Concise report

An Immunochip-based interaction study of contrasting interaction effects with smoking in ACPA-positive versus ACPA-negative rheumatoid arthritis

Xia Jiang1,*, Henrik Källberg1,*, Zuomei Chen1, Lisbeth Ärlestig2, Solbritt Rantapää-Dahlqvist2, Sonia Davila3, Lars Klareskog4, Leonid Padyukov4,† and Lars Alfredsson1,5,‡

Abstract

Objective. To investigate the gene-environment interaction between smoking and single nucleotide polymorphisms (SNPs), using Immunochip material, on the risk of developing either of two serologically defined subsets of RA.

Methods. Interaction between smoking and 133 648 genetic markers from the Immunochip was examined for two RA subsets, defined by the presence or absence of ACPA. A total of 1590 ACPA-positive and 891 ACPA-negative cases were compared with 1856 controls in the Swedish Epidemiological Investigation of RA (EIRA) case-control study. Logistic regression models were used to determine the presence of interaction. The proportion attributable to interaction was calculated for each smoking-SNP pair. Replication was carried out in an independent dataset from northern Sweden. To further validate and extend the results, interaction analysis was also performed using genome-wide association studies data on EIRA individuals.

Results. In ACPA-positive RA, 102 SNPs interacted significantly with smoking, after Bonferroni correction. All 102 SNPs were located in the HLA region, mainly within the HLA class II region, 51 of which were replicated. No additional loci outside chromosome 6 were identified in the genome-wide association studies validation. After adjusting for HLA-DRB1 shared epitope, 15 smoking-SNP pairs remained significant for ACPA-positive RA, with 8 of these replicated (loci: BTNL2, HLA-DRA, HLA-DRB5, HLA-DQA1, HLA-DOB and TAP2). For ACPA-negative RA, no smoking-SNP pairs passed the threshold for significance.

Conclusion. Our study presents extended gene variation patterns involved in gene-smoking interaction in ACPA-positive, but not ACPA-negative, RA. Notably, variants in HLA-DRB1 and those in additional genes within the MHC class II region, but not in any other gene regions, showed interaction with smoking.

Key words: rheumatoid arthritis, Immunochip, smoking, biological interaction, HLA-DRB1 shared epitope

Rheumatology key messages

- HLA region remains the most important region for gene-smoking interaction in ACPA-positive RA.
- No single nucleotide polymorphisms outside chromosome 6 interacting with smoking in ACPA-positive RA need further validation.
- Well-powered genome-wide interaction studies are needed to identify gene-smoking interactions in ACPA-negative RA.

1The Institute of Environmental Medicine, Karolinska Institute, Stockholm, 2Department of Public Health and Clinical Medicine, Rheumatology, Umeå University, Umeå, Sweden, 3Human Genetics, Genome Institute of Singapore, Singapore, 4Rheumatology Unit, Department of Medicine, Karolinska Institute and 5Centre for Occupational and Environmental Medicine, Stockholm County Council, Stockholm, Sweden

Submitted 26 November 2014; revised version accepted 8 July 2015

Correspondence to: Xia Jiang, Institute of Environmental Medicine, Karolinska Institute, Novels väg 13, Box210, Solna, 171 77 Stockholm, Sweden. E-mail: Xia.jiang@ki.se

*Xia Jiang and Henrik Källberg contributed equally to this study.

†Leonid Padyukov and Lars Alfredsson contributed equally to this study.
Introduction

RA occurs as a consequence of the combined effect of genes and environments, and can be divided into two major subsets, according to the presence or absence of ACPAs. Cigarette smoking, the most important environmental risk factor, appears to affect ACPA-positive RA patients only [1, 2]. The shared epitope (SE) of the HLA-DRB1 gene and a polymorphism of the PTPN22 gene are the two major genetic risk factors for ACPA-positive RA, whereas they only slightly increase the risk of ACPA-negative RA [3]. To date, over 100 RA risk loci have been identified due to the success of genome-wide association studies (GWASs) [4–6]. However, these genetic variants cumulatively explain only a limited proportion of ACPA-positive RA occurrence, and an even smaller proportion of ACPA-negative RA occurrence. In addition, the genome-wide approach does not take into account the complex interactions that often occur in biological systems. Based on the candidate gene approach, we and others have in earlier work described a striking smoking–SE interaction in seropositive RA [7–9]. However, these previous reports did not examine the whole genome.

Therefore, the main aim of this study was to investigate a smoking–single nucleotide polymorphism (SNP) interaction in the two RA subsets through a two-stage approach. First, smoking–SNP interactions were identified in one case–control study, using Immunochip (ICHIP) data as the source of genetic data. Replication was accomplished in a second stage, using another cohort of RA patients and controls. Finally, we performed an additional validation, using GWAS data as a basis for the genetic data from the first cohort. See supplementary Fig. S1, available at Rheumatology online, for an overview of the study design.

Methods

Study population

This study is based on two cohorts: the Epidemiological Investigation of RA (EIRA) from southern/central Sweden (discovery dataset) and a nested case–control study from northern Sweden (Umeå) (replication dataset). Ethical approval for the EIRA and Umeå cohorts as well as the analyses performed in the current study was obtained from the relevant ethical committees (EIRA: the ethical review board of Karolinska Institutet; Umeå: the ethical review board of Umeå University), and all subjects consented to participate in the EIRA and Umeå studies by signing a written consent form.

Discovery dataset (EIRA)

The EIRA, an ongoing population-based case–control study, comprises incident cases and controls aged 18–70 years. A case was defined as a subject in the study base who was given a first-time RA diagnosis according to the 1987 ACR criteria [10], by rheumatologists. Controls were randomly selected from a continuously updated national population register. Details of the study design have been reported elsewhere [11].

Replication dataset (Umeå)

The replication dataset consists of all eligible patients from the four northern-most counties of Sweden, diagnosed with early RA fulfilling the 1987 ACR criteria [10], who were consecutively included by rheumatologists at the University Hospital Umeå. Controls were randomly selected and identified from the Medical Biobank of northern Sweden. Details of the study base have been described elsewhere [12].

ELISA, HLA-DRB1 SE genotyping and imputation of HLA alleles

We analysed serum samples from the cases for ACPA status using ELISA according to the manufacturer’s instructions. Genotyping procedure for the HLA-DRB1 gene has been described in previous publications [8, 13]. In the Umeå cohort, HLA-DRB1 genotypes were available for 60% (n = 586) of the controls. We carried out imputation on the classical HLA alleles for all Umeå controls through a previously published imputation procedure, using an independent reference dataset collected by T1DGC, in the BEAGLE package [14]. We compared the true genotyped HLA-SE with the imputed HLA-SE on the four-digit resolution level of HLA-DRB1*04 alleles. The imputed data achieved a concordance rate of 97.44%.

Genotyping and quality control

Genetic information was obtained from ICHIP, a custom-made Illumina Infinium array including a variety of SNPs of specific immunologic interest, based on observations from different autoimmune diseases [15]. The EIRA ICHIP scan included 195586 genetic markers from 5043 samples. Data were filtered on the basis of both SNPs and individuals, executed in PLINK [16]. Population stratification was controlled by using the principal component approach, as implemented in EIGENSTRAT [17]. After quality control, a total of 133648 SNPs and 4337 participants (1590 ACPA-positive cases, 891 ACPA-negative cases and 1856 controls) passed filtering (supplementary Fig. S2, available at Rheumatology Online). The same ICHIP data were available in 1859 Umeå (614 ACPA-positive cases, 271 ACPA-negative cases and 974 controls) samples.

The ICHIP primarily targets the known immunity-related genes and therefore is not truly genome-wide. To further extend our findings, we included EIRA GWAS data (performed using an Illumina 300K chip, available for 1147 ACPA-positive cases, 774 ACPA-negative cases and 1079 controls, on 301171 markers after quality control; GWAS and ICHIP were overlapping for 18000 SNPs and 2458 individuals, with a genotyping concordance rate of 99.9%) in our study. The smoking–GWAS interaction analysis was performed to identify any possible signals that might be neglected due to the special design of the ICHIP.
Cigarette smoking and covariate information
We collected information regarding lifestyle-related factors from the subjects through self-reported questionnaires at baseline. Ever-smokers were defined as individuals who smoked or previously had smoked cigarettes, while never-smokers were defined as individuals who had never smoked cigarettes.

Gene-smoking interaction assessment
We estimated the additive interaction defined by Rothman, where sufficient cause interaction happens when the joint effect of two contributory causes exceeds the sum of their independent effects [18]. We calculated the odds ratios (ORs) by comparing double-exposed individuals (risk-allele carriers and smokers) and single-exposed individuals (risk-allele carriers or smokers) with the non-exposed reference group (non-carriers and non-smokers) through unconditional logistic regression with adjustment for the design variables (age, gender, residential area and study phases). We subsequently calculated the proportion attributable to interaction based on those ORs. A Bonferroni-corrected P-value (for the attributable proportion) of <0.05 (corresponding to a genome-wide P-value of 3.74 × 10⁻⁷) was used as the threshold for significance. Three genetic models (dominant, recessive and co-dominant) were applied. Since interaction is commonly examined on a multiplicative scale, we also carried out such analysis (by fitting a smoking–SNP interaction term into the model) as a complement. All analyses were performed in the SAS version 9.3 of the GEIRA program [19].

Results
Basic characteristics
The characteristics of the subjects are summarized in supplementary Table S1, available at Rheumatology Online. Among cases in EIRA, 66.9% were ever-smokers, compared with 57.1% among controls. Smoking was more common in ACPA-positive RA, as compared with ACPA-negative RA (71.2% vs 59.2%). In the Umeå data-set, 885 cases (incident RA cases) and 974 controls were available. The proportion of ever-smokers was higher in cases (65.1% vs 49.2%) and more common in ACPA-positive cases (69.2% vs 55.6%). The smoking–RA association was found only in ACPA-positive RA (OR 2.0), but not in ACPA-negative RA cases (OR 1.1) (supplementary Table S2, available at Rheumatology Online).

Identification and replication of a smoking–SNP interaction, without adjustment of HLA-SE
In EIRA, 101 SNPs from the co-dominant model, 35 SNPs from the dominant model and nine SNPs from the recessive model were found to interact significantly with smoking in ACPA-positive RA. In total, 102 non-duplicated SNPs were noted, all located in the HLA region. The lack of interaction signals outside the HLA region might be due to the selected coverage of ICHIP markers. We therefore also performed GWAS–smoking interaction analysis, but no additional signals outside chromosome 6 were obtained (supplementary Fig. S3, available at Rheumatology Online). In contrast, no single SNP was found to be interacting with smoking in the ACPA-negative RA subset. Subsequently, we restricted our analyses solely to ACPA-positive RA cases and controls.

We then analysed all 102 SNPs in the replication dataset, in the same way as in the EIRA dataset. Only SNPs in the replication sample that reached a \( P < 0.05 \) and had the same direction of interaction were considered to be truly replicated. With these criteria, 51 SNPs were replicated (supplementary Table S3, available at Rheumatology Online). No evidence of statistically significant multiplicative interactions was observed for any of these 102 smoking–SNP pairs. Based on these 51 validated interaction signals, we selected 10 genes/gene regions (supplementary Table S4, available at Rheumatology Online), and 7 of the 10 closest genes (HLA-DOB, HLA-DQA1, HLA-DQA2, HLA-DRA, HLA-DRB1, HLA-DRB5, TAP2) correspond to an antigen-presenting pathway.

Identification and replication of smoking–SNP pairs independent of HLA-SE
Given the extensive linkage disequilibrium in the HLA region, we adjusted for SE on those previously identified smoking–SNP pairs to determine possible independent interaction signals. In EIRA, the numbers of interacting SNPs were strongly reduced after such adjustment, with 12 identified from the co-dominant model and eight from the recessive model (five overlapping SNPs). As shown in Tables 1 and 2, and supplementary Table S5, available at Rheumatology Online, altogether, eight non-duplicated SNPs were successfully replicated. All were from the HLA class II region (BTNL2, HLA-DRA, HLA-DRB5, HLA-DQA1, HLA-DOB and TAP2) and were in high linkage disequilibrium (supplementary Fig. S4, available at Rheumatology Online).

Identification of interaction effects between currently known RA risk loci and smoking
Based on the literature published to date, ~125 risk loci have been identified among Caucasian and Asian populations for RA overall or for ACPA-positive RA. We found 84 of these risk loci (supplementary Table S6, available at Rheumatology Online) and carried out the same interaction analysis in EIRA. We did not observe any evidence of significant interactions of those additional non-HLA SNPs with smoking, on either the additive or the multiplicative scale, after adjustment for multiple comparisons. Several suggestive interaction signals were obtained when the \( P \)-threshold was set to 0.05. These results are summarized in supplementary Table S7, available at Rheumatology Online.

Identification of multiplicative interactions
As a complement, we analysed gene–smoking interaction on the multiplicative scale, using all the ICHIP markers in the EIRA dataset. We did not observe any significant smoking–SNP pair after accounting for the multiple
### Table 1

The independent and joint effects of SNPs and smoking in ACPA-positive RA: odds ratios from the co-dominant model

<table>
<thead>
<tr>
<th>Chr. number</th>
<th>rs number</th>
<th>Ref group</th>
<th>OR (95% CI)</th>
<th>Case / control</th>
<th>Case / control</th>
<th>Case / control</th>
<th>Case / control</th>
<th>Case / control</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>rs1048456</td>
<td>A</td>
<td>1.07 (0.72, 1.61)</td>
<td>164 (78, 292)</td>
<td>127 (53, 171)</td>
<td>1.07 (0.72, 1.61)</td>
<td>164 (78, 292)</td>
<td>127 (53, 171)</td>
</tr>
<tr>
<td>6</td>
<td>rs3104413</td>
<td>G</td>
<td>1.28 (0.67, 2.47)</td>
<td>201 (91, 229)</td>
<td>170 (74, 231)</td>
<td>1.28 (0.67, 2.47)</td>
<td>201 (91, 229)</td>
<td>170 (74, 231)</td>
</tr>
<tr>
<td>6</td>
<td>rs3129769</td>
<td>A</td>
<td>1.12 (0.71, 1.76)</td>
<td>212 (94, 240)</td>
<td>166 (68, 222)</td>
<td>1.12 (0.71, 1.76)</td>
<td>212 (94, 240)</td>
<td>166 (68, 222)</td>
</tr>
<tr>
<td>6</td>
<td>rs3998159</td>
<td>C</td>
<td>1.07 (0.67, 1.71)</td>
<td>210 (91, 229)</td>
<td>166 (68, 222)</td>
<td>1.07 (0.67, 1.71)</td>
<td>210 (91, 229)</td>
<td>166 (68, 222)</td>
</tr>
<tr>
<td>6</td>
<td>rs6931277</td>
<td>A</td>
<td>1.11 (0.59, 2.15)</td>
<td>210 (91, 229)</td>
<td>166 (68, 222)</td>
<td>1.11 (0.59, 2.15)</td>
<td>210 (91, 229)</td>
<td>166 (68, 222)</td>
</tr>
<tr>
<td>6</td>
<td>rs7454108</td>
<td>G</td>
<td>1.04 (0.67, 1.62)</td>
<td>205 (91, 229)</td>
<td>166 (68, 222)</td>
<td>1.04 (0.67, 1.62)</td>
<td>205 (91, 229)</td>
<td>166 (68, 222)</td>
</tr>
<tr>
<td>6</td>
<td>rs9275330</td>
<td>A</td>
<td>1.08 (0.68, 1.73)</td>
<td>210 (91, 229)</td>
<td>166 (68, 222)</td>
<td>1.08 (0.68, 1.73)</td>
<td>210 (91, 229)</td>
<td>166 (68, 222)</td>
</tr>
<tr>
<td>6</td>
<td>rs9275530</td>
<td>C</td>
<td>1.11 (0.69, 1.85)</td>
<td>210 (91, 229)</td>
<td>166 (68, 222)</td>
<td>1.11 (0.69, 1.85)</td>
<td>210 (91, 229)</td>
<td>166 (68, 222)</td>
</tr>
<tr>
<td>6</td>
<td>rs9275532</td>
<td>T</td>
<td>1.07 (0.67, 1.71)</td>
<td>210 (91, 229)</td>
<td>166 (68, 222)</td>
<td>1.07 (0.67, 1.71)</td>
<td>210 (91, 229)</td>
<td>166 (68, 222)</td>
</tr>
<tr>
<td>6</td>
<td>rs9784635</td>
<td>A</td>
<td>1.08 (0.68, 1.73)</td>
<td>210 (91, 229)</td>
<td>166 (68, 222)</td>
<td>1.08 (0.68, 1.73)</td>
<td>210 (91, 229)</td>
<td>166 (68, 222)</td>
</tr>
</tbody>
</table>

**For each SNP, three rows of results were shown, representing results from the discovery dataset of the Swedish EIRA study (first row), the replication dataset of the Umeå study (second row) and pooled data combining both studies (third row). In EIRA, the odds ratios were adjusted for age, gender, and shared epitope. In both Umeå and the pooled analysis, the odds ratios were adjusted for age, gender, and shared epitope.**
### Table 2: The independent and joint effects of SNPs and smoking in ACPA-positive RA: odds ratios from the recessive model

<table>
<thead>
<tr>
<th>Chr.</th>
<th>rs number</th>
<th>Risk allele</th>
<th>Ref group case/control</th>
<th>Never-smoker 0/2 risk alleles case/control</th>
<th>OR (95% CI)</th>
<th>Never-smoker 0 or 1 risk allele case/control</th>
<th>OR (95% CI)</th>
<th>Ever-smoker 0 or 1 risk allele case/control</th>
<th>OR (95% CI)</th>
<th>Ever-smoker 0/2 risk alleles case/control</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>rs1048565</td>
<td>A</td>
<td>414/732</td>
<td>16/14</td>
<td>1.22 (0.56, 2.62)</td>
<td>101/984</td>
<td>1.85 (1.55, 2.20)</td>
<td>53/10</td>
<td>6.99 (3.37, 14.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>179/456</td>
<td>4/5</td>
<td>1.43 (0.32, 6.44)</td>
<td>399/444</td>
<td>1.75 (1.34, 2.29)</td>
<td>15/3</td>
<td>7.26 (1.85, 28.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>593/1188</td>
<td>20/19</td>
<td>1.31 (0.68, 2.53)</td>
<td>1410/1428</td>
<td>2.11 (1.84, 2.43)</td>
<td>69/13</td>
<td>7.48 (4.01, 13.96)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>rs3129890</td>
<td>A</td>
<td>145/365</td>
<td>285/381</td>
<td>1.22 (0.92, 1.62)</td>
<td>310/501</td>
<td>1.56 (1.19, 2.06)</td>
<td>754/493</td>
<td>2.61 (2.01, 3.39)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>67/211</td>
<td>116/250</td>
<td>1.07 (0.71, 1.61)</td>
<td>121/233</td>
<td>1.06 (0.70, 1.61)</td>
<td>293/214</td>
<td>2.51 (1.71, 3.68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>212/576</td>
<td>401/631</td>
<td>1.16 (0.93, 1.46)</td>
<td>431/734</td>
<td>1.61 (1.30, 2.01)</td>
<td>1047/707</td>
<td>2.98 (2.43, 3.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>rs3129891</td>
<td>G</td>
<td>116/299</td>
<td>314/447</td>
<td>1.29 (0.96, 1.74)</td>
<td>237/402</td>
<td>1.57 (1.15, 2.13)</td>
<td>827/592</td>
<td>2.65 (2.01, 3.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>58/175</td>
<td>126/285</td>
<td>1.01 (0.66, 1.54)</td>
<td>87/184</td>
<td>1.01 (0.64, 1.61)</td>
<td>327/263</td>
<td>2.18 (1.46, 3.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>174/474</td>
<td>440/732</td>
<td>1.16 (0.92, 1.47)</td>
<td>324/586</td>
<td>1.34 (1.17, 1.54)</td>
<td>1154/855</td>
<td>2.85 (2.29, 3.54)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>rs3957146</td>
<td>G</td>
<td>398/729</td>
<td>32/17</td>
<td>1.70 (0.88, 3.28)</td>
<td>979/982</td>
<td>1.84 (1.54, 2.20)</td>
<td>85/12</td>
<td>8.32 (4.21, 16.44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>170/448</td>
<td>16/13</td>
<td>1.97 (0.83, 4.66)</td>
<td>379/436</td>
<td>1.72 (1.31, 2.26)</td>
<td>41/11</td>
<td>4.12 (1.92, 8.84)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>568/1177</td>
<td>48/30</td>
<td>1.93 (1.17, 3.18)</td>
<td>1358/1418</td>
<td>2.10 (1.82, 2.42)</td>
<td>126/23</td>
<td>7.13 (4.39, 11.56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>rs3998159</td>
<td>C</td>
<td>398/729</td>
<td>32/17</td>
<td>1.70 (0.88, 3.28)</td>
<td>978/982</td>
<td>1.84 (1.54, 2.20)</td>
<td>86/12</td>
<td>8.40 (4.25, 16.60)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>167/448</td>
<td>17/13</td>
<td>2.06 (0.88, 4.87)</td>
<td>370/436</td>
<td>1.72 (1.31, 2.25)</td>
<td>40/11</td>
<td>4.17 (1.94, 8.95)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>565/1177</td>
<td>49/30</td>
<td>1.99 (1.21, 3.27)</td>
<td>1348/1418</td>
<td>2.09 (1.82, 2.41)</td>
<td>126/23</td>
<td>7.20 (4.44, 11.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>rs745108</td>
<td>G</td>
<td>391/725</td>
<td>32/17</td>
<td>1.73 (0.89, 3.33)</td>
<td>967/964</td>
<td>1.89 (1.58, 2.25)</td>
<td>85/12</td>
<td>8.46 (4.28, 16.74)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>169/448</td>
<td>15/13</td>
<td>1.75 (0.72, 4.25)</td>
<td>377/436</td>
<td>1.72 (1.31, 2.26)</td>
<td>38/11</td>
<td>4.09 (1.90, 8.82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>560/1173</td>
<td>47/30</td>
<td>1.90 (1.15, 3.13)</td>
<td>1344/1400</td>
<td>2.14 (1.86, 2.46)</td>
<td>123/23</td>
<td>7.09 (4.37, 11.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>rs9268557</td>
<td>G</td>
<td>248/554</td>
<td>182/192</td>
<td>1.32 (0.99, 1.76)</td>
<td>575/759</td>
<td>1.70 (1.37, 2.10)</td>
<td>489/235</td>
<td>3.10 (2.41, 3.98)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>103/337</td>
<td>82/124</td>
<td>1.59 (1.05, 2.42)</td>
<td>216/343</td>
<td>2.14 (1.86, 2.46)</td>
<td>202/104</td>
<td>4.05 (2.76, 5.93)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>351/891</td>
<td>264/361</td>
<td>1.40 (1.11, 1.75)</td>
<td>791/1102</td>
<td>1.89 (1.59, 2.24)</td>
<td>691/339</td>
<td>3.79 (3.10, 4.63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>rs9784858</td>
<td>G</td>
<td>411/728</td>
<td>17/16</td>
<td>1.18 (0.56, 2.46)</td>
<td>1005/982</td>
<td>1.84 (1.55, 2.20)</td>
<td>55/10</td>
<td>7.57 (3.65, 15.69)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>178/456</td>
<td>4/5</td>
<td>1.42 (0.32, 6.36)</td>
<td>394/444</td>
<td>1.71 (1.31, 2.24)</td>
<td>15/3</td>
<td>7.24 (1.85, 28.38)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>589/1184</td>
<td>21/21</td>
<td>1.30 (0.69, 2.46)</td>
<td>1399/1426</td>
<td>2.09 (1.82, 2.41)</td>
<td>70/13</td>
<td>7.89 (4.23, 14.71)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For each SNP, three rows of results were shown, representing results from the discovery dataset of the Swedish EIRA study (first row), the replication dataset of the Umeå study (second row) and pooled data combining both studies (third row). In EIRA, the odds ratios were adjusted for age, gender, residential area, and shared epitope. In both Umeå and pooled analysis, the odds ratios were adjusted for age, gender and shared epitope. Chr.: chromosome; ref: reference; OR: odds ratio; SNP: single nucleotide polymorphism. The reference group was composed of never-smokers with 0 or 1 risk allele.
Our results reinforced the notion that the chromosome 6 typing, part of the HLA-SE information on Umeå controls alleles with high effects. Due to a lack of classical geno- during the QC stage may have led us to miss rare important SNPs with a very small main effect, but a genome, thus minimizing the risk of missing potentially subset of SNPs, we screened all SNPs over the performing interaction analyses within that particular decisions due to the SNP selection for the ICHIP. The SNPs from the entire genome, which shed light on potential differences due to the SNP selection for the ICHIP. The use of three genetic models provided us with the maximum amount of information concerning potential interactions. Instead of incorporating a screening step to determine the main effect of each SNP, and subsequently performing interaction analyses within that particular subset of SNPs, we screened all SNPs over the genome, thus minimizing the risk of missing potentially important SNPs with a very small main effect, but a significant interaction effect.

Our study also has limitations that need to be acknowledged. Our methods for the elimination of some of the SNPs during the QC stage may have led us to miss rare alleles with high effects. Due to a lack of classical genotyping, part of the HLA-SE information on Umeå controls was imputed and therefore was not as accurate as for the other materials. Additionally, it is possible we missed auxiliary interaction signals outside the chromosome 6p21 locus with smoking, due to the sparse coverage of GWAS data. Although we believe it is feasible that physiological effects from smoking contribute to the deregulation of the immune system and to the development of RA, we cannot exclude potential gene–gene interactions, since smoking habits, to some extent, are genetically influenced [20].

In the current study, all the results identified alleles in chromosome 6p21 only. We did not observe any genome-wide significant interaction effect between smoking and several well-validated loci from previous single SNP association analyses, including TNAIP3, CTLA4, STAT4, PTPN22 and TRAF1/C5 [4], nor for the loci described in smoking candidate gene interaction papers, including PADI4, GSTT1, GSTM1, HMOX1 and EPXH1 [21–23]. Our results reinforced the notion that the chromosome 6 HLA region provides a strong pathogenic background for ACPA-positive RA, driven by smoking [24]. A major novel finding is that several gene variants within the MHC region, in addition to the known HLA-DRB1 locus, interact with smoking, indicating that smoking is involved in the triggering of immune responses that are restricted by different MHC class II, and possibly MHC class I, molecules. The concept is consistent with the findings from purely genetic association studies that several structures in the binding grooves of different MHC classes II and I molecules are associated with an increased ACPA-positive RA risk [25]. A challenging task for further investigations will be to integrate the present findings with detailed information on the structural variants of MHC molecules and to investigate which specific immune reactions are related to smoking and specific MHC molecular structures.

To summarize, we performed a comprehensive analysis of smoking–gene interaction with regard to RA susceptibility. The results should contribute to the designing of further functional studies on the effects of MHC class II-dependent immune activation due to smoking, in ACPA-positive RA. In addition, well-powered genome-wide interaction studies are needed to identify possible gene–smoking interactions in ACPA-negative RA.

**Acknowledgements**

The authors thank Marie-Louise Serra and Lena Nise for their excellent assistance in data collection, and Lena Israelsson for performing the ELISA tests. The authors thank all the investigators and participants from the EIRA study and Umeå study. The EIRA study and Umeå studies were supported by grants from the Swedish Research Council for Health, Working Life and Welfare, the Swedish Research Council, the European Research Council, the Agency for Science & Technology and Research of Singapore, the IMI-supported project BTCu and the Swedish Rheumatism Foundation. X.J. designed the study, carried out data analysis and drafted the tables and manuscript. H.K. provided supervision in data analysis, drafted the plots and revised the manuscript. Z.C. assisted in performing the analysis and drafting the plots, as well as drafting part of the manuscript. LÄ, provided data and actively participated in the study design and coordination. S.D. participated in the planning and execution of the experiments and data management. L.K., L.P. and L.A. are principal investigators of the EIRA study and thus were involved in the study design, coordination and supervision, as well as manuscript revision. S.R.D. is the principal investigator of the Umeå study and was involved in the study design, coordination and supervision. All authors read and approved the final manuscript.

**Funding:** No specific funding was received from any funding bodies in the public, commercial or not-for-profit sectors to carry out the work described in this manuscript.

**Disclosure statement:** The authors have declared no conflicts of interest.
Supplementary data

Supplementary data are available at Rheumatology Online.

References


16. Purcell S, Neale B, Todd-Brown K et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.


