Overlap of disease susceptibility loci for rheumatoid arthritis and juvenile idiopathic arthritis

Anne Hinks,1 Steve Eyre,1 Xiayi Ke,1 Anne Barton,1 Paul Martin,1 Edward Flynn,1 Jon Packham,2 Childhood Arthritis Prospective Study (CAPS),3 UKRAG Consortium,4 BSPAR Study Group,5 Jane Worthington,1 Wendy Thomson1

ABSTRACT

Background Genome-wide association studies (GWAS) have been extremely useful in the search for susceptibility risk factors for complex genetic autoimmune diseases. As more studies are published, evidence is emerging of considerable overlap of loci between these diseases. In juvenile idiopathic arthritis (JIA), another complex genetic autoimmune disease, the strategy of using information from autoimmune disease GWAS or candidate gene studies to help in the search for novel JIA susceptibility loci has been successful, with confirmed association with two genes, PTPN22 and IL2RA. Rheumatoid arthritis (RA) is an autoimmune disease that shares similar clinical and pathological features with JIA and, therefore, recently identified confirmed RA susceptibility loci are also excellent JIA candidate loci.

Objective To determine the overlap of disease susceptibility loci for RA and JIA.

Methods Fifteen single nucleotide polymorphisms (SNPs) at nine RA-associated loci were genotyped in Caucasian patients with JIA (n = 1054) and controls (n = 3531) and tested for association with JIA. Allele and genotype frequencies were compared between cases and controls using the genetic analysis software, PLINK.

Results Two JIA susceptibility loci were identified, one of which was a novel JIA association (STAT4) and the second confirmed previously published associations of the TRAF1/C5 locus with JIA. Weak evidence of association of JIA with three additional loci (Chr6q23, KIF5A and PRKCO) was also obtained, which warrants further investigation.

Conclusion All these loci are good candidates in view of their known pathogenesis of JIA, as genes within these regions (TRAF1, STAT4, TNFAIP3, PRKCO) are known to be involved in T-cell receptor signalling or activation pathways.

INTRODUCTION

Complex genetic diseases are caused by the effects of multiple genes in combination with lifestyle and environmental factors. In the past couple of years, with the emergence of genome-wide association studies (GWAS), there has been an explosion of genetic discoveries for many complex genetic autoimmune diseases such as Crohn’s disease, type 1 diabetes, multiple sclerosis and rheumatoid arthritis (RA). Indeed, more genetic risk factors were identified for common diseases in 2007 than had been collectively reported before that date.1 It is also interesting to note that evaluating the results from the study of one disease in other complex diseases can disclose common risk factors. Thus there has been detectable overlap of loci for autoimmune diseases, with examples such as a region on chromosome 4q27 that encodes the interleukin 2 (IL2) and IL21 genes, which is associated with susceptibility to coeliac disease, multiple sclerosis, type 1 diabetes and RA,2,3 and the IL23R region on 1p31 which has been associated with Crohn’s disease,4 ulcerative colitis,4 psoriasis5 and ankylosing spondylitis.6

Juvenile idiopathic arthritis (JIA), another complex genetic autoimmune disease, is the most common chronic inflammatory rheumatic disease in children and is an important cause of short- and long-term disability.7 It is defined as arthritis of unknown aetiology that starts before the age of 16 and persists for at least 6 weeks. JIA can be subdivided into seven clinically more homogeneous subtypes, using the International League of Associations for Rheumatology (ILAR) classification system.8 As in other autoimmune diseases there are well-established associations of the HLA region with JIA.9

In JIA, only a single, comparatively small 100k single nucleotide polymorphism (SNP) study has been published to date.10 This is largely owing to the low prevalence and clinically heterogeneous nature of JIA, meaning that collection of sufficiently large and thus powerful, sample sizes has been problematic compared with other complex autoimmune diseases. However, this does not mean that data from GWAS cannot be exploited to help elucidate the underlying genetic basis of this disease. Indeed, the strategy of using information from autoimmune disease GWAS or candidate studies to help identify JIA susceptibility genes has already been successful in identifying two JIA susceptibility genes, protein tyrosine phosphatase non-receptor 22 (PTPN22)11 and more recently, IL2RA,12 which have been validated in independent cohorts.

RA is an autoimmune disease that has similar clinical and pathological features to those of JIA. In the past year a number of novel RA susceptibility genes have been identified, which have been confirmed in independent datasets, including the STAT4 gene,13 14 the TRAF1 and C5 region on chromosome 9,14–16 a region on chromosome 6q23,17 18 the KIF5A gene,19 a region on chromosome 10p15 close to the PRKCO gene,19 MME1,19 20 CD40,20 CD66C20 and a region on chromosome 19p13 close to the CCL21 gene.20 Thus, these loci are also excellent JIA candidate susceptibility loci and, therefore, the aim of this study was to determine whether
these confirmed that RA susceptibility loci are also associated with susceptibility to JIA.

**SUBJECTS AND METHODS**

**Subjects**

DNA was available for 1054 UK Caucasian patients with JIA (532 male, 715 female) from three sources: the British Society for Paediatric and Adolescent Rheumatology (BSPAR) National Repository of JIA (n=654); a cohort of UK Caucasian patients with longstanding JIA (n=201), described previously; and a third cohort collected as part of the Childhood Arthritis Prospective Study (CAPS), a prospective inception cohort study of JIA cases from five centres across the UK (n=199). JIA cases were classified according to ILAR criteria, which are a combined set of all ILAR subtypes (online supplementary table 1).

Healthy Caucasian control DNA samples were available from five centres in the UK as described previously: Manchester, 832 controls (including 228 1958 birth cohort controls); Sheffield, 929 controls; Leeds, 422 controls; Aberdeen, 825 controls; Oxford, 525 controls; total control sample size=3531). All individuals were recruited with ethical approval and provided informed consent (North-West Multi-Centre Research Ethics Committee (MREC 99/8/84) and the University of Manchester Committee on the Ethics of Research on Human Beings).

**SNP selection**

In total, 15 SNPs were selected for genotyping in nine independent genetic regions, all of which have robust confirmed evidence for association with RA. These comprised:

1. **STAT4**: Four SNPs (rs11889341, rs7574865, rs8179673 and rs10181656) located in the third intron of STAT4, which showed the strongest association with North American and Swedish RA cases; Korean RA cases were also selected.

2. **TRAF1/C5**: Three SNPs located in the TRAF1/C5 region were identified for investigation, one of which (rs10818486) showed the strongest evidence for association with Dutch RA cases. The other two SNPs, rs3761847 and rs2900180, showed the strongest evidence for association in North American and Swedish RA cases.

3. **CD40**: Association of rs4810485 on chromosome 20q13 with RA was also reported in a meta-analysis of two published GWAS.

4. **CCL21**: The SNP lies in intron 7 of the cyclin-dependent kinase 6 gene. Association of this SNP with RA was reported in the meta-analysis of two published GWAS.

5. **CDK6**: rs42041, which maps to chromosome 7q21, lies within ~0.1 kb from the CCL21 gene. Association of this SNP with RA was also reported in the meta-analysis of two published GWAS.

6. **OLIG3**

7. **TNFAIP3**: rs6920220 and rs13207033, located at the chromosome 6q23 susceptibility locus, were also investigated. The SNPs lie in an intergenic region between OLIG3 and TNFAIP3. Variants in this region have been reported to be independently associated with RA risk. The minor allele of rs6920220 is associated with increased risk of RA (minor allele frequency (MAF) > controls), whereas the minor allele of rs10499194 is associated with protection (MAF controls > cases). Although the latter was not directly genotyped, a perfect proxy SNP (rs13207033; r2=1) was.

8. **KIF5A**: Association with rs1678542, which maps to intron 15 of the KIF5A gene, on chromosome 12q13 was initially identified in the Wellcome Trust case–control consortium (WTCCC) genome-wide association screen of RA cases and controls and then validated in an independent cohort.

9. **PRKCQ**: rs4750316 maps to an intergenic region on chromosome 10p15, the closest gene, PRKCQ, being 76 kb away. Association with this SNP has been validated in independent datasets.

10. **MMEL**: Association with rs10910099, which maps to intron 11 of the MMEL (membrane metallo-endopeptidase-like) gene on chromosome 1p36 was initially identified in the WTCCC genome-wide association screen of RA cases and controls and then validated in an independent cohort.

11. **CD40**: Association of rs4810485 on chromosome 20q13 with RA was reported in a meta-analysis of two published GWAS. The SNP lies in intron 2 of the CD40 gene.

12. **CCL21**: rs2812378, which maps to chromosome 9p13, is situated ~0.1 kb from the CCL21 gene. Association of this SNP with RA was also reported in the meta-analysis of two published GWAS.

13. **CDK6**: rs42041, which maps to chromosome 7q21, lies within intron 7 of the cyclin-dependent kinase 6 gene. Association of this SNP with RA was also reported in the meta-analysis of two published GWAS.

**Genotyping**

All SNPs were genotyped in UK JIA cases and controls. Genotyping was performed using the Sequenom iPLEX platform. A 90% sample quality control rate and 90% SNP genotyping success rate was imposed on the analysis. The control genotype data have been reported previously.

**Statistical analysis**

Genotype and allele frequencies were compared between cases with JIA and controls using STATA version 9 SE (StataCorp, Texas, USA) and PLINK. Associations were expressed as two-sided p values, allelic odds ratios (ORs) and their 95% CIs. Power calculations were performed using Quanto based on previous RA studies for these SNPs. JIA is a phenotypically heterogeneous disease and can be classified into more clinically homogeneous diseases using the ILAR classification criteria. However, comparing each of the ILAR subtypes separately against controls would result in a large number of hypothesis tests. Therefore we first examined whether there was evidence of a difference in allele frequencies between the seven ILAR subtypes. Differences between subtypes were assessed using χ² tests on the 7×2 tables. Only when a difference was found (p<0.05) were separate ORs and 95% CIs calculated for the subgroups.
RESULTS
Fifteen SNPs across nine loci were genotyped in JIA cases (n=1054) and controls (n=3531) and analysed for association with JIA.

Table 1 shows the power of this study to detect the effect sizes identified in the RA studies for these SNPs, at a 5% significance level, assuming a log-additive model. As expected, despite being the largest collection of JIA cases and controls investigated to date, for many of the SNPs this study had modest power to detect an effect. A Bonferroni correction of 15 was applied to correct for the number of SNPs studied, resulting in a p value threshold of 0.003 for claims of significance.

Despite the modest power in this study, all of the four STAT4 SNPs were associated with JIA, with the strongest association for rs8179673 (allelic OR=1.25, 95% CI 1.11 to 1.4, trend p=0.0002) (table 2). These SNPs are in strong linkage disequilibrium with each other (r²>0.96). One of the three TRAF1/C5 SNPs was significantly associated with JIA, rs2900180 (allelic OR=1.21, 95% CI 1.09 to 1.35, trend p=0.0003). The associated SNP was the one for which the sample size tested had most power to detect statistical evidence for association.

The chromosome 6q23 SNP, rs6920220, also showed weak association with JIA (allelic OR=1.16, 95% CI 1.02 to 1.31, trend p=0.02) (table 2). As with RA, the SNP, rs15207053, showed a weak protective association with JIA (allelic OR=0.87, 95% CI 0.77 to 0.98, trend p=0.02). Both the KIF5A SNP, rs1678542 and the SNP, rs4750316, close to PRKCCQ, showed weak protective association with JIA (allelic OR=0.91, 95% CI 0.82 to 1.01, trend p=0.06 and allelic OR=0.88, 95% CI 0.77 to 1.0, trend p=0.05, respectively), although the evidence for association did not reach the corrected threshold for any of these loci.

The markers rs10910099 in the MMEL gene, rs42041 in the CDK6 gene, rs4810485 in the CD40 gene and rs2812378, close to the CCL21 gene, showed no association with JIA (table 2).

The associated SNPs were examined for differences in allele frequencies between the seven ILAR subtypes. In all cases this was not significant (p>0.05) (data not shown). Therefore further stratification by ILAR subtype was not performed.

Figure 1 shows a comparison between the association analysis results in RA and JIA. For all the JIA-associated SNPs the same allele is associated with JIA as was associated with RA and effect sizes are similar.

DISCUSSION
In this study we examined nine recently identified RA loci, which have been confirmed in independent datasets, for association with JIA. Using this approach we have identified two JIA susceptibility loci, of which one was a novel JIA association (STAT4); while the evidence for a true association for STAT4 is compelling, it will require validation in independent JIA case–control cohorts, to confirm it as a JIA susceptibility locus. The second association confirms previously published associations of the TRAF1/C5 locus with JIA susceptibility.27 28 However, previous studies of this locus have provided conflicting evidence for association. For example, a recent study of Dutch JIA cases and controls investigated the rs10818488 SNP in the TRAF1/C5 region,27 and found no evidence for association of this SNP in patients with JIA as a whole. Interestingly, in our UK study we also found no statistically significant evidence for association of this SNP in all JIA cases but there was a trend (allelic OR=1.09, 95% CI 0.98 to 1.21, trend p=0.1).

We performed a meta-analysis of the two studies which yielded significant evidence for association (OR=1.09, 95% CI 1.10 to 1.2, p=0.05) with no evidence for heterogeneity between the two cohorts (p=0.9). A second, small study of SNPs within TRAF1/C5 region in USA JIA cases also reported association with this region. The strongest effect in that study was with rs376187.20 This SNP was not associated in our current study, although there was association with another SNP (rs2900180). The lack of consistency between studies of TRAF1/C5 region and JIA susceptibility may be due to the lack of power of individual studies to detect the modest effect sizes reported at this locus. For the three SNPs in this study in the TRAF1/C5 region we had 86% power to detect an effect for the SNP rs2900180 and we found that this was the most strongly associated SNP with JIA for this region. For the other two SNPs we had ~60% power to detect an association and therefore the lack of association may be a false-negative result due to small sample size. This is also compounded by the fact that, as yet, the causal SNP(s) in this region has not been determined.

Although not statistically significant, we found weak evidence of association with JIA of SNPs in the chromosome 6q23 region, within KIF5A and close to PRKCCQ; in each case the effect sizes were similar to those seen in RA. Thus, given the relatively low
power within our study, they warrant further investigation in independent datasets.

We found no significant association with the SNPs in the MME1, CDK6, CCL21 or the CD40 genes, which may suggest that these genes are RA-specific loci. However, as the effect sizes for these genes in RA were modest, this study had limited power to detect an association with JIA and MME1, CDK6 and CCL21 (53%, 69% and 59% respectively). For three of these SNPs, the direction of association was opposite in the JIA samples compared with the RA studies. Larger datasets and meta-analyses will be required to completely rule out association of these SNPs with JIA.

One limitation of the study may be the use of a common set of controls between this study and two of the previous RA validation studies. This is a conventional strategy, as used by the WTCCC GWAS, to minimise genotyping, but it may be a concern. All the control genotypes used in this study were compared with the control data from the WTCCC and there was no significant difference (p<0.05) in the frequencies, suggesting these controls accurately represent UK population frequencies.

It is interesting to note that for the JIA susceptibility loci identified, the associated allele is the same and effect size similar to that seen in studies of RA (figure 1).

This finding, combined with data from previous studies, suggests that for many complex diseases there will be some shared susceptibility loci while others will be disease specific. In JIA, given the clinical heterogeneity, the picture may be even more complicated. Current classification criteria divide JIA into seven subgroups and it would be interesting to determine whether the susceptibility loci identified are common to all subgroups or restricted to only some subgroups. However, stratified analysis leads to small sample sizes for many of the subgroups as well as introducing the problem of multiple testing and thus makes interpreting these results challenging. In other cohorts it will be important to combine data in order to improve power and accurately determine whether the associations seen are consistent between the different clinical subgroups of JIA.

In RA, there is some evidence to suggest that there are distinct genetic risk factors for anti-cyclic citrullinated peptide (anti-CCP)-positive and anti-CCP-negative disease. In the UK population the TRAF1/C5 association did appear stronger in the anti-CCP-positive subgroup than in the negative subgroup; however, the strength of association of the STAT4 locus was similar in both subgroups. There is little evidence for anti-CCP antibodies in JIA, apart from in the rheumatoid factor-positive polyarthritis subgroup, though antinuclear antibodies (ANA) are often observed. However, larger sample sizes will be required to investigate any specific associations with ANA-positive and ANA-negative subgroups.

It is important to note that for all these loci the causal variant, indeed in some cases the gene, has yet to be determined. Re-sequencing and fine-mapping studies are underway in order to further characterise the genetic effects before functional studies can be undertaken. However, all these loci represent good candidates for JIA as genes within these regions (TRAFl, STAT4, TNAFla, PRKCC) known to be involved in T-cell receptor signalling or T-cell activation pathways.

In conclusion, by investigating the overlapping of RA and JIA susceptibility loci, we have identified a novel JIA susceptibility locus (STAT4) and confirmed association with a second, TRAF1/C5. It is clear that this strategy is useful in identifying autoimmune/arthritis-associated JIA susceptibility loci. However, it is likely that there are many more JIA-specific or JIA-subgroup specific genes yet to be identified and well-powered GWAS will be required to identify these.

Acknowledgements We thank David Strachan for facilitating access to the 1958 birth cohort. We acknowledge use of genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 088545/2/02.

Competing interests None.

Funding This work was supported by the Arthritis Research Campaign: arc grant reference number 17552.

Ethics approval This study was conducted with the approval of the North-West Multi-Centre Research Ethics Committee [MREC 99/6/04] and the University of Manchester Committee on the Ethics of Research on Human Beings.

Provenance and peer review Not commissioned; externally peer reviewed.

Patient consent Obtained.

REFERENCES