Genetic polymorphisms inside and outside the MHC improve prediction of AS radiographic severity in addition to clinical variables

Nerea Bartolomé1,*, Magdalena Szczypiorska1,*, Alejandra Sánchez2, Jesús Sanz2, Xavier Juanola-Roura3, Jordi Gratacos4, Pedro Zarco-Montejo5, Eduardo Collantes6, Antonio Martínez1, Diego Tejedor1, Marta Artieda1 and Juan Mulero2

Abstract

Objective. The aim of this study was to analyse if single nucleotide polymorphisms (SNPs) inside and outside the MHC region might improve the prediction of radiographic severity in AS.

Methods. A cross-sectional multi-centre study was performed including 473 Spanish AS patients previously diagnosed with AS following the Modified New York Criteria and with at least 10 years of follow-up from the first symptoms of AS. Clinical variables and 384 SNPs were analysed to predict radiographic severity [BASRI-total (BASRI-t) corrected for the duration of AS since first symptoms] using multivariate forward logistic regression. Predictive power was measured by the area under the receiver operating characteristic curve (AUC), specificity, sensitivity, positive predictive value (PPV) and negative predictive value (NPV).

Results. The model with the best fit measured radiographic severity as the BASRI-t 60th percentile and combined eight variables: male gender, older age at disease onset and six SNPs at ADRB1 (rs1801253), NELL1 (rs8176785) and MHC (rs1634747, rs9270986, rs7451962 and rs241453) genes. The model predictive power was defined by AUC = 0.76 (95% CI 0.71, 0.80), being significantly better than the model with only clinical variables, AUC = 0.68 (95% CI 0.63, 0.73), P = 0.0004. Internal split-sample analysis proved the validation of the model. Patient genotype for SNPs outside the MHC region, inside the MHC region and clinical variables account for 26, 38 and 36%, respectively, of the explained variability on radiographic severity prediction.

Conclusion. Prediction of radiographic severity in AS based on clinical variables can be significantly improved by including SNPs both inside and outside the MHC region.

Key words: ankylosing spondylitis, gene polymorphism, autoimmune diseases.

Introduction

AS is a chronic inflammatory rheumatic disease characterized by sacroiliitis, ankylosis of the spine and enthesitis. Although the aetiology of AS is unknown, several studies have demonstrated that it is highly heritable and that genetic factors contribute >90% to AS susceptibility [1]. Until now, HLA-B27 has been considered as the major genetic contributor to the disease. However, familial studies suggest that the contribution of HLA-B27 to the genetic risk for AS may not be >16–50% [2]. It has become evident that additional genetic factors influence AS development [3]. Genome-wide association studies have revealed association with AS for SNPs in the non-MHC genes ERAP1 (endoplasmic reticulum...
aminopeptidase 1) and IL-23 receptor (IL-23R), as well as, in two gene deserts, 2p15 and 21q22 [4, 5]. These associations have been consistently replicated in populations of different geographic origin [6–10]. Association with other genes like IL-1 receptor, type II (IL-1R2), anthrax toxin receptor 2 (ANTXR2), TNF superfamily, member 15 (TNFSF15), TNF receptor 1 (TNFR1) and TNF receptor type 1-associated DEATH domain protein (TRADD) has also been described [4, 5, 11].

Although it is well established that susceptibility to AS is largely genetically determined, little is known about the involvement of genetics in the radiographic severity of AS. Familial studies have demonstrated a high heritability, ~60% of radiographic severity of AS, suggesting a key role for genetics in this process [12]. Structural damage in AS may lead to fusion of the spine and impairment of spinal mobility [13], which considerably reduces the patient’s quality of life. Though the molecular basis of radiographic damage is not completely understood, there is increasing evidence that the process could be driven by the interaction between inflammation and new bone formation pathways [14, 15]. Revealing the genetic basis of radiographic severity in AS could be of great value to identify, at the moment of diagnosis, patients at high risk of radiographic damage. This could allow clinicians to select and optimize the patients’ preventive and therapeutic approach. Previous studies have shown that several clinical factors have prognostic significance in the radiographic status of the disease. Male sex and hip involvement have been associated with poorer radiographic condition [16–18]. Smoking has also been described as having a strong influence in radiographic severity [19]. However, few association studies have been performed to detect the specific genes associated with the radiographic severity of AS, and several of them have shown negative results [20, 21]. To our knowledge, only two studies have found a significant association of genetic markers with radiographic severity. On one hand, association of an SNP haplotype in the VEGF gene with radiographic severity of AS was described by Seo et al. [22], but the association was lost after correction for disease duration. On the other hand, a recent study performed by Ward et al. [23] showed association of MHC alleles with the spine radiographic severity of AS using the 75th percentile as the threshold for severe damage.

We hypothesized that, as occurs with AS susceptibility, genetic factors outside the MHC region could also be involved in AS radiographic severity. In the current study we aimed to investigate if both SNPs inside and outside the MHC region, some of them in genes related to inflammation and bone formation pathways, might improve the prediction for severity of total radiographic damage in a group of Spanish AS patients.

Methods

Study population

We carried out a cross-sectional association study with 473 Spanish AS patients. The patients, who fulfilled the Modified New York Criteria for AS diagnosis [24] and have at least 10 years of follow-up from the first symptoms of the disease, were recruited from 25 hospitals participating in the National Spondyloarthropathies Registry of Spain (REGISPONSER) [25]. All patients gave their informed consent to participate in the study, which was centrally approved by the Ethics Committee of the Puerta de Hierro Hospital.

To assess the disease structural damage of the patients, we used the total BASRI (BASRI-t) score [26]. Evaluation of the radiographs for assessment of the BASRI-t score was performed in each centre by an expert rheumatologist, previously trained in a 2-day session. The baseline clinical and demographic characteristics of the patients at the beginning of the disease were recovered from the Regisponser database, as possible prognostic indicators [25].

HLA-B27 typing and SNP genotyping

Genomic DNA was isolated from saliva samples using a DNA Self-Collection kit (DNA Genotek Inc., Ottawa, ON, Canada), following the manufacturer’s extraction protocol. All the samples were tested for the presence of the HLA-B27 allele by conventional PCR [10].

We selected 384 SNPs distributed in 190 genes to be analysed in this study. We selected SNPs previously described as associated with AS, to other SpA, to autoimmune- and bone-related diseases and SNPs from the metabolic pathways of the IL-23R and ERAP1 genes. SNPs genotyping was performed using the Illumina GoldenGate genotyping platform (Illumina, Inc., San Diego, CA, USA) [27].

Statistical analysis

To standardize the measure of radiographic severity, we corrected BASRI-t for duration of AS from the onset of the first symptoms of the disease (BASRI-t/duration) [22, 23]. We performed several association analyses with different criteria to define the radiographically severe group, based on clinicians’ opinion who estimated that 25–40% of AS patients develop severe radiographic damage. Specifically, we considered the patients in the top 75th (p75), 70th (p70), 65th (p65) or 60th (p60) percentile of BASRI-t/duration as having severe radiographic damage. The cut-off values for p60, p65, p70 and p75 of BASRI-t/duration were 0.362, 0.380, 0.410 and 0.447, respectively.

A test for deviation from Hardy–Weinberg equilibrium (HWE) was performed for each SNP in the case and control populations using the Helix Tree software version 6.4.2 (Golden Helix Inc., Bozeman, MT, USA). Pruning of the initial data set with standard parameters [28] (exclusion of SNPs with poor genotype cloud clustering, of SNPs with call rate <85%, of SNPs with severe deviation from HWE (P < 0.0001), and of samples with call rate <85%) resulted in 456 samples and 344 SNPs.

An association test for SNP allele frequencies with BASRI-t/duration was performed by the chi-square test. An association test for clinical variables was performed by
the chi-squared test for categorical variables and by the unpaired t-test for continuous variables. We developed multivariate forward logistic regression models for predicting radiographic severity (75th, 70th, 65th and 60th percentile of BASRI-t/duration) using as predictors the most significant SNPs and clinical variables, ranked by the associated P-value (P < 0.1).

We tested our model’s discrimination via the Hosmer–Lemeshow statistic and the receiver operating characteristic (ROC) curve with 95% CIs. To analyse the predictive power added to the model by the genetic variables we compared the AUC of the model based only on clinical variables with that of the model based on clinical variables and SNPs using Analyse-it software v2.09 (Analyse-it Software Ltd, Leeds, UK).

The model with the best fit was obtained for the BASRI-t/duration p60 and was internally validated by split-sample validation. The study data set was split (10 times) into one data set to train or build the model (training set 75% of the population) and another data set to validate it (validation set 25% of the population) [29, 30]. The average of the 10 training and validation ROC AUCs were compared by means of a Z-test, as described by Hanley and McNeil [31].

To describe the relative contribution of each risk factor’s predictive ability in the radiographic severity model, we constructed an X-pie chart. In order to determine the contributing chi-square value of each predictive factor, the model was calculated by eliminating one predictive factor at time [32, 33]. The reduced model chi-square value was recorded for each variable. Finally, each variable contribution was compared with the sum of all variables contribution to identify the relative contribution of each risk factor’s predictive strength in the multivariate model (supplementary Fig. S1, available as supplementary data at Rheumatology Online). The statistical analysis was performed with the SPSS v15.0 (SPSS, Chicago, IL, USA) and Helix Tree version 6.4.2 (Golden Helix Inc.) software. P < 0.05 was considered statistically significant.

Results

Clinical and demographic variables associated with radiographic severity

The study cohort included 456 AS patients (348 males and 108 females) with a mean (S.E.) age at disease onset of 26.06 (9.1) years, and an average time of evolution from disease onset of 24.7 (10.1) years. A total of 84.6% of the patients were HLA-B27 positive and 19.3% had a family history of SpA.

As a representative example of the association of the demographic and clinical variables with BASRI-t severity, the data corresponding to BASRI-t/duration p60 are shown in Table 1. We observed that male gender, older age at disease onset and neck pain at disease onset were strongly associated with radiographic severity when the patients were classified according to BASRI-t/duration p60 (Table 1) and to any of the rest of the percentiles, p65, p70 and p75 (data not shown). The number of clinical manifestations at disease onset was also found to confer risk for severity of BASRI-t/duration when we considered p60 (P = 0.002), p65 (P = 0.028) and p70 (P = 0.029). A borderline association was found for inflammatory low back pain in the case of BASRI-t/duration p60 and p65 (P = 0.081 and P = 0.065, respectively). This association reached statistical significance for BASRI-t/duration p70 (P = 0.005) and 75 (P = 0.016). The slight association noticed for HLA-B27 with BASRI-t/duration p60 (P = 0.047) was not observed for the other BASRI-t/duration percentiles (data not shown).

Model to predict radiographic severity

To search for a predictive model of radiographic severity, the clinical and demographic variables associated with AS were entered together with the most associated SNPs into multivariate forward logistic regression analysis. In all the cases, BASRI-t/duration p60, p65, p70 and p75, we obtained a model with a ROC AUC > 0.75 (data not shown). The model with the best fit was obtained when the BASRI-t/duration p60 was used as the criteria to classify the patients in the severe or mild group (Fig. 1 and Table 2). The univariate association with BASRI-t/duration p60 of the SNPs included in the predictive model is shown in Table 3. The univariate association with BASRI-t/duration p60 of the complete list of investigated SNPs is summarized in supplementary Table S1, available as supplementary data at Rheumatology Online. The validation of the predictive model was demonstrated by internal split-sample analysis. Thus our predictive model allows for the identification of patients with a future BASRI-t/duration p60 index > 0.447. As a practical example, if we considered a prediction in 10 years’ time, we could predict at the initial stages of the disease which patients would have a future BASRI-t score > 0.447 in 10 years.

The internal validated model, p60, includes two clinical variables, male gender and older age at diagnosis and six SNPs, and has a predictive power, as indicated by the ROC AUC, of 0.76 (95% CI 0.71, 0.80) (Fig. 1). Four of the six SNPs, rs1634747, rs9270986, rs7451962 and rs241453, are located within the MHC and the other two, rs1801253 and rs8176785, within the (k)adrenergic receptor (ADRB1) and nel-like 1 precursor (NELL1) genes (Table 2). We did not find any of the four SNPs in the MHC to be in linkage disequilibrium. Regarding the inherited model of association, a dominant model was found for the SNPs rs1634747 and rs7451962, and a recessive model for the SNPs rs9270986, rs1801253, rs8176785 and rs241453.

The model based on SNPs plus clinical variables (full model) significantly improved the predictive value for radiographic severity compared with the model based only in clinical variables (reduced model) (P = 0.0004). Specifically, an 8% increase in the AUC was registered (Fig. 1). This result indicates that the genetic component plays an important role in AS radiographic severity. The sensitivity, specificity and positive predictive value and negative predictive value of the full and reduced...
models at different points of the ROC curves confirm the significant improvement of the full model over the model without SNPs (Table 4). For example, the reduced model has a sensitivity of 20% for a specificity of 95%, whereas the sensitivity of the full model increases by 15% (35%) at the same value of specificity (95%) (Table 4).

The relative contribution of each genetic and clinical variable to the predictive power of the full model was calculated (supplementary Fig. S1, available as supplementary data at Rheumatology Online). Older age at disease onset and the SNP rs1801253 of the ADRB1 gene are the variables with a highest contribution to the predictive power of the model, 25 and 20% respectively. In general, the clinical variables contribute 36% to the predictive strength of the full model, the remaining 64% being explained by the genetic variables (26% explained by SNPs at genes outside the MHC region and 38% by SNPs inside the MHC region).

Discussion

Radiographic damage in the SI and spine is a hallmark of AS that ultimately leads to impaired spinal mobility and function, and considerably reduces the quality of life of patients. In this study we have developed a logistic regression model that combines genetic and baseline clinical variables to predict the radiographic severity of AS. We demonstrate for the first time that SNPs in the ADRB1 and NELL1 genes, together with SNPs in the MHC, improve the prediction ability of clinical variables. Predicting the AS severity at an early stage of the disease by means of an easy tool, as a clinical-genetic predictive model, could be of great value for clinicians to follow a personalized approach to prevention and care in each individual.

Our model predicts total radiographic damage in AS, measured as BASRI-t/duration (spine and hips), and includes two clinical and six genetic variables (Fig. 1 and Table 2). The two clinical variables are male sex and older age at disease onset. This result is in agreement with previous studies that have shown that male patients have worse radiographic outcomes than female patients [17, 18] and that older age at disease onset is associated with more severe radiographic changes in AS [23, 34]. To our knowledge, there is only one previous study, published during the development of this work, that has presented a predictive model for AS radiographic severity combining clinical and genetic variables [23].
Specifically, Ward et al. [23] presented a model that combines six alleles of the MHC with three clinical variables (male gender, older age at disease onset and smoking status) for predicting radiographic severity of the spine, measured as BASRI-spine, in AS patients. Thus our results add further evidence to the thought that MHC is involved not only in AS susceptibility, but also in AS severity. The six MHC alleles described by Ward et al. [23] correspond to the genes HLA-B, DRB1, DQA1, DPB1 and DPB1. Three of the MHC SNPs found in our study are located in genes coincident with those found by Ward et al. [23] (HLA-B, DRB1 and DQA1). However, it cannot be concluded that the SNPs described by us and the MHC alleles found by Ward et al. [23] represent the same associations, since typically a unique SNP, or even combinations of two or three SNPs, are not enough to identify MHC alleles. Together with the gender and the age at disease onset, the smoking status is also a predictor for spine AS severity in the model developed by Ward et al. [23]. We did not analyse the predictive value of smoking status, since those data were not available in our cohort. Although not remaining in our predictive model, we show for the first time that neck pain and number of clinical manifestations at disease onset are clinical variables associated with radiographic severity (Table 1).

Among the six genetic variables in our model, there are four SNPs located in the MHC region (Table 2), which have been previously described as being associated with other autoimmune diseases, but not with AS susceptibility or radiographic progression. One of them, rs9270986, is located upstream of the HLA-DRB1 gene and has been shown to be associated with susceptibility to Type 1 diabetes.

The ROC curves of the full (clinical variables plus SNPs) and reduced (clinical variables) models.

### Table 2: Multivariate logistic regression analysis to predict severe radiographic damage (BASRI-t/duration p60)

<table>
<thead>
<tr>
<th>Variables (full model)</th>
<th>Chromosome</th>
<th>Gene (NCBI Build 36)</th>
<th>Nucleotide change</th>
<th>Risk allele</th>
<th>Risk genotype</th>
<th>OR (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male gender</td>
<td>2.99</td>
<td>(1.71, 5.21)</td>
<td>1.16E-04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Older age of AS onset</td>
<td>1.08</td>
<td>(1.05, 1.11)</td>
<td>1.32E-08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1634747</td>
<td>6</td>
<td>HLA-B</td>
<td>A/G</td>
<td>A</td>
<td>AA + AG</td>
<td>2.19 (1.35, 3.55)</td>
<td>0.0016</td>
</tr>
<tr>
<td>rs9270986</td>
<td>6</td>
<td>Near to HLA-DRB1</td>
<td>A/G</td>
<td>A</td>
<td>AA</td>
<td>7.40 (2.10, 26.11)</td>
<td>0.0019</td>
</tr>
<tr>
<td>rs7451962</td>
<td>6</td>
<td>Near to HLA-DQA1</td>
<td>A/G</td>
<td>G</td>
<td>GG + AG</td>
<td>2.47 (1.33, 4.56)</td>
<td>0.0040</td>
</tr>
<tr>
<td>rs1801253</td>
<td>10</td>
<td>ADRB1 (Gly389Arg)</td>
<td>G/C</td>
<td>C</td>
<td>CC</td>
<td>2.83 (1.34, 5.98)</td>
<td>0.0063</td>
</tr>
<tr>
<td>rs8176785</td>
<td>11</td>
<td>NELL1 (Arg82Gln)</td>
<td>A/G</td>
<td>A</td>
<td>AA</td>
<td>1.90 (1.17, 3.08)</td>
<td>0.0091</td>
</tr>
<tr>
<td>rs241453</td>
<td>6</td>
<td>TAP2 (intron)</td>
<td>A/G</td>
<td>G</td>
<td>GG</td>
<td>2.36 (1.49, 3.76)</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

The variables included in the predictive model, their OR (95% CI) and P-value are shown.

### Table 3: Univariate association with radiographic severity of the six SNPs included in the predictive model (BASRI-t/duration p60)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>Chromosome</th>
<th>Position</th>
<th>Allele (P-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs8176785</td>
<td>NELL1</td>
<td>11</td>
<td>20 761 862</td>
<td>0.038</td>
</tr>
<tr>
<td>rs1801253</td>
<td>ADRB1</td>
<td>10</td>
<td>115 795 046</td>
<td>0.039</td>
</tr>
<tr>
<td>rs9270986</td>
<td>HLA-DRB1</td>
<td>6</td>
<td>32 682 038</td>
<td>0.001</td>
</tr>
<tr>
<td>rs1634747</td>
<td>HLA-B</td>
<td>6</td>
<td>31 389 855</td>
<td>0.001</td>
</tr>
<tr>
<td>rs7451962</td>
<td>HLA-DQA1</td>
<td>6</td>
<td>32 690 413</td>
<td>0.061</td>
</tr>
<tr>
<td>rs241453</td>
<td>TAP2</td>
<td>6</td>
<td>32 904 204</td>
<td>0.055</td>
</tr>
</tbody>
</table>
diabetes (T1D) and RA [4, 35]. The second one, rs7451962, is located near HLA-DQA1 and was found to be associated with multiple sclerosis (MS) [36].

rs7451962 is a key candidate gene for association with autoimmune diseases, since it is involved in the antigen presentation process. This is the first report that investigates and demonstrates the association of TAP2 SNPs with radiographic severity of AS. TAP2 has been previously shown to be associated with AS susceptibility in a Chinese cohort [38], but no association has been found in Caucasian [39] or Japanese patients [40]. TAP2 is also involved in MS [41] and T1D susceptibility in Caucasian patients [42]. Taken together, our results add evidence to the previous findings of Ward et al. [23], which showed an association between MHC alleles and AS spine radiographic severity, and suggest that the MHC is not only a pivotal genetic factor for AS susceptibility, but also for AS radiographic severity.

One of the two non-MHC SNPs in the predictive model is the non-synonymous SNP, rs8176785 (R82Q), located in NELL1 gene, which encodes a newly identified potent osteoinductive protein [43–45]. NELL1 was discovered in the sutilves of infants with craniosynostosis [46], and it has been recently demonstrated that it can significantly improve spinal fusion rates in rat and sheep models by promoting adequate new bone formation [45]. Spinal fusion is the clinical and radiological hallmark of AS, we hypothesize that SNPs in the NELL1 gene could affect the function of this enzyme, altering bone formation and leading to bone fusion in AS. This is the first evidence of association of SNPs in NELL1 with AS radiographic severity; additional research in this field would be of great value to confirm the results and to explain the functional role of NELL1 in radiographic severity of AS. NELL1 gene has also been associated with the autoimmune disorder IBD [47].

The SNP rs1801253 (G389R) is the other non-MHC SNP in the predictive model and is located in the ADRB1 gene. To our knowledge there are no previous associations between the ADRB1 gene and AS or other SpAs. The ADRB1 gene is mostly known for its role in the regulation of cardiac output, peripheral resistance and obesity. However, several studies have suggested a role for adrenergic signalling in the regulation of bone remodelling [48, 49]. The molecular mechanisms by which ADRB1 regulates bone turnover remains to be elucidated, but it has been suggested that it could stimulate bone formation systemically, since it is not well expressed in bone, and that its effects could be mediated by the growth hormone–IGF-1 axis [48, 50]. Clearly, further research is needed in this area to establish the link between SNPs in the ADRB1 gene, the implication of ADRB1 in the regulation of bone formation and the radiographic severity of AS.

We are aware of the limitations of our study, coming mainly from the lack of an independent cohort to validate the observed clinical/genetic associations with radiographic severity. Further research in other cohorts would be of great value to corroborate the univariate SNP associations and to discard the possibility of spurious associations due to multiple testing. Our predictive model has been internally validated by split-sample analysis; however, a validation study in an independent cohort would also be of interest to confirm the results. Another limitation of this study comes from the fact that we do not have reliable data about the therapy strategies of the patients before their inclusion in the study. Although the impact of NSAID(s) and anti-TNF treatment in AS radiographic severity remains a controversial aspect, some studies have found evidence of disease-modifying action for these agents [51–53]. Since we do not have reliable treatment information for our cohort, the impact of the treatment in the final results cannot be completely excluded; therefore a validation in an independent cohort would be desirable.

To summarize, in this article we present a predictive model for radiographic severity of AS based on the combination of baseline clinical variables and SNPs in the MHC and the ADRB1 and NELL1 genes, which could aid clinicians in identifying high-risk patients at early stages of the disease, allowing for a personalized approach to prevention and care in each individual. Our results, in a Spanish population, add evidence to the thought that genetics plays an important role in AS radiographic severity and that genomics is the basis towards a personalized medicine approach in AS prognosis. Whether our predictive model could be helpful in other ethnically different populations needs to be explored.

### Rheumatology key messages

- We present a predictive model for AS radiographic severity that combines clinical and genetic variables.
- SNPs in non-MHC genes have been associated with AS radiographic severity for the first time.
- Our results support that genetics plays an important role in AS radiographic severity.
SNPs improve prediction of AS radiographic severity

Acknowledgements

We are grateful to Iratxe Ciriza and Jokin del Amo for helpful suggestions at the beginning of the study, to Rute Rodrigues for help in the writing process of the manuscript and to all the clinicians and patients who have participated in the study. The work presented in this report is the subject of a pending patent filed by Progenika Biopharma SA.

Funding: This work was supported by the Ministerio de Ciencia e Innovación of Spain (Project PSE-01000-2006-1) and by Progenika Biopharma SA.

Disclosure statement. M.S., A.M., N.B., D.T. and M.A. are employees of Progenika Biopharma SA. A.S. is supported by an unrestricted grant from Pfizer. All other authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at Rheumatology Online.

References

35 Nejentsev S, Howson JM, Walker NM et al. Type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. Nature 2007;450:887–92.