Association of IL23R and ERAP1 genes with ankylosing spondylitis in a Portuguese population

F.M. Pimentel-Santos1,2,3, D. Ligeiro4, M. Matos5, A.F. Mourão3, E. Sousa5,6, P. Pinto7, A. Ribeiro8, M. Sousa9, A. Barcelos10, F. Godinho11, M. Cruz12, J.E. Fonseca5,6, H. Guedes-Pinto2, H. Trindade1,4, D.M. Evans13, M.A. Brown14*, J.C. Branco1,3* and the CORPOREA Study Group

1Faculdade de Ciências Médicas da Universidade Nova de Lisboa, Lisboa, Portugal; 2Instituto de Biotecnologia e Bioengenharia, Centro de Genética e Biotecnologia, da Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal; 3Centro Hospitalar Lisboa Ocidental, Hospital de Egas Moniz, Lisboa, Portugal; 4Centro de Histocompatibilidade do Sul, Lisboa, Portugal; 5Unidade de Investigação em Reumatologia, Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal; 6Hospital de Santa Maria, Lisboa, Portugal; 7Hospital de São Marcos, Braga, Portugal; 8Centro Hospitalar do Alto Minho, Hospital de Ponte de Lima, Ponte de Lima, Portugal; 9Instituto Português de Reumatologia, Lisboa, Portugal; 10Hospital Infantile D. Pedro, Aveiro, Portugal; 11Hospital Garcia de Orta, Almada, Portugal; 12Centro Hospitalar das Caldas da Rainha, Caldas da Rainha, Portugal; 13MRC Centre for Causal Analyses in Translational Epidemiology, Department of Social Medicine, University of Bristol, Bristol, United Kingdom; 14Diamantina Institute for Cancer, Immunology and Metabolic Medicine, Princess Alexandra Hospital, University of Queensland, Brisbane, Australia.

Abstract

Objective

Association between ankylosing spondylitis (AS) and two genes, ERAP1 and IL23R, has recently been reported in North American and British populations. The population attributable risk fraction for ERAP1 in this study was 25%, and for IL23R, 9%. Confirmation of these findings to ERAP1 in other ethnic groups has not yet been demonstrated. We sought to test the association between single nucleotide polymorphisms (SNPs) in these genes and susceptibility to AS among a Portuguese population. We also investigated the role of these genes in clinical manifestations of AS, including age of symptom onset, the Bath Ankylosing Spondylitis Disease Activity, Metrology and Functional Indices, and the modified Stoke Ankylosing Spondylitis Spinal Score.

Methods

The study was conducted on 358 AS cases and 285 ethnically matched Portuguese healthy controls. AS was defined according to the modified New York Criteria. Genotyping of IL23R and ERAP1 allelic variants was carried out with TaqMan allelic discrimination assays. Association analysis was performed using the Cochrane-Armitage and linear regression tests of genotypes as implemented in PLINK for dichotomous and quantitative variables respectively. A meta-analysis for Portuguese and previously published Spanish IL23R data was performed using the StatsDirect® Statistical tools, by fixed and random effects models.

Results

A total of 14 nsSNPs markers (8 for IL23R, 5 for ERAP1, 1 for LN-PEP) were analysed. Three markers (2 for IL23R and 1 for ERAP1) showed significant single-locus disease associations, confirming that the association of these genes with AS in the Portuguese population. The strongest associated SNP in IL23R was rs1004819 (OR=1.4, p=0.0049), and in ERAP1 was rs30187 (OR=1.26, p=0.035). The population attributable risk fractions in the Portuguese population for these SNPs are 11% and 9.7% respectively. No association was seen with any SNP in LN-PEP, which flanks ERAP1 and was associated with AS in the British population. No association was seen with clinical manifestations of AS.

Conclusions

These results show that IL23R and ERAP1 genes are also associated with susceptibility to AS in the Portuguese population, and that they contribute a significant proportion of the population risk for this disease.

Key words

Ankylosing spondylitis, ERAP1, IL23R.

**Introduction**

Ankylosing spondylitis (AS) is a chronic inflammatory disorder with an estimated prevalence of 0.1–0.9% in Caucasian populations (1, 2). Although the contribution of the HLA-B27 allele to the overall genetic predisposition has been estimated at 20-30% and the contribution of all genes in the HLA region estimated at 40-50% (3), genes outside the major histocompatibility complex are strongly implicated in the aetiology of the disease. In particular, the genes *IL23R* and *ERAP1* have recently been demonstrated to be associated with AS in British and North American Caucasians (4). The population attributable risk fraction for *ERAP1* in this study was 25%, and for *IL23R*, 9% (4). Association of the *IL23R* findings has recently been confirmed in Canadian (5) and Spanish populations (6), but not yet been demonstrated for *ERAP1*. We sought to test the association between single nucleotide polymorphisms (SNPs) in these genes and susceptibility to AS in the Portuguese population. We also investigated the role of these genes in the pattern of AS clinical manifestations, including age of symptom onset and the Bath Ankylosing Spondylitis Disease Activity (BASDAI) (7), Functional (BASFI) (8) and Metrology (BASMI) (9) Indices, and the modified Stoke Ankylosing Spondylitis Spinal Score (10).

**Methods**

**Subject**

The study group comprised 358 unrelated AS patients and 285 ethnically matched healthy controls. Individuals included in the study were of Portuguese ancestry and came from mainland Portugal. All cases were diagnosed as having AS according to the modified New York Criteria (11). Cases were recruited from hospital outpatient departments; controls were healthy Portuguese bone marrow donors. This study was approved by the Ethics Board of the involved centres, and written informed consent was obtained from the individuals involved in this study. Patients completed a questionnaire containing a self-assessment of clinical features, including the BASDAI and the BASFI. Metrology was performed by one of the investigators (FS), to obtain the BASMI. Age at disease onset was defined as the age at onset of clinical symptoms. Similarly, disease duration was defined as the period (years) after the onset of clinical symptoms.

**Statistical analysis**

SNP genotype data was assessed for missingness (overall, and differences, cases and controls assessed by \( \chi^2 \) test) and for Hardy-Weinberg equilibrium in controls. Individuals with >10% missingness were excluded (n=5 controls, 2 cases). Association analysis was performed using the Cochrane-Armitage test as implemented in PLINK (12). Association between SNPs and the quantitative variables age of symptom onset, BASDAI, BASFI, BASMI and mSASSS were tested by linear regression assuming an additive model using PLINK, taking into account gender and disease duration as covariates. Imputation analyses were carried out using Markov Chain Haplotyping software.
Association between AS and IL23R and ERAP1 genes in Portugal / F.M. Pimentel-Santos et al.

Table I. IL23R, ERAP1 and LN-PEP genetic variants analysed in AS patients and controls. Positions are given as per dbSNP build 129.

<table>
<thead>
<tr>
<th>Genes</th>
<th>Chromosome</th>
<th>Position (bp)*</th>
<th>Gene location</th>
<th>Taqman ASSAY ID</th>
<th>NCBI SNP Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL23R</td>
<td>1</td>
<td>67442801</td>
<td>Intron</td>
<td>C_1272321_10</td>
<td>rs1004819</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>67460937</td>
<td>Intron</td>
<td>C_30279129_20</td>
<td>rs10489629</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>67475114</td>
<td>Intron</td>
<td>C_31222838_10</td>
<td>rs11465804</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>67491717</td>
<td>Intron</td>
<td>C_8367043_10</td>
<td>rs1343151</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>67497708</td>
<td>Exon</td>
<td>C_11283764_10</td>
<td>rs10889677</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>67487546</td>
<td>Exon</td>
<td>C_1272298_10</td>
<td>rs11209026</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>67526096</td>
<td>Intergenic</td>
<td>C_8361864_10</td>
<td>rs1495965</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>67512680</td>
<td>Intergenic</td>
<td>C_2720238_10</td>
<td>rs11209032</td>
</tr>
<tr>
<td>ERAP1</td>
<td>5</td>
<td>96150086</td>
<td>Intergenic</td>
<td>C_3056885_10</td>
<td>rs30187</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>96147966</td>
<td>Intergenic</td>
<td>C_3056876_10</td>
<td>rs10050860</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>96155291</td>
<td>Intergenic</td>
<td>C_3056893_10</td>
<td>rs2287987</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>96144608</td>
<td>Intergenic</td>
<td>C_3056870_10</td>
<td>rs27044</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>96144622</td>
<td>Intergenic</td>
<td>C_3056871_10</td>
<td>rs17482078</td>
</tr>
<tr>
<td>LN-PEP</td>
<td>5</td>
<td>96376466</td>
<td>Missense Mutation</td>
<td>C_25649482_10</td>
<td>rs2303138</td>
</tr>
</tbody>
</table>

Table II. Characteristics of the Portuguese AS cases and controls*.

<table>
<thead>
<tr>
<th>N (%) males/ N (%) females</th>
<th>AS patients (n= 358)</th>
<th>Controls (n=285)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>226 (63%) / 132 (37%)</td>
<td>127 (46.6%) /158 (55.4%)</td>
</tr>
<tr>
<td>Age, years</td>
<td>45.4 ± 13.3</td>
<td>35.9 ± 11.1</td>
</tr>
<tr>
<td>Disease duration, years</td>
<td>19.1 ± 12.6</td>
<td></td>
</tr>
<tr>
<td>BASDAI</td>
<td>4.2 ± 2.3</td>
<td></td>
</tr>
<tr>
<td>BASFI</td>
<td>4.1 ± 2.7</td>
<td></td>
</tr>
<tr>
<td>BASMI</td>
<td>4.0 ± 2.5</td>
<td></td>
</tr>
<tr>
<td>sMASSS</td>
<td>20.9 ± 22.9</td>
<td></td>
</tr>
</tbody>
</table>

*Except where indicated otherwise, values are the mean ± SD (standard deviation). AS- Ankylosing spondylitis, BASDAI- Bath AS Disease activity Index; BASFI- Bath AS Functional Index; BASMI- Bath AS Metabolic Index.

Table III. Frequency of ERAP1 and LNPEP minor allele frequencies in the Portuguese AS cohort.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Minor Allele</th>
<th>Case MAF</th>
<th>Control MAF</th>
<th>OR</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERAP1</td>
<td>rs27044</td>
<td>G</td>
<td>0.37</td>
<td>0.32</td>
<td>1.26 (1.10-1.60)</td>
<td>0.044</td>
</tr>
<tr>
<td>ERAP1</td>
<td>rs17482078</td>
<td>T</td>
<td>0.16</td>
<td>0.19</td>
<td>0.80 (0.76-0.83)</td>
<td>0.096</td>
</tr>
<tr>
<td>ERAP1</td>
<td>rs10050860</td>
<td>T</td>
<td>0.16</td>
<td>0.20</td>
<td>0.76 (0.71-0.82)</td>
<td>0.057</td>
</tr>
<tr>
<td>ERAP1</td>
<td>rs30187</td>
<td>T</td>
<td>0.47</td>
<td>0.41</td>
<td>1.26 (1.10-1.57)</td>
<td>0.035</td>
</tr>
<tr>
<td>ERAP1</td>
<td>rs2287987</td>
<td>C</td>
<td>0.16</td>
<td>0.20</td>
<td>0.77 (0.75-0.88)</td>
<td>0.074</td>
</tr>
<tr>
<td>LNPEP</td>
<td>rs2303138</td>
<td>A</td>
<td>0.08</td>
<td>0.07</td>
<td>1.23 (0.80-1.88)</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Both fixed and random effects analysis was performed; non-combinability of studies was assessed using the Cochran Q statistic, and the extent of heterogeneity between studies assessed using the I² statistic. Power calculations were performed using the Genetic Power Calculator (14).

Results
The Portuguese AS cohort population (Table II) included 228 (63.5%) men and 132 (36.5%) women with a mean age of 45.4 (±13.2) years (range 20-79 years) and a mean disease duration of 19.1 (±12.6) years (range 0-60 years), of whom 82% were HLA-B27 positive. Epidemiological data of the cases and controls are summarized in Table II.

All the studied genetic markers were in Hardy-Weinberg equilibrium in the controls group, had missings rates <10%, and none had differential missingness in cases and controls (p<10⁻²). The minor allele frequencies (MAF) of the 14 SNPs are presented in Tables III and IV.

Single-marker association tests revealed significant AS associations for the ERAP1 SNP rs30187 (odds ratio (OR) 1.26, p=0.035) and rs27044 (OR 1.26, p=0.044) (Table III). Unlike in the previous British and North American studies, no association was seen with the SNP rs2303138, which lies within LN-PEP, which flanks ERAP1. Of the 119 SNPs imputed in and around ERAP1, 32 demonstrated nominal association with AS (p<0.05), with the strongest association being with rs41135 (p=0.014) (Fig. 1).

Two SNPs in and around IL23R demonstrated significant association (rs1004819 OR=1.44, p=4.9×10⁻³; rs10889677 OR=1.41, p=5.8×10⁻³) (Table IV). The strongest associated SNP reported in both Crohn’s disease (15) and psoriasis (16) (Arg381Gln; rs11209026) did not show any protective effect in our population. No imputed SNP was more strongly associated with AS than these two genotyped SNPs, but many SNPs in a block extending from rs10889667 as far as rs11465817 (67493685 bp from the p-telomere) were associated with AS with p<0.01 (Fig. 2). Of the 49 imputed SNPs in and around IL23R, 23 demonstrated nominal association with AS (p<0.05).
The population attributable risk fraction in the Portuguese population for rs1004819 is 11%, and for rs30187 is 9.7%.

The meta-analysis performed between Portuguese and Spanish populations (Table V) revealed significant AS associations, through the fixed effects, for SNPs rs1004819, rs11209026, and rs1343151. However, for markers rs11209026, rs1343151 and rs210889617 in particular, there were significant differences in the findings in the Spanish and Portuguese populations, reflected by significant Cochran Q statistics (rs11209026, \(p = 0.017\); rs1343151, \(p = 0.023\); rs10889617, \(p = 0.0005\)), suggesting that a random effects model should be applied. Considering the random effects models, the associations with AS remain significant for rs1004819.

No association was observed between *IL23R* or *ERAP1* variants and age of symptom onset, BASDAI, BASFI, BASMI or mSASSS (data not shown).

The study had 80% power to detect associations with these quantitative variables at a significance level of \(\alpha = 0.05\) for SNPs contributing >4% of the trait variance, assuming linkage disequilibrium between the marker and disease-associated variant of \(D' > 0.8\) and that the marker and disease-associated allele frequencies are equal. Considering the case-control analysis of disease susceptibility, assuming a population prevalence of disease of 0.4%, minor allele frequencies of 0.1–0.5, and \(D' > 0.8\), the study had 80% power to detect an additive association with heterozygote odds ratio of 1.6-1.8.

**Discussion**

Many candidate genes outside the MHC have been evaluated in different studies regarding AS susceptibility and/or phenotype associations. Recently, association has been demonstrated and confirmed with SNPs in and around the genes *ERAP1* and *IL23R* in British and North American populations (4).

The *IL23R* association with AS was also recently replicated in the Canadian (5) and Spanish (6) populations. The present study has replicated this association in the Portuguese population. The peak association in our cohort is seen with rs1004819. This is different from the UK, United States, Canada and Spanish data sets, where the peak association was observed for different SNPs, although the minor allele frequencies (MAF) that we have observed for SNPs in Portuguese were similar to those reported in other British and North Americans. Furthermore, the association observed in the Portuguese population had a similar magnitude of effect to the one described in those other populations, as can be appreciated by the attributable risk for rs1004819, which is very similar to the one reported for the most strongly associated SNP (rs11020932) in the British/North American populations (4). Interestingly, no association was established with rs11209032 in Portuguese or Spanish populations, which was strongly associated with AS in other Caucasian populations (4). Consistent with the Alberta

**Table IV.** Frequency of *IL-23R* minor allele frequencies in the Portuguese and Spanish cohorts.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Minor Allele</th>
<th>Portuguese Cohort</th>
<th>Spanish Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Case MAF</td>
<td>Control MAF</td>
<td>OR</td>
</tr>
<tr>
<td><em>IL23R</em></td>
<td>rs1004819</td>
<td>A</td>
<td>0.35</td>
<td>0.27</td>
</tr>
<tr>
<td><em>IL23R</em></td>
<td>rs7517847</td>
<td>C</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><em>IL23R</em></td>
<td>rs10489629</td>
<td>C</td>
<td>0.43</td>
<td>0.45</td>
</tr>
<tr>
<td><em>IL23R</em></td>
<td>rs11465804</td>
<td>G</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td><em>IL23R</em></td>
<td>rs11209026</td>
<td>A</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td><em>IL23R</em></td>
<td>rs1343151</td>
<td>A</td>
<td>0.33</td>
<td>0.34</td>
</tr>
<tr>
<td><em>IL23R</em></td>
<td>rs10889677</td>
<td>A</td>
<td>0.28</td>
<td>0.36</td>
</tr>
<tr>
<td><em>IL23R</em></td>
<td>rs11209032</td>
<td>A</td>
<td>0.35</td>
<td>0.32</td>
</tr>
<tr>
<td><em>IL23R</em></td>
<td>rs1495965</td>
<td>C</td>
<td>0.45</td>
<td>0.45</td>
</tr>
</tbody>
</table>

The population attributable risk fraction in the Portuguese population for rs1004819 is 11%, and for rs30187 is 9.7%.

The meta-analysis performed between Portuguese and Spanish populations (Table V) revealed significant AS associations, through the fixed effects, for SNPs rs1004819, rs11209026, and rs1343151. However, for markers rs11209026, rs1343151 and rs210889617 in particular, there were significant differences in the findings in the Spanish and Portuguese populations, reflected by significant Cochran Q statistics (rs11209026, \(p = 0.017\); rs1343151, \(p = 0.023\); rs10889617, \(p = 0.0005\)), suggesting that a random effects model should be applied. Considering the random effects models, the associations with AS remain significant for rs1004819.

No association was observed between *IL23R* or *ERAP1* variants and age of symptom onset, BASDAI, BASFI, BASMI or mSASSS (data not shown).
both the discovery and replication set of the study by the WTCCC/TASC in the British and North Americans (4), and in marked contrast to those reported in the Spanish population (6). We observed a MAF in cases and controls of 34% and 28% for this SNP respectively, compared with 31% and 35% in the Spanish population. In the WTCCC/TASC study, the MAF for this SNP in cases and controls respectively were 36% and 31% in the discovery cohort, and 37% and 29% in the control cohort. Association was seen between rs10889677 and AS in the Spanish study (OR=0.81 with minor allele ‘A’, P=0.039), but curiously it was in the opposite direction to the two cohorts reported by the WTCCC/TASC study, and our own Portuguese study.

Numerous studies have demonstrated association of IL23R SNPs with susceptibility to Crohn’s disease (CD) (17), as well as to psoriasis (18) and psoriatic arthritis (19-21). IL23R therefore seems to be a common susceptibility factor for the major seronegative diseases, at least partially explaining their co-occurrence. In contrast, ERAP1 is not associated with inflammatory bowel disease, and it is unknown whether it is associated with psoriasis or psoriatic arthritis. Whether IL23R or ERAP1 polymorphisms influence clinical manifestations of disease such as age of symptoms onset, disease activity or severity is unknown. In the current study no association was observed with these traits, but the study power was only adequate for large genetic effect sizes (>4% of the trait variance).

This study confirms the association of ERAP1 variants with AS, with a similar magnitude of effect to that seen for IL23R, as assessed by the population attributable risk fraction. We are not aware of other papers that have replicated this finding in populations other than British or North Americans. The strongest associated in ERAP1 was rs27044 (OR=1.29, 95% CI=1.02-1.63; p=0.032). In contrast with IL23R, the association of ERAP1 seems to be confined to AS. No association was observed between ERAP1 SNPs and either Crohn’s disease or ulcerative colitis in the WTCCC/TASC study (4). Whether
ERAPI SNPs are associated with psoriasis or psoriatic arthritis is unknown. No association was seen with the marker rs2303138 lying in LNPEP, providing further support to the hypothesis that at least a component of the association observed between this SNP and AS previously reported in British Caucasians is due to linkage disequilibrium with ERAPI polymorphisms. The primary associated variant(s) in ERAPI remain uncertain. In this study, nominal association was seen between the SNPs rs26509 (96108436 from the p-telomere) and rs190298 (96191986 from the p-telomere), an interval of 84 kb. Very broad association was seen in the WTCCC/TASC study as well. Further fine-mapping and resequencing studies will be required to narrow the associated region and identify the key associated variant(s) to inform more targeted functional analysis of the mechanisms of involvement of this gene and AS.

Conclusions
These results show that IL23R and ERAPI genes are also associated with susceptibility to AS in the Portuguese population, and that they contribute a significant proportion of the population risk for this disease.

Authors’ contributions
FPS, DL, HT, HGP, MAB and JCB participated in the design of the study. Experiments were performed by FPS, DL and MM. Statistical analysis was carried out by FPS, DME and MB, FPS, AFM, ES, PP, AR, MS, AB, FG and MC contributed by providing human samples. Analysis of data was carried out by FPS, DL, MM and MB. Intellectual contributions to the manuscript were provided by FPS, JEF, HGP, MAB and JCB. All authors read and approved the final manuscript.

Acknowledgements
We would like to thank the individuals who shared their clinical data and blood samples with us to complete this study and to ANEA (Ankylosing Spondylitis Portuguese Patients Association).

References
13. StatsDirect: www.statsdirect.com
15. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447: 661-78.
APPENDIX

CORPOREA Study Group

Centro Hospitalar de Lisboa Ocidental, Hospital de Egas Moniz, Lisboa
António Alves de Matos
Célia Ribeiro
José Bravo Pimentão
Margarida Mateus
Patricia Nero
Paula Araújo
Rita Barros
Sandra Falcão
Teresa Laura Pinto
Walter Castelão

Unidade de Investigação em Reumatologia, Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa; Joana Caetano-Lopes

Instituto Português de Reumatologia, Lisboa
Cândida Silva
Eugénia Simões
Helena Madeira
Helena Santos
José Vaz Patto
Julia Ferreira
Manuela Micaelo
Maria Jesus Mediavilla

Hospital Curry Cabral, Lisboa
Pedro Soares Branco

Hospital Garcia de Orta, Almada
José Canas da Silva
Sandra Garcês
Viviana Tavares

Centro Hospitalar do Alto Minho, Hospital de Ponte de Lima, Ponte de Lima
Domingos Araújo
José A. Costa
Lucía Costa
Maria Carmo Afonso
Mónica Bogas
Sergio Alcino

Hospital Infante D. Pedro, Aveiro
Inês Cunha

Hospital de São Marcos, Braga
José Redondo

Hospital Central de Faro, Faro
Ana Rita Cravo
Graça Sequeira

Hospital Militar Principal, Lisboa
Rui André Santos

Gabinete Investigação Clínica (GIC) / Hospital de Dia, Instituto Português de Oncologia Francisco Gentil (IPOFG), Lisboa
M. Elisabete Carvalho