                 | C 370 (86%)                  | 394 (91%)                         | 280 (86%)                    | 293 (90%)                         | 81 (86%)                         | 91 (97%)                         |
|            |                 | T 60 (14%)                   | 40 (9%)                           | 46 (14%)                     | 33 (10%)                          | 13 (14%)                         | 3 (3%)                           |
| °P         |                 | 0.03                         | 0.12                             | 0.009                        |                                   |                                   |                                   |
| CTLA-4b    | Genotype        | AA 35 (17%)                  | 46 (22%)                          | 40 (25%)                     | 29 (18%)                          | 7 (15%)                          | 12 (25%)                         |
|            |                 | AG 106 (50%)                 | 96 (45%)                          | 69 (42%)                     | 80 (50%)                          | 26 (54%)                         | 19 (40%)                         |
|            |                 | GG 71 (33%)                  | 70 (33%)                          | 54 (33%)                     | 51 (32%)                          | 15 (31%)                         | 17 (35%)                         |
| Missing    |                 | 7                            | 7                                | 6                            | 9                                | 1                                | 1                                |
| °P         |                 | 0.37                         | 0.27                             | 0.28                         |                                   |                                   |                                   |
| Allele     |                 | A 176 (42%)                  | 188 (44%)                         | 149 (46%)                    | 138 (43%)                         | 40 (42%)                         | 43 (45%)                         |
|            |                 | G 248 (58%)                  | 236 (56%)                         | 177 (54%)                    | 182 (57%)                         | 56 (58%)                         | 53 (55%)                         |
| °P         |                 | 0.41                         | 0.51                             | 0.66                         |                                   |                                   |                                   |
| PADI-4c    | Genotype        | GG 76 (35%)                  | 76 (36%)                          | 57 (35%)                     | 64 (39%)                          | 22 (45%)                         | 18 (38%)                         |
|            |                 | GA 107 (49%)                 | 101 (47%)                         | 68 (41%)                     | 73 (44%)                          | 16 (33%)                         | 21 (45%)                         |
|            |                 | AA 34 (16%)                  | 37 (17%)                          | 39 (24%)                     | 28 (17%)                          | 11 (22%)                         | 8 (17%)                          |
| Missing    |                 | 2                            | 5                                | 5                            | 4                                | 0                                | 2                                |
| °P         |                 | 0.87                         | 0.30                             | 0.47                         |                                   |                                   |                                   |
| Allele     |                 | G 259 (60%)                  | 253 (59%)                         | 182 (55%)                    | 201 (61%)                         | 60 (61%)                         | 57 (61%)                         |
|            |                 | A 175 (40%)                  | 175 (41%)                         | 146 (45%)                    | 129 (39%)                         | 38 (39%)                         | 37 (39%)                         |
| °P         |                 | 0.87                         | 0.16                             | 0.93                         |                                   |                                   |                                   |
| HLA-SEd    | 0 copy          | 84 (38%)                     | 122 (56%)                         | 99 (60%)                     | 114 (70%)                         | 23 (47%)                         | 30 (63%)                         |
|            | 1 copies        | 99 (45%)                     | 83 (38%)                          | 52 (32%)                     | 40 (24%)                          | 17 (35%)                         | 17 (35%)                         |
|            | 2 copies        | 36 (16%)                     | 14 (6%)                           | 13 (8%)                      | 10 (6%)                           | 9 (18%)                          | 1 (2%)                           |
| Missing    |                 | 0                            | 0                                | 5                            | 5                                | 0                                | 1                                |
| °P         |                 | 0.0001                       | 0.22                             | 0.03                         |                                   |                                   |                                   |

aUnable to genotype 12 cases and 10 controls. bUnable to genotype 14 cases and 17 controls. cUnable to genotype 7 cases and 11 controls. dUnable to genotype 5 cases and 5 controls. eP values estimated from the Fisher’s exact test. fP values estimated from the χ² test. NHS, Nurses’ Health Study; RA, rheumatoid arthritis. OK.
oral contraceptive use did not affect risk estimates either and these were not included in the final models. As is evident comparing the results of a model taking only the matching factors into account, adjustment for potential confounders did not significantly influence results. The risk for RA associated with the PTPN22 variant T allele was elevated (OR = 1.46 [95% CI = 1.02 to 2.08]) in a NHS and NHSII pooled multivariable dominant model. These results may have been influenced by the high OR observed in the smaller NHSII cohort (OR = 8.77).

The CTLA-4 variant G allele was associated with an increased RA risk among women in the NHS cohort (multivariable dominant model OR = 1.92 [95% CI = 1.10 to 3.37]), but not in the NHSII cohort or pooled results (multivariable dominant model OR = 1.27 [95% CI = 0.88 to 1.84]; the OR was similar [1.29, 95% CI = 0.54 to 3.08] in a random effects meta-analytically pooled dominant model). The PADI-4 genotype was not associated with risk for RA in the NHS and NHSII cohorts in any of the models.

To pursue potential associations of these polymorphisms with different RA phenotypes, we conducted analyses stratified by RF positivity, because many risk factors, including HLA-SE and cigarette smoking, have been shown to be more strongly associated with RF-seropositive RA [35,36]. Results of these analyses are shown in Table 5. The effect of the PTPN22 polymorphism was seen primarily for the development of RF-seropositive RA (OR = 1.75 [95% CI = 1.18 to 2.59]).

Cigarette smoking is a strong environmental risk factor for the development of RA, in particular RF-positive RA, and amount and duration are associated with increased risk [35]. We thus investigated potential interactions between the three polymorphisms of interest and the amount and duration of cigarette smoking.
smoking at the time of RA diagnosis. Table 6 presents the results of analyses in which we tested for both multiplicative and additive interactions between smoking, categorized as ever/never smoking and then dichotomized as ≤10 or ≥10 pack-years of smoking, for each of the genotypes. Among those with the CC genotype of PTPN22, a modest effect of heavy smoking was observed (OR = 1.22 [95% CI = 0.81 to 1.83]). However, among those with the PTPN22 T risk allele, the effect of heavy smoking was much more pronounced (OR = 2.50 [95% CI = 1.25 to 5.00]). We observed significant additive and multiplicative gene-environment interactions between heavy cigarette smoking and the presence of the PTPN22 T allele (additive interaction: $P = 0.0006$; multiplicative interaction: $P = 0.04$). When smoking was dichotomized as never/ever, there was marginal evidence for additive but not multiplicative interaction. We also tested for genotype-smoking interactions in RF-positive and RF-negative RA cases separately. In stratified analyses, we found significant additive but not multiplicative interactions between the PTPN22 risk allele and heavy smoking for both seropositive and seronegative RA. We did not observe similar gene-smoking interactions for CTLA-4 or PADI-4, for the overall risk for RA, or for RF-positive or RF-negative RA separately. No additive or multiplicative interactions were observed between PTPN22 and HLA-SE (Table 7). (Given potential difficulties with HLA-SE genotyping)

Table 5

| Gene/genotype | RF status of RA cases (versus controls) |
|---------------|----------------------------------------|
|               | RF-positive cases | RF-negative cases |

PTPN22

| Genotype | RF status of RA cases (versus controls) |
|----------|----------------------------------------|
| CC       | 1.00 (ref.) | 1.00 (ref.) |
| CT/TT    | 1.75 (1.18–2.59) | 1.32 (0.85–2.06) |

CTLA-4

| Genotype | RF status of RA cases (versus controls) |
|----------|----------------------------------------|
| AA       | 1.00 (ref.) | 1.00 (ref.) |
| AG/GG    | 1.18 (0.78–1.80) | 1.11 (0.71–1.72) |

PADI-4

| Genotype | RF status of RA cases (versus controls) |
|----------|----------------------------------------|
| GG       | 1.00 (ref.) | 1.00 (ref.) |
| GA/AA    | 1.10 (0.79–1.54) | 0.96 (0.66–1.38) |

The analysis was unconditional logistic regression adjusting for year of birth, pack-year smoking, parity, breast-feeding, menstrual irregularity, age at menarche, menopausal status and postmenopausal hormone use. Values are expressed as odds ratio (95% confidence interval). CI, confidence interval; NHS, Nurses’ Health Study; OR, odds ratio; RA, rheumatoid arthritis; RF, rheumatoid factor.

Table 6

| Smoking status/pack-years | PTPN22 Genotype | All cases | RF-positive cases | RF-negative cases |
|---------------------------|-----------------|-----------|------------------|------------------|
|                           |                 |           |                  |                  |
|                           | Never           | CC        | 1.00 (ref.)      | 1.00 (ref.)      |
|                           | Ever            | CC        | 1.10 (0.79–1.53) | 1.37 (0.93–2.01) | 1.02 (0.68–1.53) |
|                           | Never           | CT/TT     | 1.07 (0.61–1.86) | 1.57 (0.86–2.87) | 0.87 (0.43–1.77) |
|                           | Ever            | CT/TT     | 1.90 (1.14–3.16) | 2.56 (1.47–4.46) | 1.67 (0.91–3.06) |
| Additive interaction      |                 |           |                  |                  |
|                           |                 |           | AP = 0.39 (95% CI 0.04 to +0.82); $P_{add}^{a} = 0.08$ | |
| Multiplicative interaction| $P_{multi}^{a} = 0.20$ | $P_{multi}^{a} = 0.67$ | $P_{multi}^{a} = 0.18$ |

Pack-years smoking

| Smoking status | PTPN22 Genotype | All cases | RF-positive cases | RF-negative cases |
|----------------|-----------------|-----------|------------------|------------------|
| Never          | CC              | 1.00 (ref.) | 1.00 (ref.)      | 1.00 (ref.)      |
| Ever           | CC              | 1.25 (0.89–1.75) | 1.40 (0.96–2.05) | 1.22 (0.81–1.83) |
| Never          | CT/TT           | 1.09 (0.69–1.73) | 1.48 (0.89–2.47) | 1.00 (0.55–1.82) |
| Ever           | CT/TT           | 3.05 (1.64–5.66) | 3.27 (1.74–6.15) | 2.50 (1.25–5.00) |
| Additive interaction |                 |           |                  |                  |
|                           |                 |           | AP = 0.56 (95% CI 0.24 to 0.88); $P_{add}^{a} = 0.0006$ | |
| Multiplicative interaction| $P_{multi}^{a} = 0.04$ | $P_{multi}^{a} = 0.28$ | $P_{multi}^{a} = 0.13$ |

Values are expressed as odds ratio (95% confidence interval). $^{a}$Conditional logistic regression, adjusting for parity, breast-feeding, menstrual irregularity, and age at menarche, menopausal status and postmenopausal hormone use. $^{b}$Unconditional logistic regression adjusting for year of birth, parity, breast-feeding, menstrual irregularity, and age at menarche, menopausal status and postmenopausal hormone use. $^{c}$RF-positive RA cases and all controls. $^{d}$RF-negative RA cases and all controls. $^{P_{add}}$ is the $P$ value for attributable proportion (AP), one of the indices of additive interactions between binary smoking variable and binary PTPN22 genotype. $^{P_{multi}}$ is the $P$ value for multiplicative interaction term between binary smoking variable and binary PTPN22 genotype with one degree of freedom. NHS, Nurses’ Health Study; RA, rheumatoid arthritis; RF, rheumatoid factor.
The C→T polymorphism at position 1858 of the PTPN22 gene interferes with the function of the PTPN22/Csk complex, which is an important inhibitor of T-cell signaling, hindering its ability to suppress T-cell activation [9,25,55]. Similar to past reports, we have found the elevated risk to be primarily for RF-positive disease [9,25,56-58]. Several reports and a meta-analysis have found the elevated risk to be primarily for RF-positive disease [9,26,55]. Similar to past reports, we have found that cigarette smoking, parity, total duration of breastfeeding, menstrual irregularity, and age at menarche, menopausal status and postmenopausal hormone use. We did not confirm that the R620W polymorphism was associated with increased risk for RA. We did not confirm that the PADI-4 (rs2240340) or the CTLA-4 (rs3087243) polymorphism were associated with increased risk for RA or for RF-positive RA in this population. We did not find that cigarette smoking, parity, total duration of breastfeeding, age at menarche, regularity of menses, menopausal status, or postmenopausal hormone use – all associated with risk for RA in past studies – were important confounders of the relationships between these genotypes and RA. However, we did uncover a significant multiplicative gene-environment interaction between heavy smoking and PTPN22 in determining RA risk.

In the case of HLA-SE, it is hypothesized that cigarette smoking leads to inflammation and citrullination of certain peptides, which – when presented within the context of HLA-DR4 molecules – are specifically recognized, contributing to anti-citrulline autoimmunity [37]. The newly described interaction between PTPN22 and heavy cigarette smoking suggests that the smoking/citrullination/T-cell recognition and activation pathway in RA pathogenesis may be influenced by both PTPN22 and HLA-SE.

The CTLA-4 gene is an attractive candidate gene for RA susceptibility, given the role played by CTLA-4 (cytotoxic T-lymphocyte associated 4) in T-cell activation and that a CTLA-4-IgG1 fusion protein is very effective in treating RA [62]. The

Table 7

PTPN22 genotype and HLA-SE interactions according to RF status in NHS/NHSII pooled samples

| PTPN22 | HLA-SE | All cases | RF-positive cases | RF-negative cases |
|--------|--------|-----------|------------------|------------------|
| CC     | None   | 1.00 (ref.) | 1.00 (ref.)      | 1.00 (ref.)      |
| CC     | Any    | 1.97 (1.39–2.78) | 2.37 (1.62–3.46) | 1.33 (0.89–1.99) |
| T carrier | None | 1.41 (0.87–2.27) | 1.76 (0.99–3.13) | 1.40 (0.78–2.51) |
| T carrier | Any  | 2.73 (1.55–4.81) | 3.58 (2.02–6.35) | 1.47 (0.74–2.92) |

Additive interaction

| Additive interaction | AP = 0.13 (95% CI -0.41 to +0.67); P add = 0.62 | AP = 0.13 (95% CI -0.41 to +0.67); P add = 0.64 | AP = -0.18 |

Multiplicative interaction

| Multiplicative interaction | P mul = 0.97 | P mul = 0.72 | P mul = 0.60 |

Values are expressed as odds ratio (95% confidence interval). Conditional logistic regression, adjusting for parity, pack-year smoking, breastfeeding, menstrual irregularity, and age at menarche, menopausal status and postmenopausal hormone use. Unconditional logistic regression adjusting for year of birth, parity, pack-year smoking, breastfeeding, menstrual irregularity, and age at menarche, menopausal status and postmenopausal hormone use. RF-positive RA cases and all controls. RF-negative RA cases and all controls. 

In these two cohorts of women followed prospectively for the development of RA and for multiple potential environmental exposures, we have confirmed that the R620W polymorphism in the PTPN22 gene is associated with increased risk for RA. Kallberg and colleagues [38] recently explored potential gene-environment and gene-gene interactions in RA susceptibility, combining data from three large RA cohort studies. The results of their study are slightly different from ours, in that they did not find a significant interaction between the presence of the PTPN22 polymorphism and smoking in determining RA risk. Their gene-smoking interaction analyses used data from the Swedish Environmental Investigations in RA incident RA cohort, in which participants were asked to recall past smoking and were classified as ever or never smokers. Using the detailed prospective data regarding smoking amount and duration available for NHS and NHSII participants, we demonstrated a multiplicative interaction between the presence of the PTPN22 risk allele and heavy cigarette smoking of ≥10 pack-years in this female cohort. In past studies, we have found that the risk for RA was significantly elevated with ≥10 pack-years [35]. Our results now suggest that it may be necessary to exceed a threshold of heavy smoking to trigger a biologic pathway in RA pathogenesis involving the PTPN22 gene. Both HLA-SE and PTPN22 primarily affect the risk for RF-positive and anti-CCP-positive RA [25,36,59-61].
Citrullination of peptides to which anti-CCP antibodies are 71% power to detect an effect of 40% or greater (OR = 1.4).

Among controls and a two-sided type I error rate of 0.05, we had formed, is encoded by the PADI-4 gene. The PADI4_94 SNP we have investigated was associated with RA in Japanese subjects (OR = 1.97 [95% CI = 1.44 to 2.69]) [8]. The effect sizes observed in previous replication studies in Caucasians have been small (pooled OR = 1.1 [95% CI = 1.0 to 1.2]) [25, 64]. We were unable to detect an effect of this polymorphism on the risk for RA in these cohorts of women, and this could reflect inadequate power to detect a risk estimate of that magnitude. Given that the allele frequencies in the controls were similar in each of the cohorts to that reported in the literature, the significant P value for heterogeneity across the cohorts we observed was probably due to small sample size.

Limitations of this study that should be noted include the fact that, in the NHS and NHSII cohorts, the presence or absence of RF in the blood among RA cases was confirmed by medical record review at diagnosis, and thus not assayed at the same laboratory, and was not assayed in controls. Rheumatoid nodules and radiographic erosions are likewise documented at the time of diagnosis from thorough medical record review, but cohort participants have not been followed longitudinally for RA disease activity or complications. Similarly, we have limited data in the medical record on antibodies to CCP among the cases, which is important in the subphenotyping of RA [64], because the dates of diagnosis for most of the RA cases in this cohort preceded the clinical use of anti-CCP. Further analysis by anti-CCP status could be potentially informative.

Although all participants included in this analysis were of self-reported Caucasian ancestry, potential population stratification, or confounding by ethnicity, still exists, in particular if the inclusion of individuals of Northern compared with Southern European origin varied between cases and controls [48, 65, 66]. We assessed the potential for this bias in two ways. First, we examined and did not find significant differences in the more precise racial backgrounds reported by the Caucasian women included as cases or controls in these analyses. Second, we genotyped the lactase gene, which is known to exhibit substantial variation in allele frequency from Northern to Southern Europe [47, 48], and found no significant differences in allele frequencies between cases and controls. A recent whole-genome association study investigating breast cancer risk alleles [67] found no evidence of population stratification among self-reported Caucasian women in the NHS cohort.

This study is unique in that the participants were followed for many years, in great detail, before the onset on their RA, and environmental and reproductive risk factors for RA have been well studied in this cohort [35, 41]. This has allowed the investigation of possible gene-environment interactions with each of these recently described polymorphisms, and known and suspected RA risk factors assessed prospectively, such as cigarette smoking and menopausal status.

Conclusion
Our data confirm that the PTPN22 R620W polymorphism is a strong risk factor for RF-positive RA, and that presence of this polymorphism interacts with heavy cigarette smoking in a multiplicative manner. These findings contribute to the growing understanding of how genetic and environmental factors interact in RA pathogenesis, and suggest that heavy cigarette smoking and PTPN22 may be acting in a similar mechanistic pathway.

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
KHC was responsible for study design, data acquisition, analysis and interpretation of data, manuscript preparation, and statistical analysis. S-CC was responsible for analysis and interpretation of data, manuscript preparation, and statistical analysis. IDV was responsible for analysis and interpretation of data, manuscript preparation, and statistical analysis. RP was responsible for study design, analysis and interpretation of data, and manuscript preparation. EWV was responsible for study design, data acquisition, analysis and interpretation of data, manuscript preparation, and statistical analysis.

Acknowledgements
Supported by NIH grants CA87969, P60 AR047782, R01 AR49880, K24 AR0524-01 and BIRCWH K12 HD051959 (supported by the NIMH, NIAID, NICHD, and OD). Dr Costenbader is the recipient of an Arthritis Foundation/American College of Rheumatology Arthritis Investigator Award and a Katherine Swan Ginsburg Memorial Award.

The authors gratefully acknowledge the participants in the NHS studies for their continuing cooperation. The authors also thank Frank Speizer and Walter Willett. We are grateful to Gideon Aweh and Karen Corsano for their technical assistance.

References
1. Gabriel SE, Crowson CS, O’Fallon WM: The epidemiology of rheumatoid arthritis in Rochester, Minnesota, 1955–1985. Arthritis Rheum 1999, 42:145-20.
2. Coménés F, Fauré S, Martinez M, Prud’homme JF, Fritz P, Dib C, Alves H, Barrera P, de Vries N, Balsa A, Pascual-Salcedo D, Maenaut K, Westhovens R, Migliorini P, Tran TH, Delaye A, Prince N, Lefèvre C, Thomas G, Poirier M, Soubigou S, Albert O, Lasbliez S, Foux S, Bouchier C, Lutè F, Lose MN, Lepage V, Charon D, Gypay G, et al.: New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. Proc Natl Acad Sci USA 1998, 95:10746-50.
3. Jawaheer D, Gregersen PK: Rheumatoid arthritis. The genetic components. Rheum Dis Clin North Am 2002, 28:1-15. v
4. Gregersen PK, Silver J, Winchester RJ: The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum 1987, 30:1205-13.

5. Thomson W, Harrison B, Ollier B, Wiles N, Payton T, Barrett J, Thomson W, Harrison B, Ollier B, Wiles N, Payton T, Barrett J, Symmons D, Silman A: Quantifying the exact role of HLA-DRB1 alleles in susceptibility to inflammatory polyarthritides: results from a large, population-based study. Arthritis Rheum 1999, 42:757-762.

6. Hantshel H, Wagner U: Analysis of families in the Multiple Autoimmune Disease Genetics Consortium (MADGC) collection: the PTPN22 620W allele associates with rheumatoid arthritis. J Hum Genet 2004, 50:399-400.

7. Oliveira SA, Damasceno FA, Carvalho BR, Dias AA: Cigarette smoking is strongly associated with rheumatoid arthritis and systemic lupus erythematosus. Arthritis Rheum 2005, 52:1694-1699.

8. Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M, MacMurray J, Meloni GF, Lucarelli P, Pellecchia M, Eisenbarth GS, Comings D, Mustelin T: A functional variant of lymphoid tyrosine phosphatase is associated with type 1 diabetes. Ann Hum Genet 2007, 71:337-338.

9. Zheng W, She JX: Genetic variation and chronic inflammatory arthropathies in Northern Ireland. Arthritis Rheum 2005, 52:306-16.

10. Gregersen-PK, Silver, J., Winchester, R.J.: The shared epitope hypothesis. Arthritis Rheum 1987, 30:1205-13.

11. Thomson, W., Harrison, B., Ollier, B., Wiles, N., Payton, T., Barrett, J., Symmons, D., Silman, A.: Quantifying the exact role of HLA-DRB1 alleles in susceptibility to inflammatory polyarthritides: results from a large, population-based study. Arthritis Rheum 1999, 42:757-762.

12. Hantshel, H., Wagner, U.: Analysis of families in the Multiple Autoimmune Disease Genetics Consortium (MADGC) collection: the PTPN22 620W allele associates with rheumatoid arthritis. J Hum Genet 2004, 50:399-400.

13. Oliveira, S.A., Damasceno, F.A., Carvalho, B.R., Dias, A.A.: Cigarette smoking is strongly associated with rheumatoid arthritis and systemic lupus erythematosus. Arthritis Rheum 2005, 52:1694-1699.

14. Zheng, W., She, J.X.: Genetic variation and chronic inflammatory arthropathies in Northern Ireland. Arthritis Rheum 2005, 52:306-16.
tion based case-control study, using incident cases. Ann Rheum Dis 2003, 62:835-841.
34. Karlson EW, Mandl LA, Aweh GN, Grodstein F: Coffee consumption and risk of rheumatoid arthritis. Arthritis Rheum 2003, 48:3055-3060.
35. Costenbader KH, Feskanich D, Mandl LA, Karlson EW: Smoking intensity, duration, and cessation, and the risk of rheumatoid arthritis in women. Am J Med 2006, 119:303e1-9.
36. Padyukov L, Silva C, Stolt P, Alfredsson L, Klareskog L: A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. Arthritis Rheum 2004, 50:3089-3092.
37. Karlsson EW, Stolt P, Lind L, Leborg K, Kallberg H, Bengtsson C, Grunewald J, Ronnelid J, Erdalnsness H, Ulfkaren AG, Rantanap-Dahlqvist S, Eklund A, Padyukov L, Alfredsson L: A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DRB1 shared epitope restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum 2006, 54:38-46.
38. Kallberg H, Padyukov L, Plenge RM, Ronnelid J, Gregersen PK, Helm-van Mil AH van der, Toes RE, Huizinga TW, Klareskog L, All HJ, Walker TM, Nemanic-Dragic M, Hugot JP, Notton CH: Association with involvement HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. Am J Hum Genet 2007, 80:867-875.
39. Hernandez-Avila M, Liang MH, Willett WC, Stamper MJ, Colditz GA, Rosner B, Chang RW, Hennekens CH, Speizer FE: Exogenous sex hormones and the risk of rheumatoid arthritis. Arthritis Rheum 1990, 33:947-953.
40. Hernandez-Avila M, Liang MH, Willett WC, Stamper MJ, Colditz GA, Rosner B, Roberts WN, Hennekens CH, Speizer FE: Reproductive factors, smoking, and the risk for rheumatoid arthritis. Epidemiology 1990, 1:289-291.
41. Karlson EW, Mandl LA, Hankinson SE, Grodstein F: Do breastfeeding and other reproductive factors influence future risk of rheumatoid arthritis? Results from the Nurses’ Health Study, Arthritis Rheum 2004, 50:3458-3467.
42. Karlson EW, Daltroy LH, Lew RA, Willett EA, Partridge AJ, Roberts WN, Stern SH, Straatov KP, Wacholtz MC, Grofisam JM: The independence and stability of socioeconomic predictors of morbidity in systemic lupus erythematosus. Arthritis Rheum 1996, 38:267-273.
43. Amett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, Healey LA, Kaplan SR, Liang MH, Luthra HS: The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988, 31:315-324.
44. Hankinson SE, Colditz GA, Hunter DJ, Manson JE, Willett WC, Stamper MJ, Longcope C, Speizer FE: Reproductive factors and family history of breast cancer in relation to plasma estrogen and prolactin concentrations and risk of breast cancer. Cancer Causes Control 1995, 6:217-224.
45. Tworoger SS, Sluss P, Hankinson SE: Association between plasma prolactin concentrations and risk of breast cancer among predominantly premenopausal women. Cancer Res 2006, 66:2476-2482.
46. Ueda H, Howson JM, Esposito L, Heward J, Snook H, Chamberlain G, Rainford DB, Hunter KM, Smith AN, Di Genova G, Herr MH, Dahlman I, Payne I, Smyth D, Lowe C, Twells RC, Howlett SR, Heal Y, Nutland S, Rance HE, Everett V, Smink LJ, Lam AC, Cordell H, Walker TM, Nemanic-Dragic M, Hugot JP, Notton CH, et al.: Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. Nature 2003, 423:506-511.
47. Beraugliot T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, Drake JA, Rhodes M, Reich DE, Hirschhorn JN: Genetic signatures of strong recent positive selection at the lactase gene. Am J Hum Genet 2004, 74:1111-1120.
48. Campbell CD, Ogburn EL, Lunetta KL, Lyon HN, Freedman ML, Groop LC, Altschuler D, Ardlie KG, Hirschhorn JN: Demonstrating stratification in a European American population. Nat Genet 2005, 37:868-872.
49. Colditz GA, Stamper MJ, Willett WC, Stason WB, Rosner B, Hennekens CH, Speizer FE: Reproducibility and validity of self-reported menopausal status in a prospective cohort study. Am J Epidemiol 1987, 126:319-325.
50. DerSimonian R, Laird N: Meta-analysis in clinical trials. Control Clin Trials 1988, 7:171-188.
51. Anderson A, Wilt, Alfredsson L, Kallberg H, Zdravkovic S, Ahlborn A: Calculating measures of biological interaction. Eur J Epidemiol 2005, 20:575-579.
52. Ahlborn A, Alfredsson L: Interaction: a word with two meanings creates confusion. Eur J Epidemiol 2005, 20:563-564.
53. Lundberg M, Fredlund P, Hallqvist J, Didierchsen F: A SAS program calculating three measures of interaction with confidence intervals. Epidemiology 1999, 7:655-656.
54. Hosmer DW, Lemeshow S: Confidence interval estimation of interaction. Epidemiology 2002, 13:82-85.
55. Closter JF, Vilella A: Cooperative inhibition of T-cell antigen receptor signaling by a complex between a kinase and a phosphatase. J Exp Med 1999, 189:111-121.
56. Dieudef G, Garrier S, Michou L, Petit-Teixeira E, Gliksman E, Pietrot C, Rivesbliez S, Bard F: Rheumatoid arthritis serositive for the rheumatoid factor is linked to the protein tyrosine phosphate nonreceptor 22-620W allele. Arthritis Rheum 2005, 58:358-360.
57. Steer S, Lad B, Gramley JA, Kingsley GH, Fisher SA: Association of the R602W in a protein tyrosine phosphatase gene with risk of rheumatoid arthritis in a British population: evidence for an early onset/disease severity effect. Arthritis Rheum 2005, 52:358-360.
58. Wesoly J, Helm-van Mil AH van der, Toes RE, Chokkalingam AP, Corrigan VE, Begovic T, Abu, Huizinga TW: Association of the PTPN22 C1858T single-nucleotide polymorphism with rheumatoid arthritis phenotypes in an inception cohort. Arthritis Rheum 2005, 52:2948-2950.
59. Huizinga TW, Amos CI, Helm-van Mil AH van der, Chen W, van Graa- flein FA, Jawaeher D, Schreuder GM, Wener M, Breedveld FC, Ahmad N, Lum RF, de Vries RR, Gregersen PK, Toes RE, Criswell LA: Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum 2005, 52:343-349.
60. Irigoyen P, Lee AT, Wener MH, Li W, Kern M, Batliwalla F, Lum RF, Massarotti E, Weisman M, Bombardier C, Remmers EF, Kastner DL, Selin MF, Criswell LA, Gregersen PK: Regulation of anti-cyclical citrullinated peptide antibodies in rheumatoid arthritis: contrasting effects of HLA-DR3 and the shared epitope alleles. Arthritis Rheum 2005, 52:3813-3818.
61. Johansson AS, Stenberg G, Wijmer M, Mennvik B, Strubinger: Structure-activity relationships and thermal stability of human g- lactocidase transerase P1-1 covered by the H-site residue 105. J Mol Biol 1998, 278:687-698.
62. Kremer JM, Westhovens R, Leon M, Giorgio E, Alten R, Steinfeld S, Russell A, Douagados M, Emery P, Naumah IF, Williams GR, Becker JC, Hagerdy DT, Moreland LW: Treatment of rheumatoid arthritis by selective inhibition of T-cell activation with fusion protein CTLA4ig. N Engl J Med 2003, 349:1907-1915.
63. Lee YH, Rho YH, Choi SJ, Ji JD, Song GG: PADI4 polymorphisms and rheumatoid arthritis susceptibility: a meta-analysis. Rheumatol Int 2007, 27:827-833.
64. Berulin E, Padyukov L, Sundin U, Hallmehn L, Vanrenooij WJ, Klareskog L, Dahlqvist SR: A combination of autoantibodies to cyclic citrullinated peptide (CCP) and HLA-DRB1 locus antigens is strongly associated with future onset of rheumatoid arthritis. Arthritis Res Ther 2004, 6:R303-R308.
65. Epstein MP, Allen AS, Saten GA: A simple and improved correction for population stratification in case-control studies. Am J Hum Genet 2007, 80:921-930.
66. Wacholder S, Rothman N, Gapor N: Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. J Natl Cancer Inst 2000, 92:1115-1158.
67. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A, Wang J, Yu K, Chatterjee N, Orr N, Willett WC, Colditz GA, Ziegler RA, Berg CD, Buys SS, McCarty CA, Fiegelson HS, Calle EE, Thun MJ, Hayes RB, Tucker M, Gerhard DS, Fraumeni JF Jr, Hoover RN, Thomas G, Chanock SJ: A genome-wide association study identifies alleles in FGF22 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 2007, 39:870-874.