CONCISE REPORT

Ankylosing spondylitis is associated with the anthrax toxin receptor 2 gene (ANTXR2)

T Karaderi,1 S M Keidel,1 J J Pointon,1 L H Appleton,2 M A Brown,3 D M Evans,3,4 B P Wordsworth1,2

ABSTRACT

Objectives ANTXR2 variants have been associated with ankylosing spondylitis (AS) in two previous genome-wide association studies (GWAS) (p<9×10⁻⁸). However, a genome-wide significant association (p<5×10⁻⁸) was not observed. We conducted a more comprehensive analysis of ANTXR2 in an independent UK sample to confirm and refine this association.

Methods A replication study was carried out with 2978 cases and 8365 controls. Then, these were combined with non-overlapping samples from the two previous GWAS in a meta-analysis. Human leukocyte antigen (HLA)-B27 stratification was also performed to test for ANTXR2-HLA-B27 interaction.

Results Out of nine single nucleotide polymorphisms (SNP) in the study, five SNPs were nominally associated (p<0.05) with AS in the replication dataset. In the meta-analysis, eight SNPs showed evidence of association, the strongest being with rs12504282 (OR=0.88, p=6.7×10⁻⁸). Seven of these SNPs showed evidence for association in the HLA-B27-positive subgroup, but none was associated with HLA-B27-negative AS. However, no statistically significant interaction was detected between HLA-B27 and ANTXR2 variants.

Conclusions ANTXR2 variants are clearly associated with AS. The top SNPs from two previous GWAS (rs4333130 and rs4389526) and this study (rs12504282) are in strong linkage disequilibrium (r²=0.76). All are located near a putative regulatory region. Further studies are required to clarify the role played by these ANTXR2 variants in AS.

INTRODUCTION

The genetic association between ankylosing spondylitis (AS) and human leukocyte antigen (HLA)-B27 was first established 40 years ago.1 Subsequently, it has become clear that AS is a polygenic disease with over 40 variants at 28 loci involved.1 These include anthrax toxin receptor 2 (ANTXR2), also known as capillary morphogenesis gene 2 (CMG2) to take account of its functions in basement membrane matrix assembly, angiogenesis and embryonic development.2 Rare ANTXR2 mutations cause the recessive Mendelian conditions juvenile hyaline fibromatosis and infantile systemic hyalinosis,3,4 while common upstream variants show suggestive association with myopia.5 ANTXR2 has no obvious role in AS, but it contains two single nucleotide polymorphisms (SNP), rs4333130 and rs4389526, which are in strong linkage disequilibrium (LD) in Europeans.6,7 that have shown suggestive association with AS in two previous genome-wide association studies (GWAS). In 2010, the Triple A Australo-Anglo-American Spondyloarthritis Consortium (TASC) reported association with the intronic SNP, rs4333130 (p=9.3×10⁻⁸); and in 2011, the Wellcome Trust Case Control Consortium 2 (WTCCC2) showed an association with rs4389526 (meta-analysis p=9.4×10⁻⁸).6 In neither study did the evidence for association reach the threshold for genome-wide significance (p<5×10⁻⁸). Subsequently, two smaller Chinese studies failed to show association with ANTXR2.8,9 Unfortunately, neither of the two strongly associated ANTXR2 SNPs was included in the recent Immunochip study designed to replicate and refine suspected genetic associations with AS.7 Consequently, we sought to clarify the potential association through a more comprehensive analysis of ANTXR2 in another independent UK sample prior to starting more detailed analysis of this region.

MATERIALS AND METHODS

The study was approved by the National Research Ethics Service, Cambridgeshire 4 Research Ethics Committee, UK (MREC project number 98/5/23). We studied 2978 AS cases fulfilling the 1984 modified New York Criteria10 who were not in the previous TASC 20107 or WTCCC2 analyses.6 Cases were genotyped on nine SNPs within ANTXR2 that had previously shown evidence of association with AS (p<0.05) in the TASC 2010 study.9 SNPs were genotyped using KASP technology (competitive allele-specific polymerase chain reaction amplification) by LGC Genomics (Hoddesdon, UK). Genotyping assays were validated on KASP panels, and cluster plots were checked for clearly separated genotype clusters.

Cases were compared to 8365 children from the Avon Longitudinal Study of Parents And Children who had previously been genotyped on the Illumina 550K platform (rs4234848, rs11098965, rs4444771, rs10000471, rs4333130, rs6839672) and then imputed to HapMap 2 as previously described (rs12504282, rs6534639 and rs4640621; imputation quality R²≥0.98).11,12 Please note that the study website contains details of all the data that is available through a fully searchable data dictionary (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/).

Control allele frequencies were checked for Hardy–Weinberg equilibrium (all p>0.05). Allele
frequencies between cases and controls were compared by logistic regression analysis. Where SNPs were imputed, ‘best guess’ genotypes were used in the calculation (all $R^2 \geq 0.98$). All statistical analyses were carried out using the software programme PLINK (http://pngu.mgh.harvard.edu/purcell/plink). Expected power of the study was based on the ORs from the TASC 2010 GWAS discovery sample. We estimate that our study had 98% statistical power to detect an association corresponding to an allelic OR of 0.82, with a minor allele frequency of 0.31, an $\alpha$ of 0.05 (two-sided), and a disease prevalence of 4/1000, assuming an additive model for disease risk (http://www.dartmouth.edu/~eugened/power-sample-size.php).

### Stratified analyses
Cases and controls were stratified by HLA-B27 status, and tested for association using logistic regression analysis. HLA-B27 status of cases and controls was determined using the rs4349859 marker, which tags HLA-B27 in Europeans with very high sensitivity and specificity.

### Table 1

<table>
<thead>
<tr>
<th>Replication study</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) All cases vs controls (2978 vs 8365)</td>
</tr>
<tr>
<td>SNP</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>rs6534639</td>
</tr>
<tr>
<td>rs4334848</td>
</tr>
<tr>
<td>rs11098965</td>
</tr>
<tr>
<td>rs4444771</td>
</tr>
<tr>
<td>rs12504282</td>
</tr>
<tr>
<td>rs10000471</td>
</tr>
<tr>
<td>rs4640621</td>
</tr>
<tr>
<td>rs4333130</td>
</tr>
<tr>
<td>rs4839672</td>
</tr>
</tbody>
</table>

Nominally significant (p<0.05) associations are shown (*).
Number of cases and controls in each analysis is also shown in brackets. (B27: HLA-B27). HLA, human leukocyte antigen; MAF, minor allele frequency; SNP, single nucleotide polymorphism.

### Meta-analysis

ORs from the TASC 2010 (1236 cases, 3979 controls), WTCCC2 2011 (1787 cases, 5162 controls) and this study were combined in an inverse variance meta-analysis assuming a fixed effects model. Combinaibility of the studies was determined using Cochran’s Q test (p>0.05), and the inconsistency measure ($I^2$) was used to estimate the amount of heterogeneity across the studies in the meta-analysis. For TASC, the analysis included correction for four principal components due to the different ethnicities in that study. Genome-wide significance level (p<5×10^{-8}) was used for declaring statistically significant genetic association. Meta-analysis conditioning on rs12504282 was also performed to detect secondary signals at the locus. Forest plots were produced using the metafor package in R.

### ANTXR2-HLA-B27 interaction analysis

Logistic regression analyses were performed in each cohort using the model (logOdds=B0+HLA-B27(dominant)+ANTXR2(additive)+ANTXR2xHLA-B27). The logORs and SEs for the interaction terms were combined in an inverse variance meta-analysis.

### Table 2

<table>
<thead>
<tr>
<th>Meta-analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) All cases vs controls (6001 vs 17 506)</td>
</tr>
<tr>
<td>SNP</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>rs6534639</td>
</tr>
<tr>
<td>rs4234848</td>
</tr>
<tr>
<td>rs11098965</td>
</tr>
<tr>
<td>rs4444771</td>
</tr>
<tr>
<td>rs12504282</td>
</tr>
<tr>
<td>rs10000471</td>
</tr>
<tr>
<td>rs4640621</td>
</tr>
<tr>
<td>rs4333130</td>
</tr>
<tr>
<td>rs6839672</td>
</tr>
</tbody>
</table>

Three SNPs reached genome-wide significance (p<5×10^{-8}) in the overall meta-analysis. Number of cases and controls in each analysis is also shown in brackets. (B27: HLA-B27). AS, ankylosing spondylitis; HLA, human leukocyte antigen; SNP, single nucleotide polymorphism.
meta-analysis to test for the presence of an ANTXR2-HLA-B27 interaction.

Bioinformatics
LocusZoom was used to display a regional association plot showing the pairwise LD between the top SNP and the other SNPs in the meta-analysis (1000 Genomes March 2012 European panel).14 The University of California, Santa Cruz (UCSC) genome browser (http://genome-euro.ucsc.edu/index.html) was used to interrogate the Encyclopaedia of DNA elements (ENCODE) data for regions of interest containing the associated SNPs.15 16 Public eQTL databases were used to determine possible effects of SNPs in the ANTXR2 region on gene expression (eQTL databases: http://www.sph.umich.edu/csg/liang/imputation/ and http://www.hsp.harvard.edu/timing-liang/software/eqtl/; GTEx: http://www.gtexportal.org; Genevar18: http://www.sanger.ac.uk/resources/software/genevar/).

RESULTS
Replication study
Five SNPs showed nominal evidence of association (p<0.05) with AS in the expected direction (table 1A). Four of these SNPs were also nominally associated with AS in the HLA-B27-positive subgroup analysis (table 1B), whereas only one of these showed nominal evidence of association in HLA-B27-negative AS (table 1C). There appeared to be a difference in the point estimates of the ORs between HLA-B27-positive and negative individuals (with HLA-B27-positive individuals showing stronger association for most SNPs) although the CIs around the estimates were wide.

Meta-analysis
In the meta-analysis, three SNPs showed strong evidence of association with AS (p<5×10^{-8}), while five SNPs exhibited suggestive associations (4.5×10^{-6}≤p≤0.0007; table 2A). Forest plots revealed that the direction and strength of the association was consistent across most SNPs (figure 1). Conditional analysis indicated no associations independent of rs12504282 (all p>0.1). Seven of the SNPs showed nominal evidence for association in the HLA-B27-positive AS subgroup (table 2B), but none was associated with HLA-B27-negative AS (p>0.05; table 2C). However, no statistically significant interaction was detected between HLA-B27 and ANTXR2 variants (p>0.05).

DISCUSSION
This study convincingly replicated the association between ANTXR2 variants and AS in an independent UK sample with three variants reaching genome-wide significance level in the meta-analysis. In the replication study, the strongest association was with rs12504282 (p=0.0001, OR=0.89, 95% CI 0.84 to 0.94), but there is a high degree of LD across this locus and the five SNPs showing significant association (table 1A) are distributed across the locus (figure 2). In the meta-analysis, eight SNPs showed evidence of association, three SNPs reaching genome-wide significance; rs12504282 again showed the strongest association (p=6.7×10^{-9}, OR=0.88). Forest plots show that direction and magnitude of association across the studies in the meta-analysis are in agreement with overlapping CIs (figure 1).

The strength of evidence for association appeared to be stronger in HLA-B27-positive patients compared with the HLA-B27 negatives. None of the seven SNPs, which showed nominal evidence of association with HLA-B27-positive AS showed...
nominal association with HLA-B27-negative disease, including the top SNP rs12504282 (table 1B, C). However, failure to detect nominal association with ANTXR2 in HLA-B27-negative AS probably reflects reduced statistical power rather than a genuine epistatic interaction. Consistent with this interpretation, the ANTXR2-HLA-B27 interaction analysis statistically testing for an epistatic effect revealed no significant findings. By contrast, the well-established epistatic genetic effect between HLA-B27 and ERAP1 in AS can be explained by the functional synergy between ERAP1 and HLA class I molecules in antigen processing and presentation.19 It is much less easy to explain a putative genetic interaction with ANTXR2 where there is no known functional interplay with HLA-B27 at the molecular level.

Biological explanations for the consistent associations of ANTXR2 variants with AS have not yet been forthcoming. The top SNPs from three different studies, including this one, are in close proximity and strongly correlated (rs4333130, rs4389526 and rs12504282; r²>0.76). Furthermore, these SNPs are near a putative transcription factor-binding region, and may therefore affect the expression level of ANTXR2 or another gene. Of interest, rs4333130 and rs12504282 have trans-eQTL effects on the expression of NSMAF on chromosome 8 (allele A, β=−0.27, p=0.00025) and GPR89A on chromosome 1 (allele T, β=−0.25, p=0.0088), respectively, in lymphoblastoid cell lines. NSMAF encodes a WD-repeat protein potentially playing a role in AS pathology.

Figure 2  Regional association plot showing ANTXR2 (CMG2) and all the single nucleotide polymorphisms (SNPs) in the study on the x-axis and their corresponding −log₁₀ p values on the y-axis. Historic sex-averaged recombination rates are also shown in blue (top panel). Linkage disequilibrium (r², 1000 Genomes March 2012 European reference panel) between the top SNP rs12504282 (purple) and the other SNPs are also shown with colour-coding. ENCODE data showing that rs12504282 and the previous top genome-wide association studies (GWAS) SNPs are near a region of possible enhancer binding are presented in the bottom panel (A) Gene structure; (B) H3K27Ac mark, ie, often found near active regulatory elements; (C) DNaseI hypersensitivity clusters; (D) Transcription factor ChIP-seq; (E) Chromatin state segmentation; (F) Transcription factor binding sites by epitope-tag; G: Top SNPs from the previous GWAS (rs4333130 (TASC, 2010) and rs4389526 (WTCCCq, 2011)). These three top SNPs from three different studies are strongly correlated (rs4389526-rs4333130 r²=1.0, D'=1.0; rs4389526-rs12504282 r²=0.8, D'=1.0; rs4333130-rs12504282 r²=0.76, D'=1.0).
in regulating tumour necrosis factor-induced cellular responses, such as inflammation. GPR89A encodes G protein-coupled receptor 89A involved in Golgi apparatus acidification possibly modulating its functions.\textsuperscript{20} Substantially more work will be required before the nature of this tantalising genetic association with AS can be explained at a functional level.

Acknowledgements We are extremely grateful to all the families who took part in the Avon Longitudinal Study of Parents And Children (ALSPAC), the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. This publication is the work of the authors and DME will serve as guarantor for the contents of this paper. The authors are also grateful for the support of the National Ankylosing Spondylitis Society UK and the NIHR Thames Valley CLRN for their roles in facilitating the recruitment of cases.

Contributors TK performed the statistical analyses and wrote the manuscript. SMK and JJP contributed to study design and HLA genotyping. MAB contributed to critical revision of the manuscript. BPW and DME contributed to subject recruitment/data collection and supervised the study. All authors participated in manuscript drafting and final manuscript review.

Funding The UK Medical Research Council, the Wellcome Trust (Grant ref: 092731) and the University of Bristol provide core support for ALSPAC. TK was funded by the National Ankylosing Spondylitis Society. SK and JJP were funded by the NIHR Oxford Thames Valley CLRN for their roles in facilitating the recruitment of cases.

Ethics approval National Research Ethics Service, Cambridgeshire 4 Research Ethics Committee, UK (MREC project number 98/5/23); ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. This publication is the work of the authors and DME will serve as guarantor for the contents of this paper. The authors are also grateful for the support of the National Ankylosing Spondylitis Society UK and the NIHR Thames Valley CLRN for their roles in facilitating the recruitment of cases.

Competition interests None.

REFERENCES

20 Maeda Y, Ide T, Koike M, et al. GPHR is a novel anion channel critical for regulation of tumour necrosis factor-induced cellular responses, such as inflammation. GPR89A encodes G protein-coupled receptor 89A involved in Golgi apparatus acidification possibly modulating its functions.\textsuperscript{20} Substantially more work will be required before the nature of this tantalising genetic association with AS can be explained at a functional level.