Variants in linkage disequilibrium with the late cornified envelope gene cluster deletion are associated with susceptibility to psoriatic arthritis

John Bowes,1 Edward Flynn,1 Pauline Ho,1,2 Batool Aly,1 Ann W Morgan,3 Helena Marzo-Ortega,3 Laura Coates,3 Ross McManus,4 Anthony W Ryan,4 David Kane,5 Eleanor Korendowych,6,7 Neil McHugh,6,7 Oliver FitzGerald,8 Jonathan Packham,9 Ian N Bruce,1,2 Anne Barton1,2

ABSTRACT

Objective A common deletion mapping to the psoriasis susceptibility locus 4 on chromosome 1q21, encompassing two genes of the late cornified envelope (LCE) gene cluster, has been associated with an increased risk of psoriasis vulgaris (PsV). One previous report found no association of the deletion with psoriatic arthritis (PsA), suggesting it may be a specific risk factor for PsV. Given the genetic overlap between PsA and PsV, a study was undertaken to investigate whether single nucleotide polymorphisms (SNPs) mapping to this locus are risk factors for PsA in a UK and Irish population.

Methods Three SNPs with prior evidence of association with susceptibility to PsV were genotyped in 1057 patients with PsA using Sequenom iPLEX chemistry and genotype frequencies compared with data available for 5575 healthy controls. Two of the SNPs, rs4112788 and rs4085613, were reported to be highly correlated with the LCE deletion. The third SNP, rs6701216, was previously reported to be associated with PsV in a US population.

Results Alleles tagging the deletion for both rs4112788 and rs4085613 were found to be enriched in cases compared with controls (69% vs 65%) and significantly associated with increased susceptibility to PsA (p trend = 0.001, OR 1.19 and p trend = 0.001, OR 1.18, respectively). No association was observed with rs6701216.

Conclusions The evidence presented here supports LCE deletion as a risk factor for PsA in a UK and Irish population. It suggests that this locus is a risk factor within a shared aetiological pathway that contributes to psoriatic skin disease in both PsV and PsA.

INTRODUCTION

Psoriatic arthritis (PsA) is an inflammatory joint condition associated with psoriasis vulgaris (PsV); it is estimated that up to 30% of patients with PsV also have articular disease. PsA is considered to be a complex disease with contributions from both environmental and genetic risk factors.1 The genetic component is estimated to be strong, with a sibling recurrence RR (λ s) of 47 in a UK population.2 Genetic investigations have revealed a number of risk factors that are common to both PsV and PsA. These established risk loci, HLA-Cw6, IL12B and IL23R, do not allow them to be genetically discriminated.3 4 A number of studies have recently identified genetic variants mapping to chromosome 1q21 within the previously identified psoriasis susceptibility locus 4 as being associated with susceptibility to PsV. A genome-wide scan for copy number variation identified an association between a 32.2 kb deletion (LCE3C_LCE3B-del) that encompasses two genes of the late cornified envelope (LCE) gene cluster (LCE3B and LCE3C), with increased susceptibility to PsV in multiple populations.3 This deletion was found to be highly correlated with the single nucleotide polymorphism (SNP) rs4112788 (r 2 =0.93) with the major allele (C) tagging LCE3C_LCE3B-del. This association is further supported by evidence from an independent SNP genome-wide association study (GWAS) for PsV that reported association for rs4085613 mapping to this locus.6 The SNPs rs4112788 and rs4085613 are in high linkage disequilibrium (LD) (r2>0.95) and, as such, are not considered to be independently associated. A further independent GWAS investigating susceptibility to PsV identified an association to rs6701216 that maps to the LCE1C gene,7 which represents an effect independent of rs4112788 and rs4085613 as there is little LD (r2<0.05) between the variants.9 From these genetic studies alone it is not clear whether the deletion itself or the SNPs in LD with it are primarily responsible for the association seen, and functional studies will be required to determine the mechanism for the disease predisposition.

Intriguingly, unlike the previously identified psoriasis loci, LCE3C_LCE3B-del and correlated SNPs were associated with PsV but not PsA susceptibility in a German cohort, suggesting this may be the first risk factor identified that predisposes to skin-type psoriasis only, implying that the skin disorder in PsV and PsA have differing underlying genetic aetiologies.8 9 Identification of genetic factors that differentiate the two phenotypes could help to reveal distinct aetiological pathways. Validation of such a finding in independent studies is therefore essential, and the aim of the work presented here was to investigate association with SNP markers in LD with LCE3C_LCE3B-del in a large case–control study of subjects with PsA and controls from the UK and the Republic of Ireland (ROI).
Extended report

METHODS

Patient samples
Genomic DNA was available from 1057 Caucasian patients with PsA recruited from three UK rheumatology centres and one centre in the ROI (885 patients in the UK and 172 in the ROI), details of which have been described previously.10–12 PsA was defined as the presence of both psoriasis and inflammatory arthritis regardless of rheumatoid factor status, and all had peripheral arthritis. The majority of samples satisfied the CASPAR (CLASsification criteria for Psoriatic ARthritis) classification system, although some were collected prior to the introduction of this classification system.13

SNP selection
We selected three SNPs for genotyping that map to the 1q21 locus (see figure 1 in online supplement). Two SNPs, rs4085613 and rs4112788, have previously been shown to be in high LD with LCE3C_LCE3B-del and were reported to be the most associated variants from two independent studies.5,6 The third SNP, rs6701216, represents a reported effect that is potentially independent of LCE3C_LCE3B-del.7 In addition, we genotyped 12 autosomal SNPs that were reported to be informative of geographical variation across the UK.14

Control samples
Data for healthy controls were sourced from the Wellcome Trust Case-Control Consortium 2 project (www.wtccc.org.uk). This dataset consists of samples from the 1958 British Birth Cohort and the UK Blood Service Collection genotyped on the Illumina Human1M-Duo BeadChip or Affymetrix Human SNP Array 6.0. In addition, 375 control samples were available from the ROI.

Genotyping
SNP genotyping of the PsV and ROI control samples was performed using Sequenom’s MassARRAY system (San Diego, California, USA) according to the manufacturer’s specifications for the iPLEX chemistry using 10 ng genomic DNA. Genotype cluster plots were evaluated prior to analysis to ensure satisfactory assay performance.

Statistical analysis
Statistical analyses were performed using the PLINK software package.15 Quality control (QC) inclusion thresholds for both individual sample and SNP genotyping missing data rates were set at >90%. Test statistics were calculated for deviation from Hardy–Weinberg equilibrium (HWE) using an exact test, the Cochran-Armitage trend test, ORs (including 95% CI) and LD (r²). The trend test was selected as the most appropriate genetic model as previous evidence in psoriasis studies suggested a dosage effect based on allele carriage,7 thus providing the maximal power for association testing.16

The primary analysis investigated the dataset in its entirety with UK and ROI samples treated as one common population. In the interests of identifying heterogeneity between the UK and ROI cohorts, we analysed these components separately. The results were subsequently combined using an inverse variance meta-analysis under the assumption of fixed effects. Significant heterogeneity between the two groups was estimated using the Cochran Q and P statistics.

Previous work in PsV has found that the major genetic susceptibility locus, HLA-Cw*06, is associated with early onset (type I PsV, age at onset <40 years) but not late onset (type II) PsV. Stratification analysis was therefore undertaken to explore whether any associations were restricted to one subphenotype. Patients with PsA are usually but not always seronegative for rheumatoid factor. It is possible that those with a positive rheumatoid factor are actually patients with rheumatoid arthritis and co-existent psoriasis, which may bias any analysis. Subphenotype analysis was therefore also undertaken in patients with PsA seronegative for rheumatoid factor to determine whether the presence of this antibody was a significant confounder. All subphenotype analyses were performed in UK samples only. No information was available regarding the extent of psoriasis or the presence of spondyloarthritis.

Finally, we analysed data from 12 autosomal SNPs identified as showing significant variation across geographical regions of the UK.14 Our interpretation of these results would be that multiple associations to these SNPs is indicative of regional ascertainment bias of case samples given that control samples are known to be collected from across the UK. This analysis was restricted to the UK samples.

RESULTS

Genotyping
All LCE SNPs demonstrated satisfactory clustering of genotypes with clear and highly distinct clusters. Following the application of the data quality control criteria, the dataset consisted of a maximum of 982 cases and 5574 controls with a genotyping success rate of 99.9% in the remaining samples. The genotyping for the 12 SNPs with prior evidence of geographical variation across the UK consisted of a maximum of 819 cases and 5379 controls following QC with a genotyping success rate of 99.9% in the remaining samples.

Statistical analysis
For all LCE SNPs in the primary analysis the control sample groups conformed to HWE (table 1). We observed similar genotype frequencies for rs4112788 and rs4085613 and they were found to be highly correlated (D=1, r²=0.99). Allele frequencies were found to be within the range of previously reported frequencies for the Caucasian population.5,8 We observed enrichment of deletion-correlated alleles in the case group compared with the control group (69% vs 65% for both SNPs). This enrichment was found to be significantly associated with PsA susceptibility under an additive model (table 1). The association remained when the analysis was restricted to the UK-only sample set (rs4112788, p_trend=0.008; rs4085613, p_trend=0.009). No association to any SNP was observed in the ROI-only analysis and deviation from HWE (p=0.02) was noted in this control group. Meta-analysis of the independent UK and ROI results supported the association (rs4112788, p=0.005; rs4085613, p=0.006) with no evidence of heterogeneity (Q=0.95, I²=0) for both SNPs. No evidence was found to support association to rs6701216.

Stratification analysis did not reveal qualitatively different associations in subgroups of patients with PsA with type I or type II PsV or in those seronegative for rheumatoid factor (table 2). It is important to note that the sample numbers were reduced owing to incomplete phenotype data.

Analysis of the 12 geographical markers revealed association to only a single SNP, rs3873375 (p_trend=1.53×10−7) (see table 1 in online supplement).

DISCUSSION
We have found that SNPs previously reported to be in high LD with LCE3C_LCE3B-del on chromosome 1q21 are associated...
with PsA in a large case–control cohort recruited from the UK and ROI.

This deletion is of particular interest for psoriatic phenotypes as the genes of the LCE cluster encode stratum corneum proteins that play an important role in epidermal terminal differentiation, disruption of which has been proposed to lead to abnormal differentiation and psoriatic skin lesions.\(^5\) We did not find evidence to support an independent association to rs6701216. Stratification analysis by subphenotype demonstrated no qualitative difference between type I and type II psoriatic subphenotypes when considering the ORs. However, no association was observed in the type II group, which is likely to be attributable to low power owing to reduced sample size. A stronger effect was observed in the seronegative subphenotype, but this could be due to imprecision in the effect estimates owing to smaller sample numbers reflected by the wider CIs compared with the analysis of the whole UK group.

Table 1
Summary of genotype and association results for the three SNPs mapping to the 1q21 locus

<table>
<thead>
<tr>
<th>SNP</th>
<th>Primary analysis</th>
<th>UK analysis</th>
<th>Meta-analysis (IV)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Genotype</td>
<td>Control</td>
<td>Case</td>
<td>HWE</td>
</tr>
<tr>
<td>rs6585613 CC</td>
<td>1914 (34.0)</td>
<td>3755 (36.0)</td>
<td>0.98 (0.90 to 1.08)</td>
<td>2.02</td>
</tr>
<tr>
<td>rs6585613 CT</td>
<td>2190 (41.4)</td>
<td>3755 (36.0)</td>
<td>0.88 (0.79 to 0.98)</td>
<td>2.02</td>
</tr>
<tr>
<td>rs6585613 TT</td>
<td>1410 (27.0)</td>
<td>3755 (36.0)</td>
<td>0.98 (0.90 to 1.07)</td>
<td>2.02</td>
</tr>
</tbody>
</table>

Analysis of the complete dataset is presented in the primary results. Results for the UK and ROI are presented individually and combined via inverse variance meta-analysis. Genotype counts are presented with frequencies in parentheses. The C alleles of rs4112788 and rs4085613 are in high linkage disequilibrium with the LCE3C_LCE3B-del and HLA-Cw6 allele.

Our findings differ from the results of a recent study performed in a German population which found no evidence to support such an association in PsA.\(^5\)\(^9\) This discrepancy could be explained by any one of three scenarios. First, it may be a false positive (type I error) in our study. However, we have genotyped two highly correlated SNPs that demonstrate comparable results, thus making genotyping error an unlikely factor. It is also important to note that our genotype frequencies, p values and effect estimates for rs4112788 closely match the German study. While the study had 95% power to detect the maximal reported effect size, there is still potential for type II errors in such well-powered studies. This point is well illustrated with examples from a recent GWAS in the UK.\(^14\) This study failed to identify two robust rheumatoid arthritis risk loci—tumour necrosis factor receptor-associated factor 1 and signal transducer and activator of transcription 4—which were found in linkage disequilibrium with the LCE3C_LCE3B-del and HLA-Cw6 alleles.\(^5\)\(^18\) The clinical characteristics of the two cohorts may be substantially different. For example, they may consist of different proportions of type I and type II psoriasis or HLA-Cw6 positivity, with association restricted to a particular subphenotype. This may be of particular importance if the evidence for an epistatic interaction between LCE3C_LCE3B-del and HLA-Cw6 holds true upon further validation.\(^5\)\(^18\) Although the clinical characteristics of the British Isles and German cohorts appear similar, we do not currently have the HLA-Cw6 data that would allow us to thoroughly explore this possibility.

This study has two potential limiting factors. The first is the risk of confounding by population stratification as the cohort is drawn from two populations of European ancestry. This issue was addressed in a number of ways. First, we reduced the potential for heterogeneity by restricting the analysis to only UK samples. Both rs4112788 and rs4085613 remained significantly associated with a marginal fall in significance attributable to the smaller sample size. No association was observed in the ROI-only analysis, conceivably due to the small sample size. Second, we attempted to control for heterogeneity by employing an inverse variance meta-analysis on summary statistics from the individual UK and ROI dataset. Both SNPs retained significant evidence for association with no evidence for heterogeneity, thus supporting the use of a combined dataset. Although the Cochran Q and P\(^2\) test statistics are not informative of underlying population structure, they would...
reveal differences in allele frequencies between the groups with the potential to confound. Finally, in the UK cohort we analysed 12 autosomal SNPs known to be informative of geographical region. Only one SNP was found to be associated (rs873375), but this maps 11 kb upstream of HLA-C which is a known PsV risk locus. From this we conclude that there is no evidence of regional ascertainment bias between the UK case and control groups. More sophisticated forms of correction for population stratification do exist—for example, correcting with principal component analysis. However, these methods typically require dense SNP datasets, such as those emerging from GWAS, which are currently not available for this cohort. Given the frequency differences between the UK and ROI controls, the most robust association statistics could be considered to be those emerging from the UK-only analysis based on the larger sample size.

The second potential issue is that our study did not directly type LCE3C_LCE3B-del, but we did genotype two SNPs shown in previous studies to be in high LD with the deletion (rs4112788, r2>0.90). The use of SNPs as accurate proxies for deletions is supported by evidence from studies exploring their type with principal component analysis. However, these methods data for the deletion would represent the ideal scenario as any deviation of the proxy from perfect LD would result in the loss of power. However, studies have highlighted the technical difficulties involved in the accurate typing of copy number variants using methods such as quantitative PCR and the paralogue ratio test. Indeed, the accurate genotyping of LCE3C_LCE3B-del in a recent PsV study required the application of multiple methods to achieve an acceptable success rate. We therefore preferred the robust technology of SNP genotyping, accepting a potential minor reduction of power.

Assessing the published reports to date, it is unclear if the variant traits are SNPs, the deletion or a tagged unknown variant. A comprehensive fine mapping study of the region is therefore required that incorporates both SNPs and structural variation in order to identify the causal variant(s) before proceeding to functional studies aimed at understanding the mechanism by which the variants predispose to PsV, and whether these differ in any way from those predisposing to PsA. In turn, this may open up the possibility of different therapies being developed based on the novel target.

In conclusion, this study provides further genetic support that the psoriasis seen in dermatology clinics does not differ from that seen by rheumatologists, and that patients with PsA can be included in genetic studies of psoriasis per se. This example also further illustrates the importance of large sample sizes for the detection of the modest effect sizes typical of complex traits.

Acknowledgements The authors acknowledge the support of the NIHR Manchester Biomedical Research Centre and NIHR Leeds Musculoskeletal Biomedical Research Unit. This study makes use of data generated by the Wellcome Trust Case-Control Consortium. A full list of the investigators who contributed to the generation of the data is available from www.wtcg.org.uk.

Funding JB, INB and AB are funded by Arthritis Research UK (arc grant 17552). EF is supported by the European Community’s Sixth Framework Programme AutoCure funding. Funding for the project was provided by the Wellcome Trust under award 076113 and 085475. This work was funded by the Arthritis Research UK.

Ethics approval This study was conducted with the approval of the Central Manchester Research Ethics Committee MREC 99/8/84 and all subjects provided informed consent.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES


Table 2 Summary of results for phenotype subgroup analysis

<table>
<thead>
<tr>
<th>SNP</th>
<th>Type I psoriasis (n=519)</th>
<th>Type II psoriasis (n=176)</th>
<th>Seronegative (n=228)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs4085613</td>
<td>CC 234 (45.2) 1.16 (1.01 to 1.33)</td>
<td>85 (48.3) 1.15 (0.92 to 1.45)</td>
<td>157 (47.9) 1.23 (1.04 to 1.47)</td>
</tr>
<tr>
<td></td>
<td>CA 240 (46.3) 0.03</td>
<td>70 (39.9) 0.22</td>
<td>143 (43.6) 0.015</td>
</tr>
<tr>
<td></td>
<td>AA 44 (8.3)</td>
<td>21 (11.9)</td>
<td>28 (8.5)</td>
</tr>
<tr>
<td>rs4112788</td>
<td>CC 233 (44.9) 1.16 (1.01 to 1.33)</td>
<td>85 (48.3) 1.17 (0.93 to 1.47)</td>
<td>156 (47.6) 1.24 (1.04 to 1.47)</td>
</tr>
<tr>
<td></td>
<td>CT 243 (46.8) 0.03</td>
<td>71 (40.3) 0.18</td>
<td>145 (44.2) 0.014</td>
</tr>
<tr>
<td></td>
<td>TT 43 (8.3)</td>
<td>20 (11.4)</td>
<td>27 (8.2)</td>
</tr>
</tbody>
</table>

Control frequencies are those reported in the UK analysis reported in table 1. frq, frequency; n, number in subgroup; SNP, single nucleotide polymorphism.

