Combined analysis of three whole genome linkage scans for Ankylosing Spondylitis

K. W. Carter, A. Pluzhnikov[1], A. E. Timms[2], C. Miceli-Richard[3], C. Bourgain[4], B. P. Wordsworth[2], H. Jean-Pierre[5], N. J. Cox[1], L. J. Palmer, M. Breban[6], J. D. Reveille[7] and M. A. Brown[2,8]

Objective. Ankylosing spondylitis (AS) is a debilitating chronic inflammatory condition with a high degree of familiality ($\lambda_s = 82$) and heritability (>90%) that primarily affects spinal and sacroiliac joints. Whole genome scans for linkage to AS phenotypes have been conducted, although results have been inconsistent between studies and all have had modest sample sizes. One potential solution to these issues is to combine data from multiple studies in a retrospective meta-analysis.

Methods. The International Genetics of Ankylosing Spondylitis Consortium combined data from three whole genome linkage scans for AS ($n = 3744$ subjects) to determine chromosomal markers that show evidence of linkage with disease. Linkage markers typed in different centres were integrated into a consensus map to facilitate effective data pooling. We performed a weighted meta-analysis to combine the linkage results, and compared them with the three individual scans and a combined pooled scan.

Results. In addition to the expected region surrounding the HLA-B27 gene on chromosome 6, we determined that several marker regions showed significant evidence of linkage with disease status. Regions on chromosome 10q and 16q achieved 'suggestive' evidence of linkage, and regions on chromosomes 1q, 3q, 5q, 6q, 9q, 17q and 19q showed at least nominal linkage in two or more scans and in the weighted meta-analysis. Regions previously associated with AS on chromosome 2q (the IL-1 gene cluster) and 22q (CYP2D6) exhibited nominal linkage in the meta-analysis, providing further statistical support for their involvement in susceptibility to AS.

Conclusion. These findings provide a useful guide for future studies aiming to identify the genes involved in this highly heritable condition.

KEY WORDS: Ankylosing Spondylitis, genome scans, meta-analysis.

Ankylosing spondylitis (AS) is a common, chronic inflammatory arthropathy that primarily affects the spinal joints and sacroiliac joints, causing symptoms of pain and stiffness, and progressive fusion of involved joints. Peripheral arthritis and/or uveitis (inflammation of the eye) complicates approximately 40–50% of cases; less common non-spinal manifestations include enthesitis, aortitis and pneumonitis. Associations between AS and psoriasis and inflammatory bowel disease are well described. AS has an estimated prevalence of 1/1000 to 4/1000 in white populations [1, 2].

Genetic factors have long been implicated in AS, with recognition of association with the HLA-B27 gene over 30 yrs ago [3, 4]. Whilst over 90% of AS patients carry the gene, only approximately 1–6% of the general population HLA-B27 carriers develop AS [1, 2]. Twin and family studies suggest that this can be explained by the additional involvement of several other genes [5, 6]. There is evidence that other genes lying within the major histocompatibility complex (MHC) [7], interleukin-1 gene cluster [8–11], and the gene CYP2D6 influence susceptibility to AS [12, 13], but clearly a substantial proportion of the non-B27 heritability of AS is as yet unexplained.

The International Genetics of Ankylosing Spondylitis (GAS) Consortium was established in 2003 as an international effort to define the candidate regions and genes for AS with greater precision and reliability than can be achieved by smaller individual studies. We have gathered the results of three separate whole genome scans for AS and examined them for evidence of linkage with disease. These are the 'Oxford' [14], Groupe Français d'Etude Génétique des Spondyloarthropathies (GFEGS) [15], and North American Spondyloarthritis Consortium (NASC) datasets [16]. The 'Oxford' cohort consists of two subsets, the findings of a genomewide scan on the first of which was first reported in 1998 [17], and the second in 2001 [14]. The two sets were pooled for the current study. These datasets represent all currently published genome-wide linkage data in AS. Whilst a previous meta-analysis has been reported in AS, it did not use the original linkage data as we have here, but analysed only data found in figures and tables from previous linkage screens, did not take into account the differences in phenotype definition in the different datasets, nor was able to identify and correct marker position and order errors in the original datasets [18].

In addition to examining studies individually, we examined a combined dataset of the three studies pooled together giving each dataset equal weighting (referred to as the 'pooled' analysis), and compared this with a weighted meta-analysis of the three studies to explore the overall combined effect across the studies (referred to as the 'weighted' meta-analysis). This weighted meta-analysis takes into account the differences in size, marker density and linkage information content extracted by each screen. The use of meta-analysis techniques should improve on the power of the individual datasets, allowing us to examine in finer detail the genetic effects across three independent populations.

Patients and methods

Study populations

There are significant differences in the disease definitions used, types of families studied, and marker sets genotyped between the

Laboratory for Genetic Epidemiology, Western Australian Institute for Medical Research, UWA Centre for Medical Research, University of Western Australia, Ground Floor, B Block, Hospital Avenue, Nedlands, Western Australia 6009, Australia, [1]Department of Human Genetics, The University of Chicago, Chicago, USA, [2]Botnar Research Centre, Institute of Musculoskeletal Sciences, University of Oxford, Nuffield Orthopaedic Centre, Windmill Road, Headington, OX3 7LD, UK, [3]Laboratoire de Génétique des Maladies Inflammatoires de l'Intestin et Fondation Jean Dausset/CEPH, Paris, [4]INSERM U535, Villejuif, [5]INSERM AVENIR U458 and AP-HP, Hôpital Robert Debré, 48 Bd Sérurier, 75019 Paris, [6]Département d'Immunologie, Institut Cochin, INSERM U567, CNRS UMR 8104, IFR116, and Hôpital Ambroise Paré, Assistance Publique-Hôpitaux de Paris, Université Paris, Ile de France Ouest Paris, France, [7]The University of Texas Health Science Center at Houston, Houston, Texas, USA and [8]Centre for Immunology and Cancer Research, Princess Alexandra Hospital, Woolloongabba, Queensland, Australia.

TABLE 1. Family compositions of genome screens employed in meta-analysis[a]

Families	GFEGS				
	AS	SpA	NASC	Oxford	Total
Overall	27	84	232	197	488
Number of affected sibling pairs					
1	18	51	154	154	341
2	0	1	2	4	7
3	6	15	21	17	47
4	1	3	0	1	3
More	1	5	4	5	13
Sibpairs	46	139	245	251	589
Affected individuals	74	244	485	461	1116
Male (%): female (%)	51:49	50:50	58:42	66:34	60:40

[a]SpA includes all cases meeting the ESSG Criteria for Spondyloarthritis (i.e. including those also meeting the modified New York Criteria for AS).

three studies. Both the Oxford and NASC cohorts consisted entirely of families with AS defined by the modified New York criteria [19]. In contrast, the GFEGS dataset included a substantial proportion of cases with 'spondyloarthritis' (SpA), as defined by the ESSG diagnostic criteria [20]. These cases may have AS, or may have features of seronegative arthritis but may not meet the strict 'AS' radiographic criteria of the modified New York criteria, and are clinically and genetically related to AS. The GFEGS dataset was studied either as a whole (GFEGS-SpA, including all those meeting the ESSG Criteria and modified New York Criteria for AS), or considering only those cases with AS, as defined by the modified New York criteria (GFEGS-AS). As it is likely that in time many cases with SpA will progress and meet the modified New York Criteria for AS, the SpA group not currently meeting the AS criteria were not analysed separately.

The GFEGS dataset consisted of 431 individuals from 86 pedigrees with a total of 139 affected sibling pairs, genotyped with a total of 346 markers from the Applied Biosystems Prism Linkage Mapping Set Version 2.0 (LMSV2) (Applied Biosystems). The NASC dataset was comprised of 2203 individuals from 232 families with a total of 245 sibling pairs, genotyped using the ABI LMSV2 marker set. The Oxford dataset consisted of 1102 individuals from 198 pedigrees with a total of 251 affected sibling pairs. These were genotyped using both the ABI LMSV2 and the Medical Research Council (UK) microsatellite marker sets [21]. The family makeup of each screen is given in Table 1.

All study populations were collected under appropriate institutional ethics approvals, as detailed in the original papers (referenced earlier).

Consensus map

A consensus map that incorporated all typed markers in each of the three datasets was constructed to pool raw genotype data across samples. This map provided a framework of markers upon which typed markers were integrated using PYGMALION [22]; markers were ordered using the deCODE recombination marker map [23].

Linkage analysis

Possible errors in the pedigree and genotype data were assessed using the RELCHECK [24], PEDCHECK [25] and MERLIN [26] programs. These analyses suggested that the maximal potential error rate in any one dataset was less than 1% of all genotypes. Our error-checking procedure resulted in the exclusion of very few families at a given marker from any of the datasets. Marker allele frequencies were estimated using the MERLIN program by maximum likelihood estimation [27].

Multipoint non-parametric linkage analysis was performed using the MERLIN v0.10.2 pedigree analysis software package [27]. *P*-values were calculated by simulation and then converted to a $-\log_{10}$ scale and plotted using JLGraph [28]. Because the dataset was simulated 10 000 times, the minimum reported *P*-value is $<10^{-4}$, and maximum $-\log_{10}$ (*P*-value) is 4.

Pooled analysis

The three individual AS-alone datasets were combined to create a single pooled AS-alone dataset comprising all markers and families. This pooled dataset was analysed for linkage using Merlin (with the common integrated map) and empirical *P*-values calculated as previously described. A pooled SpA dataset was created by combining the markers and families from the AS alone and GFEGS-SpA datasets (i.e. all cases from the three datasets that meet the ESSG Criteria, most of whom also meet the modified New York Criteria), and analysed similarly. Because of the strong linkage between AS and markers contained in the MHC on chromosome 6 and the limits of our simulations, six markers contained in the Oxford dataset and four markers contained in the NASC dataset produced a *P*-value that was effectively zero. As this result could not be plotted on a $-\log_{10}$ scale, these markers were plotted with the arbitrary value of 10 (relative to surrounding markers). Seven markers in this same location in the both the AS-alone and SpA pooled datasets were plotted in the same manner.

Meta-analysis

We analysed each of three datasets for linkage to the genotyped markers. In order to perform a meta-analysis of the whole genome linkage scan results (for AS only affection), we then pooled linkage statistics across the datasets using a weighted combination of *z*-score approach. Given the varying information content and population sizes across the three genome scans, we chose an approach based on weighting for both information content and study size as previously described by [29].

Using the weighted meta-analysis approach, the *z*-scores for each of the 338 common markers across the scans were combined. These weighted *z*-scores follow a standard normal distribution (mean = 0, S.D. = 1), from which *P*-values were determined. As with the empirical linkage results, we converted the resulting *P*-values to a $-\log_{10}$ scale for plotting using JLGraph. Three marker locations contained in chromosome 6 produced *P*-values of zero, and as described previously were assigned the value of 10 for plotting purposes.

IBD sharing by affected sibling pairs and $\lambda_{(\text{LOCUS})}$ values were determined using MapMaker/Sibs [30] weighted according to the number of affected sibling pairs available in each cohort. The contribution of each locus to the overall sibling recurrence risk was calculated assuming a sibling recurrence risk ratio of 82 [6], and multiplicative interaction between loci [31], the most likely model in AS [6].

Results

The results of the three individual whole genome linkage scans for AS are presented in Fig. 1. The results of the pooled AS-alone analysis and weighted meta-analyses are compared in Fig. 2 and Fig. 3 presents the results of the pooled AS-alone and the pooled SpA analyses. Significant results are indicated in Figs 1–3 by small *P*-values (large values on a $-\log_{10}$ scale) forming peaks over relevant markers. Note that the positioning of results within each of the figures reflect that most p-terminal markers used in each study are not located at 0 cM on the chromosomal genetic map.

Individual scans

Considering the individual scans (Fig. 1), each shows highly significant linkage to the region of chromosome 6 encoding the MHC ($P < 10^{-5}$). 'Nominal' or greater evidence of linkage
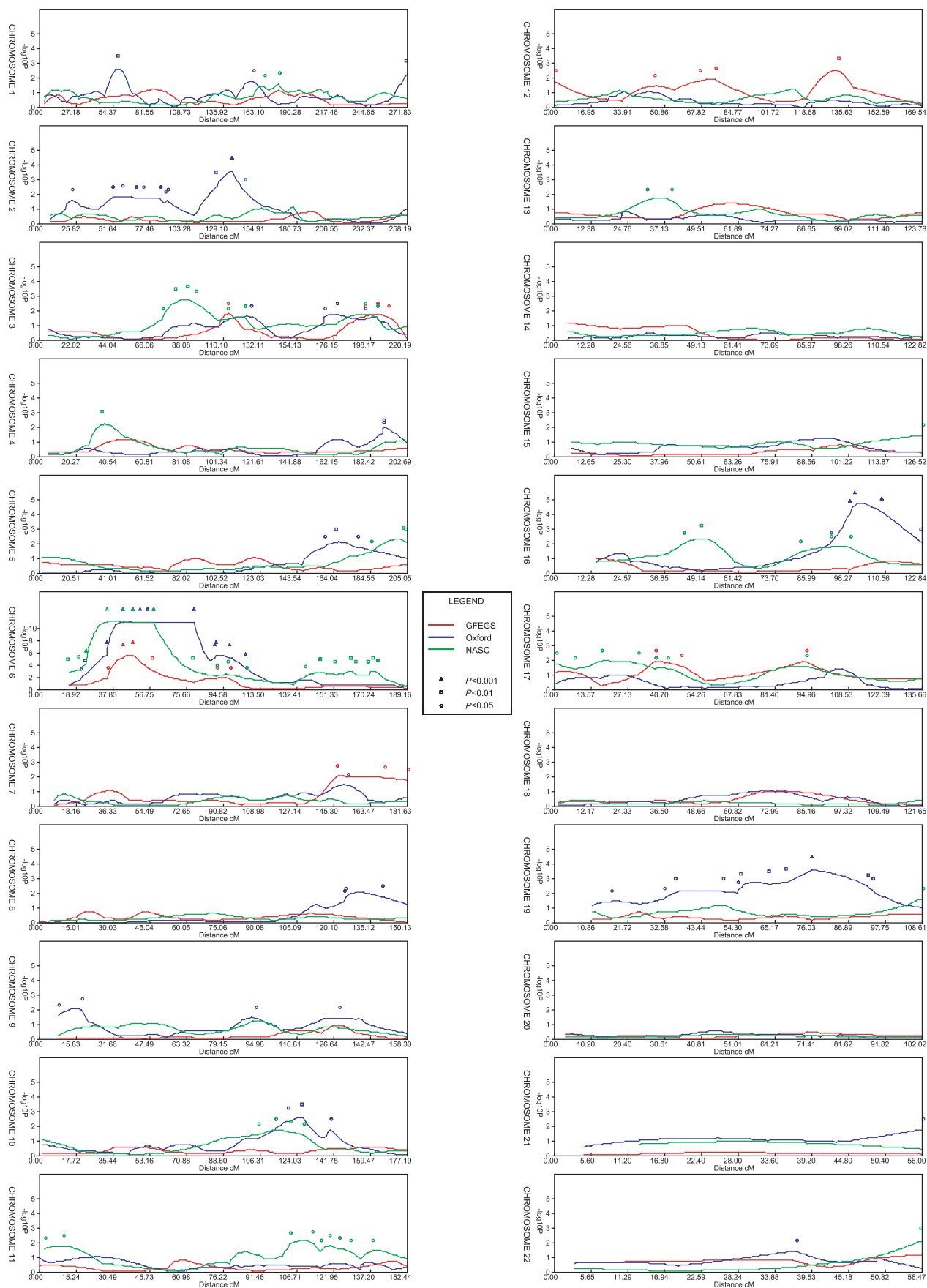
Fig. 1. Ankylosing spondylitis whole genome linkage scans. A circle indicates the marker has a *P*-value < 0.05, a square indicates *P*-values < 0.01, a triangle indicates *P*-values < 0.001. Linkage with AS affection status is shown for GFEGS, Oxford and NASC datasets in red, blue and green respectively. Please note, chromosome 6 has a different $-\log_{10}$ *P*-value scale.
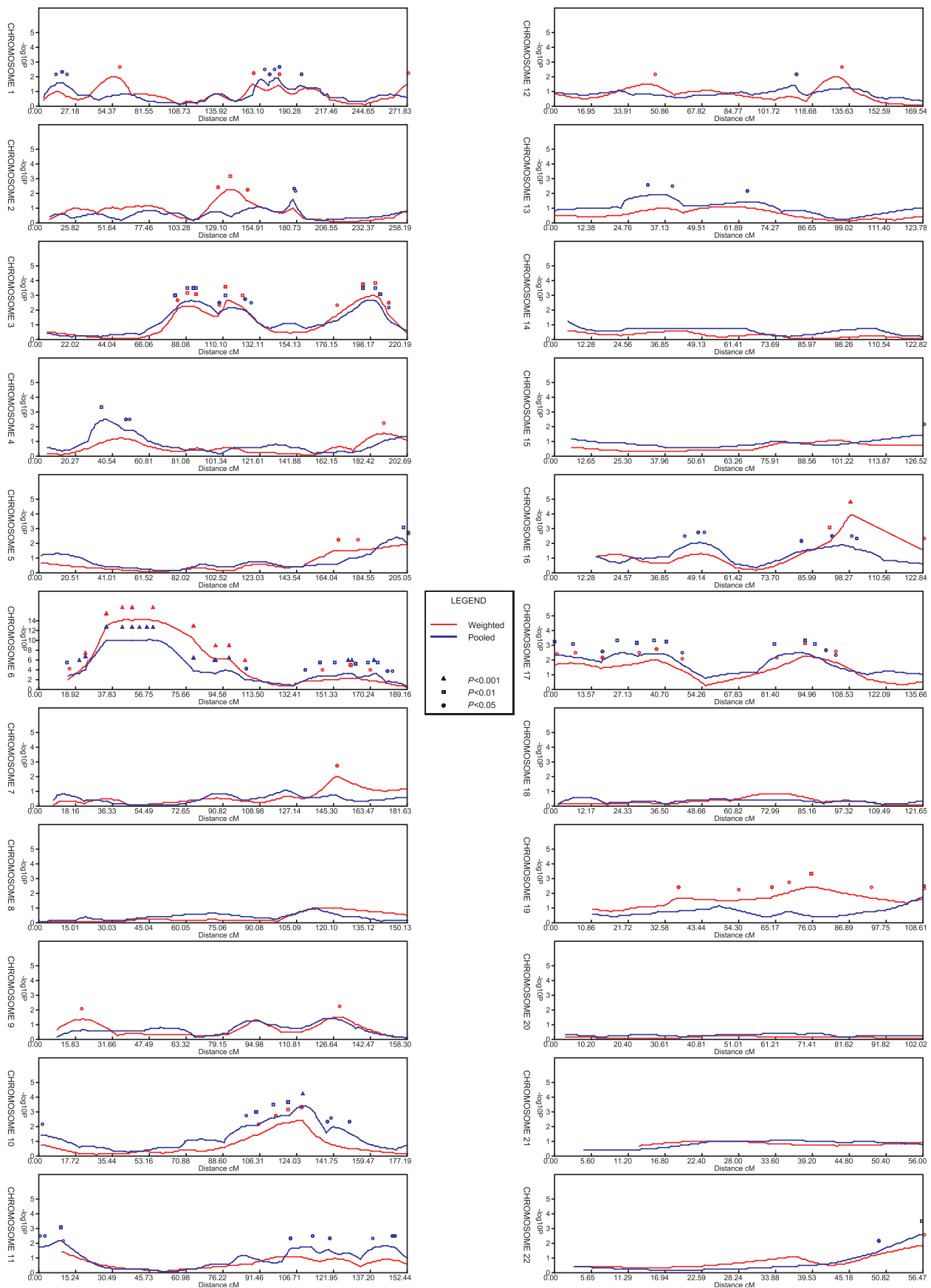
FIG. 2. Ankylosing spondylitis weighted meta-analysis *vs* pooled dataset. A circle indicates the marker has a *P*-value < 0.05, a square indicates *P*-values < 0.01, a triangle indicates *P*-values < 0.001. Weighted meta-analysis is shown in red, while pooled linkage scan is shown in blue. Please note, chromosome 6 has a different −log₁₀ *P*-value scale.
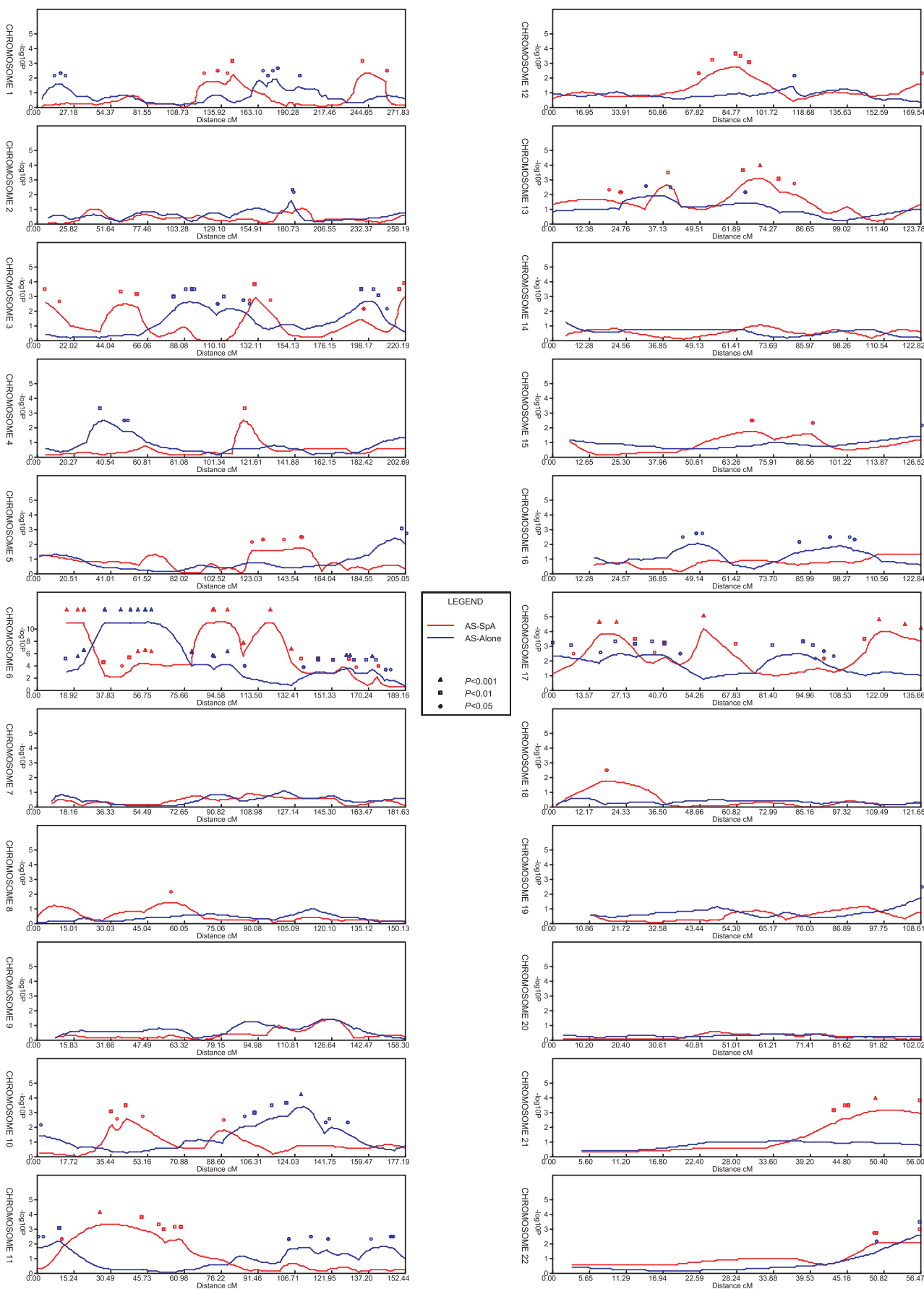
Fig. 3. Pooled SpA dataset *vs* pooled AS-alone dataset. A circle indicates the marker has a *P*-value < 0.05, a square indicates *P*-values < 0.01, a triangle indicates *P*-values < 0.001. Pooled SpA is shown in red, while pooled AS-alone scan is shown in blue. Please note, chromosome 6 has a different −log$_{10}$ *P*-value scale.

TABLE 2. Markers achieving $P \leq 0.05$ in three or more datasets/analyses[a]

Chromo-some	Marker	Location	Pooled-AS	Weighted-AS	GFEGS-AS	GFEGS-SpA	NASC-AS	Oxford-AS
3	D3S1261	93.51	0.003	0.03			0.003	
3	D3S1566	94.76	0.003	0.01			0.002	
3	D3S3681	107.43	0.02	0.01			0.004	
3	D3S1271	112.28	0.009	0.05	0.02	0.04	0.05	
3	D3S1278	122.74	0.01	0.04			0.04	0.05
3	D3S1303	126.07	0.02	0.007			0.03	0.03
3	D3S1262	194.33	0.003			0.03	0.02	0.04
3	D3S1580	202	0.003	0.005	0.02		0.03	0.03
3	D3S1601	208.14	0.04		0.03	0.04		
4	D4S419	34.97	0.004	0.03			0.007	
4	D4S391	47.33	0.02	0.03			0.05	
5	D5S408	205.05	0.01	0.03		0.04	0.009	
6	D6S309	20.5	0.0008				0.001	0.04
6	D6S470	23.75	0.0001	0.0005		0.03	0.0002	0.005
6	D6S289	34.61	0	0.00000002	0.03	0.002	0	0.00001
6	D6S422	42.83	0	0	0.00002	0.00001	0	0
6	D6S276	47.93	0	0	0.00001	0	0	0
6	D6S1610	58.47	0	0	0.002	0.00002	0	0
6	D6S257	79.52	0.0003	0.000002		0.02	0.002	0
6	D6S460	90.58	0.0008	0.000007	0.03		0.01	0.00001
6	D6S462	97.76	0.0002	0.00004	0.03		0.007	0.00002
6	D6S434	106.2	0.01	0.004			0.03	0.0007
6	D6S292	136.39	0.02				0.02	0.05
6	D6S441	160.64	0.0007	0.03		0.01	0.002	
7	D7S684	146.68		0.02	0.01			0.05
9	D9S1776	121.62	0.05	0.04				0.05
9	D9S1682	128.77	0.05	0.04		0.0003		0.04
10	D10S1686	105.03	0.009	0.04			0.04	
10	D10S185	113.34	0.003	0.02			0.02	0.02
10	D10S192	120.75	0.002	0.01			0.03	0.005
10	D10S597	126.7	0.0005	0.008			0.04	0.003
13	D13S171	31.07	0.01			0.03	0.03	
13	D13S218	39.34	0.02			0.04	0.03	
16	D16S515	92.06	0.02	0.01			0.02	0.01
16	D16S516	98.52	0.02	0.0002			0.02	0.0001
17	D17S831	7.22	0.007	0.04			0.04	
17	D17S1852	30.67	0.006	0.04			0.02	
17	D17S799	37	0.004		0.01	0.04	0.04	
17	D17S787	81.32	0.008		0.05	0.03	0.05	
17	D17S944	92.76	0.004		0.01	0.005	0.03	
17	D17S949	102.96	0.03	0.02		0.008		0.05

[a]Full results for all markers achieving $P < 0.05$ are available online (www.cicr.uq.edu.au/igas/meta-analysis).

($P < 0.05$) was observed for two or more scans on chromosomes 1, 3, 5, 7, 9, 10, 13, 16, 17, 19 and 22. Regions achieving at least moderate evidence of linkage ($P < 0.05$ in three or more individual datasets or the meta-analyses) are given in Table 2. These were observed in the GFEGS-SpA scan on chromosomes 9 (129 cM, $P = 0.0003$), 11 (53 cM, $P = 0.003$), 12 (31 cM, $P = 0.002$), and 17 (93 cM, $P = 0.005$). When considering cases of AS alone, of these linkages only the chromosome 12 linkage remained significant ($P < 0.01$). In the NASC dataset, moderate or greater linkage was observed on chromosome 3 (95 cM, $P = 0.002$), 4 (35 cM, $P = 0.007$), 5 (204 cM, $P = 0.007$), 6q (161 cM and 172 cM, $P = 0.002$), 16 (50 cM, $P = 0.005$), and 22 (56 cM, $P = 0.009$). In the Oxford families, moderate or greater linkage was observed on chromosomes 1 (59 cM, $P = 0.003$; 272 cM, $P = 0.006$), 2 (135 cM, $P = 0.0003$), 5 (166 cM, $P = 0.009$), 10 (127 cM, $P = 0.003$), 16 (101 cM, $P = 3 \times 10^{-5}$), and 19 (76 cM, $P = 0.0003$).

Pooled analysis

In addition to the very significant evidence of linkage with the 30–70 cM region of chromosome 6 containing the MHC, the pooled analysis results show markers with moderate evidence of linkage ($P < 0.01$) in chromosomes 3, 4, 5, 6, 10, 11, 13, 16, 17 and 22 (Table 2). Clusters of significant marker results were seen on chromosome 1 at approximately 170 cM ($P < 0.05$), chromosome 3 at 88–130 cM ($P < 0.01$) and 198 cM ($P < 0.05$), and with several markers across chromosome 17.

Consistent with each of the individual scans and the AS-alone weighted meta-analysis, the pooled analysis exhibited significant evidence of linkage ($P < 0.00001$) with the 30–90 cM region of

chromosome 6 in the AS-alone and AS-SpA datasets. Clusters of significant marker results are also seen on chromosomes 1 ($P < 0.01$), 3 ($P < 0.01$), 10, 11, 12, and 13 ($P < 0.01$), 17 ($P < 0.001$), 21 ($P < 0.01$) and 22 ($P < 0.01$).

Weighted meta-analysis

Considering the analyses of AS alone, highly significant linkage was observed with all three individual scans and with the combined analyses ($P < 10^{-5}$) to the 37–75 cM region of chromosome 6, consistent for linkage with the location of the HLA-B27 marker. In the weighted AS-alone meta-analysis, the next strongest evidence of linkage was found at chromosome 16q (maximum linkage at 99 cM, $P = 1.8 \times 10^{-4}$). Evidence for linkage between markers in this this region with AS was seen in both the NASC and Oxford datasets ($P = 0.02$ at 99 cM and $P = 3 \times 10^{-5}$ at 101 cM, respectively). Other regions achieving moderate evidence of linkage ($P < 0.01$) were identified on chromosomes 3, 10 and 19. On chromosome 3, peak linkage was seen at 202 cM ($P = 0.005$), with support in the GFEGS, NASC and Oxford datasets ($P = 0.02$, 0.03 and 0.03, respectively, each at 202 cM). On chromosome 10, peak linkage was seen at 127 cM ($P = 0.008$), with support in the NASC and Oxford datasets ($P = 0.02$ at 113 cM and $P = 0.003$ at 127 cM, respectively). A broad region of linkage was observed on chromosome 19, the peak of which lay at 76 cM ($P = 0.005$), with support in the NASC and Oxford datasets, although the peaks of linkage in these studies were quite separated ($P = 0.03$ at 109 cM and $P = 0.0003$ at 76 cM, respectively).

The $\lambda_{(LOCUS)}$ value for chromosome 6p was 4.5, representing 34% of the familiality of AS, assuming a multiplicative interaction between loci. For chromosome 16q, the $\lambda_{(LOCUS)}$ value was 1.6, representing 11% of the familiality of AS.

Previously associated regions

Association has previously been reported between AS and the IL-1 gene cluster on chromosome 2q13 (at 126 cM), and the gene CYP2D6 on chromosome 22q13.2 (at 50 cM). Considering the IL-1 gene cluster, this region showed evidence of linkage in the Oxford dataset (135 cM, $P = 0.0003$) and in the weighted meta-analysis (135 cM, $P = 0.01$), but not in other individual datasets. The CYP2D6 region showed nominal evidence of linkage in the North American and Oxford datasets (56 cM, $P = 0.009$ and 37 cM, $P = 0.04$, respectively) and in the pooled analysis (56 cM, $P = 0.003$).

Examining heterogeneity of findings between the three studies (for AS-alone), using the 338 common markers, we found that 114 (approx 33%) of the marker locations had substantial heterogeneity (>50%) using the I2 measure [32], and 105 show substantial heterogeneity ($P < 0.1$) using Cochran's Q test. Groups of markers showing substantial heterogeneity are discussed further in the Discussion.

Discussion

This study was designed to investigate the evidence for linkage to AS phenotypes in a retrospective analysis of combined genome-wide data from three separate studies. Meta-analysis can potentially improve on the power of individual genome screens to identify linkages, and help clarify the significance of linkage results where there is inconsistent replication across screens (such as [29, 33–35]). We have performed meta-analysis by both pooling results and a weighted combination of linkage statistics across studies. These two approaches each have strengths and weaknesses. Weighted meta-analysis more correctly accounts for the difference in size and information content between the different scans, but only uses markers common to the three studies ($n = 338$). The pooled analysis uses all markers available potentially increasing its power, but weights each study equally. Where there is discordance between the studies this can lead to the findings from smaller studies having a disproportionate effect on pooled results.

The study confirms on a genome-wide level the 'highly significant' linkage ($P < 3 \times 10^{-7}$) of the region of chromosome 6 encoding the MHC, and also identifies 'suggestive' linkage ($P < 7.4 \times 10^{-4}$) on chromosomes 6q, 10q and 16q. Several other regions, whilst not achieving 'suggestive' or greater evidence of linkage, found support in two or more screens such as to suggest that it is likely that they are true positive findings.

In addition to the highly significant linkage observed on chromosome 6p around the location of HLA-B27, there are a number of other interesting marker patterns on the chromosome 6 which illustrate the advantage of the meta-analysis approach compared with comparisons of individual screens. Between 140 cM and 175 cM in the NASC linkage results in Fig. 1, there are eight markers with significant ($P < 0.01$) linkage to AS. In contrast, the Oxford and GFEGS datasets return no significant markers for this region. The pooled analysis, as expected given the fact that the NASC dataset contributes 42% of the total studied, also returns significant results for the region. The weighted meta-analysis is therefore of particular interest, given the lack of replication across this region in the remaining two studies, identifying four significant markers for linkage in this region, the most significant result being at 160.64 cM (marker name AFM269ze1/D6S441, $P = 0.007$). Examining the three linkage scan results individually could have missed this linkage, as the results were not reproduced across studies, demonstrating

the value of the meta-analysis method. These findings strongly support the existence of further non-MHC genes involved in susceptibility to AS on chromosome 6q.

Linkage to chromosome 16q in AS was first reported by the Oxford group in 1998 [17], with further support for this being a true positive finding coming from their further genome screen [14] and the NASC study [16]. However, no linkage was observed in this region in the GFEGS study. In the weighted meta-analysis this region is the most significant non-MHC region, achieving $P = 0.0002$, the peak lying at 99 cM. Results for chromosome 10 surrounding 124 cM were also significant in the NASC and Oxford but not the GFEGS dataset. In particular, the marker at location 127 cM (AFM331xa9/D10S597) was significant in the meta-analysis ($P = 0.004$) and the pooled analysis ($P = 0.0005$). Thus, despite the inconsistent replication of this finding in different genome screens, the meta-analysis findings suggest that these are probably true positive findings, warranting further study.

Other regions potentially containing AS genes may only fulfil less stringent statistical threshold but are worth pursuing, particularly if there is support from more than one screen. Other than the regions described earlier, nominal or greater evidence of linkage ($P < 0.05$) was observed for two or more scans on chromosomes 1, 3, 5, 9, 13, 17, 19 and 22. Although there is clearly less certainty about these findings, there is at least a moderate probability that they contain disease-susceptibility genes.

Substantial differences were observed between the findings of the GFEGS study and those from the NASC and Oxford cohorts, the main exception being the consistent linkage to the MHC in all three screens. Potential explanations for this include the lower sample size of the GFEGS study, differences in structure of the families studied (the GFEGS cohort included a higher proportion of large pedigrees rather than affected-sibling pair families than the other screens), or the diagnostic scheme applied. The GFEGS–AS dataset is significantly smaller than the NASC or Oxford datasets, which likely explains the difference in the findings for that study. These differences between the scans at least partly explain the differences in finding between the pooled analysis and weighted meta-analyses.

Examining the differences between the pooled AS-alone and SpA analyses, a number of interesting patterns occur. As shown in Fig. 3, the evidence for linkage of chromosome 16q with disease is evident in the AS-alone dataset and not in the pooled SpA dataset. However, as noted previously, the GFEGS dataset is smaller than the NASC and Oxford datasets, likely explaining the difference between the pooled analysis and weighted meta-analysis findings at this region. On chromosome 21 at ~40–56 cM the individual scans and the weighted meta-analysis reveal no evidence for significant linkage, while the pooled SpA analysis reveals a cluster of markers suggestive of linkage, peaking at marker D21S268 (49 cM, $P = 0.0009$). Similarly, a closer examination of markers in the 35–50 cM region of chromosome 10 also shows evidence for linkage ($P < 0.01$) with markers in the pooled SpA analysis, that do not appear in the weighted meta-analysis. These results warrant further investigation to determine whether this is a result of the combined pooled analysis composition, or reflects heterogeneity in disease definition between the pooled datasets, or some other confounding factor.

Comparison of pooled analysis and weighted meta-analysis results are helpful in the interpretation of chromosomal regions suggestive of linkage with AS-alone that are not replicated in the SpA datasets. Linkage observed with chromosome 10 at 127 cM in the weighted meta-analysis of AS-alone and SpA, and in the pooled AS-alone analysis, returned no significant evidence for linkage within the pooled SpA analysis. The broad region of chromosome 19 identified for linkage in the AS-alone meta-analysis was not identified at all within the SpA dataset (either in the meta-analysis or the pooled studies). Notably neither the

GFEGS-SpA nor GFEGS-AS datasets provided evidence for linkage to this chromosome. This absence of linkage would least affect the AS-alone weighted meta-analysis. Thus, this disparity may be due to identification of marker locations specific to AS-alone rather that the more general condition of SpA, an observation warranting further study.

A closer examination of the population heterogeneity calculation across the studies reveals that the majority of markers showing substantial heterogeneity fall in regions where no significant evidence for linkage was discovered. The only major block of linkage results that overlay a large region of markers showing heterogeneity is on chromosome 12 at ∼66–112 cM in the pooled SpA analysis. However, the weighted meta-analysis and pooled AS-alone analysis reveal no significant evidence for linkage in this region, suggesting that perhaps the linkage with SpA in this region may reflect population heterogeneity. Other large blocks of markers showing substantial heterogeneity are located at the q-terminal of chromosome 13 and across the entire chromosome 14, where no significant evidence for linkage was revealed in any of the analyses.

In comparison with a previously published meta-analysis [36], this study is more accurate and powerful. The previous study employed data extrapolated from published graphs and tables and did not accurately differentiate between AS and SpA, the AS families from the GFEGS screen being excluded from the AS-alone analysis. In contrast, we have used the original genotype data allowing for better genotype error checking, correction of marker maps to take into account the most recent genome maps, precise multipoint analysis rather than division into 30 cM bins, and pooled and weighted analyses. AS and SpA analysis have been presented separately. Lee *et al.*'s study report only three loci achieving $P < 0.01$ (on chromosomes 6, 16 and 19), which were each identified in this screen. Two other regions achieved nominal levels of significance with the two statistical scores used in this study (*Psumrnk*, *Pord* < 0.05, >0.01), lying on chromosome 17p and 9q. These areas overlap with regions identified in the current study. However, the level of significance we report is generally greater than reported by Lee *et al.*, which also did not identify several regions we report, including on chromosomes 3, 5, 10 and 13 which were present in three or more datasets or analyses (Table 2).

In conclusion, this analysis provides an important step forward for narrowing down chromosomal locations that may hold the key to discovering the genetic linkages of AS. By performing a meta-analysis, we have identified several common markers that bear further investigation. In particular, we found 'suggestive' linkage on chromosomes 10q and 16q, and moderate evidence of linkage ($P < 0.01$) on chromosomes 3, 4, 5, 6, 10, 11, 16, 17 and 22. Further, regions on chromosome 2 (the IL-1 gene cluster) and 22 (the gene *CYP2D6*), previously reported to be associated with AS, showed at least nominal evidence of linkage ($P < 0.05$) with AS in the meta-analysis. These findings strongly support that this disease is at least oligogenic, and provide a valuable road-map for future studies aiming to identify the genes involved in this condition.

Rheumatology key messages

- Linkage of AS and SpA with the MHC is strong, but the condition is most likely polygenic, with loci on chromosomes 1q, 3q, 5q, 6q, 9q, 10q, 16q, 17q and 19q likely to harbour additional susceptibility genes.
- Regions previously associated with AS on chromosome 2q (the *IL*-1 gene cluster) and 22q (*CYP2D6*) exhibited nominal linkage in the meta-analysis, providing further statistical support for their involvement in susceptibility to AS.

References

1. van der Linden S, Valkenburg H, Cats A. The risk of developing ankylosing spondylitis in HLA-B27 positive individuals: a family and population study. Br J Rheumatol 1983;22(4 Suppl 2):18–9.
2. Braun J, Bollow M, Remlinger G, Eggens U, Rudwaleit M, Distler A, Sieper J. Prevalence of spondylarthropathies in HLA-B27 positive and negative blood donors. Arthritis Rheum 1998;41:58–67.
3. Brewerton DA, Hart FD, Nicholls A, Caffrey M, James DC, Sturrock RD. Ankylosing spondylitis and HL-A 27. Lancet 1973;1:904–7.
4. Schlosstein L, Terasaki PI, Bluestone R, Pearson CM. High association of an HL-A antigen, W27, with ankylosing spondylitis. N Engl J Med 1973;288:704–6.
5. Brown MA, Kennedy LG, MacGregor AJ, Darke C, Duncan E, Shatford JL, Taylor A, Calin A, Wordsworth P. Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment. Arthritis Rheum 1997;40:1823–8.
6. Brown MA, Laval SH, Brophy S, Calin A. Recurrence risk modelling of the genetic susceptibility to ankylosing spondylitis. Ann Rheum Dis 2000;59:883–6.
7. Brown MA, Kennedy LG, Darke C, Gibson K, Pile KD, Shatford JL, Taylor A, Calin A, Wordsworth BP. The effect of HLA-DR genes on susceptibility to and severity of ankylosing spondylitis. Arthritis Rheum 1998;41:460–5.
8. McGarry F, Neilly J, Anderson N, Sturrock R, Field M. A polymorphism within the interleukin 1 receptor antagonist (IL-1Ra) gene is associated with ankylosing spondylitis. Rheumatology 2001;40:1359–64.
9. van der Paardt M, Crusius JB, Garcia-Gonzalez MA *et al.* Interleukin-1beta and interleukin-1 receptor antagonist gene polymorphisms in ankylosing spondylitis. Rheumatology 2002;41:1419–23.
10. Maksymowych WP, Reeve JP, Reveille JD *et al.* High-throughput single-nucleotide polymorphism analysis of the IL1RN locus in patients with ankylosing spondylitis by matrix-assisted laser desorption ionization-time-of-flight mass spectrometry. Arthritis Rheum 2003;48:2011–8.
11. Timms AE, Crane AM, Sims AM *et al.* The interleukin 1 gene cluster contains a major susceptibility locus for ankylosing spondylitis. Am J Hum Genet 2004;75:587–95.
12. Brown MA, Edwards S, Hoyle E *et al.* Polymorphisms of the CYP2D6 gene increase susceptibility to ankylosing spondylitis. Hum Mol Genet 2000;9:1563–6.
13. Beyeler C, Armstrong M, Bird HA, Idle JR, Daly AK. Relationship between genotype for the cytochrome P450 CYP2D6 and susceptibility to ankylosing spondylitis and rheumatoid arthritis. Ann Rheum Dis 1996;55:66–8.
14. Laval SH, Timms A, Edwards S *et al.* Whole-genome screening in ankylosing spondylitis: evidence of non-MHC genetic-susceptibility loci. Am J Hum Genet 2001;68:918–26.
15. Miceli-Richard C, Zouali H, Said-Nahal R *et al.* Significant linkage to spondyloarthropathy on 9q31–34. Hum Mol Genet 2004;13:1641–8.
16. Zhang G, Luo J, Bruckel J *et al.* Genetic studies in familial ankylosing spondylitis susceptibility. Arthritis Rheum 2004;50:2246–54.
17. Brown MA, Pile KD, Kennedy LG *et al.* A genome-wide screen for susceptibility loci in ankylosing spondylitis. Arthritis Rheum 1998;41:588–95.
18. Lee YH, Rho YH, Choi SJ, Ji JD, Song GG. Ankylosing spondylitis susceptiblity loci defined by genome-search meta-analysis. J Hum Genet 2005;50:453–9.
19. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361–8.
20. Dougados M, van der Linden S, Juhlin R *et al.* The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy. Arthritis Rheum 1991;34:1218–27.
21. Reed PW, Davies JL, Copeman JB *et al.* Chromosome-specific microsatellite sets for fluorescence-based, semi-automated genome mapping. Nat Genet 1994;7:390–5.

22 Pluzhnikov A, Wen W, Luban S, Chen Y, Cox NJ. Efficient construction of genetic maps based on publicly available physical and genetic map data. Am J Hum Genet 2003: Suppl A438.

23 Kong A, Gudbjartsson DF, Sainz J et al. A high-resolution recombination map of the human genome. Nat Genet 2002;31:241–7.

24 Broman KW, Weber JL. Estimation of pairwise relationships in the presence of genotyping errors. Am J Hum Genet 1998;63:1563–4.

25 O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 1998;63:259–66.

26 Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin–rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002;30:97–101.

27 Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin–rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002;30:97–101.

28 Carter KW, Palmer LJ. Java linkage graph (online). In. http:/www.genepi.com.au/projects/jlgraph; 2004.

29 Dempfle A, Loesgen S. Meta-analysis of linkage studies for complex diseases: an overview of methods and a simulation study. Ann Hum Genet 2004;68:69–83.

30 Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and non-parametric linkage analysis: a unified multipoint approach. Am J Hum Genet 1996;58:1347–63.

31 Risch N. Assessing the role of HLA-linked and unlinked determinants of disease. Am J Hum Genet 1987;40:1–14.

32 Higgins J, Thompson S. Quantifying heterogeneity in a meta-analysis. Statistics in medicine 2002;21:1539–58.

33 Fisher SA, Abecasis GR, Yashar BM et al. Meta-analysis of genome scans of age-related macular degeneration. Hum Mol Genet 2005;14:2257–64.

34 Koivukoski L, Fisher SA, Kanninen T et al. Meta-analysis of genome-wide scans for hyptertension and blood pressure in Caucasians shows evidence of susceptibility regions on chromosomes 2 and 3. Hum Mol Genet 2004;13:2325–32.

35 van Heel DA, Fisher SA, Kirby A, Daly MJ, Rioux JD, Lewis CM, Consortium GSM-AGotIIG. Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. Hum Mol Genet 2004;13:763–70.

36 Lee YH, Rho YH, Choi SJ, Ji JD, Song GG. Ankylosing spondylitis susceptibility loci defined by genome-search meta-analysis. J Hum Genet 2005;50:453–9.