# Conditional analysis identifies three novel major histocompatibility complex loci associated with psoriasis

**Jo Knight[1,2], Sarah L. Spain[3], Francesca Capon[3], Adrian Hayday[4], Frank O. Nestle[3], Alex Clop[3], Wellcome Trust Case Control Consortium[†], Genetic Analysis of Psoriasis Consortium[†], I-chip for Psoriasis Consortium[†], Jonathan N. Barker[3,5], Michael E. Weale[3,‡], and Richard C. Trembath[3,6,∗,‡]**

[1]National Institute for Health Research (NIHR), Biomedical Research Centre, Guy's and St Thomas' NHS Foundation Trust, [2]Neuroscience Research, Centre for Addiction and Mental Health, Toronto, ON, Canada M5T 1R8, [3]Division of Genetics and Molecular Medicine, King's College London School of Medicine, Guy's Hospital, London, UK, [4]Division of Immunology, Infection and Inflammatory Disease and [5]St John's Institute of Dermatology, King's College London, London, UK and [6]Queen Mary University of London, Barts and the London School of Medicine and Dentistry, London, UK

**Psoriasis is a common, chronic, inflammatory skin disorder. A number of genetic loci have been shown to confer risk for psoriasis. Collectively, these offer an integrated model for the inherited basis for susceptibility to psoriasis that combines altered skin barrier function together with the dysregulation of innate immune pathogen sensing and adaptive immunity. The major histocompatibility complex (MHC) harbours the psoriasis susceptibility region which exhibits the largest effect size, driven in part by variation contained on the *HLA-Cw*∗*0602* allele. However, the resolution of the number and genomic location of potential independent risk loci are hampered by extensive linkage disequilibrium across the region. We leveraged the power of large psoriasis case and control data sets and the statistical approach of conditional analysis to identify potential further association signals distributed across the MHC. In addition to the major loci at *HLA-C* ($P = 2.20 \times 10^{-236}$), we observed and replicated four additional independent signals for disease association, three of which are novel. We detected evidence for association at SNPs rs2507971 ($P = 6.73 \times 10^{-14}$), rs9260313 ($P = 7.93 \times 10^{-09}$), rs66609536 ($P = 3.54 \times 10^{-07}$) and rs380924 ($P = 6.24 \times 10^{-06}$), located within the class I region of the MHC, with each observation replicated in an independent sample ($P \leq 0.01$). The previously identified locus is close to *MICA*, the other three lie near *MICB*, *HLA-A* and *HCG9* (a non-coding RNA gene). The identification of disease associations with both *MICA* and *MICB* is particularly intriguing, since each encodes an MHC class I-related protein with potent immunological function.**

## INTRODUCTION

Psoriasis is a common, chronic, inflammatory skin disease characterized by painful, red, indurated scaly plaques that vary in extent yet can cover the entire skin surface. The underlying cutaneous pathological features are epidermal proliferation, increased vascularity and inflammation. Chronic psoriasis is associated with several important co-morbidities that include the development of a poly-articular arthritis and an increased predisposition to cardiovascular disease (1).

Recent progress has been made towards elucidating the genetic architecture of psoriasis (2). Both Linkage and

---

**Table 1.** Independent association signals in the discovery arm from stepwise conditional analysis, using both SNPs and imputed classical alleles

| ID | Position (bp), Build 37 | Total MAF (case–control) | Single marker P-value | Conditional P-value* | Odds ratio (95% CI) | Full-model P-value | Full-model odds ratio (95% CI) |
|---|---|---|---|---|---|---|---|
| *HLA-Cw*0602* | 31 236 526– 31 239 913 | 0.16 (0.31−0.09) | $2.20 \times 10^{-236}$ | $2.20 \times 10^{-236}$ | 4.37 (3.98−4.79) | $3.64 \times 10^{-148}$ | 4.75 (4.48−5.05) |
| **rs2507971** | 31 461 372 | 0.37 (0.31−0.39) | $8.04 \times 10^{-26}$ | $6.73 \times 10^{-14}$ | 1.45 (1.35−1.57) | $1.84 \times 10^{-8}$ | 1.30 (1.24−1.36) |
| rs142291993 | 31 386 333 | 0.02 (0.02−0.01) | $5.00 \times 10^{-8}$ | $1.15 \times 10^{-13}$ | 1.74 (1.35−2.25) | $2.57 \times 10^{-12}$ | 3.92 (3.23−4.77) |
| **rs9260313** | 29 916 885 | 0.26 (0.22−0.28) | $3.60 \times 10^{-14}$ | $7.93 \times 10^{-09}$ | 0.71 (0.66−0.78) | $1.18 \times 10^{-5}$ | 0.78 (0.73−0.82) |
| **rs66609536** | 31 362 120 | 0.26 (0.34−0.22) | $2.21 \times 10^{-58}$ | $3.54 \times 10^{-07}$ | 1.85 (1.71−2.00) | $1.23 \times 10^{-7}$ | 1.32 (1.26−1.40) |
| **rs380924**[a] | 29 939 885 | 0.37 (0.30−0.40) | $8.75 \times 10^{-29}$ | $6.24 \times 10^{-06}$ | 1.55 (1.43−1.67) | $3.88 \times 10^{-6}$ | 1.25 (1.19−1.31) |

MAF, minor allele frequency.
[a]SNP rs380924 did not reach the threshold of significance required in this analysis but it did in the analysis of the data without the classical alleles and was therefore taken forward for replication. Boldfaced SNPs were taken forward for replication.
*P-value from stepwise regression conditional on all SNPs in rows above.

association studies have provided robust evidence for a major disease susceptibility locus within the major histocompatibility complex (MHC) termed *PSORS1* (psoriasis susceptibility locus 1), corroborating early serological studies which consistently showed an association between psoriasis and the *HLA-Cw6* serotype. *PSORS1* has been estimated to account for between 35 and 50% of heritability (3). Mapping studies have identified a 250 kb critical interval within the class I MHC region, containing at least nine genes, including *HLA-C*. Identification of the causal DNA variation within *PSORS1* has been compounded, at least in part, by the extensive long-range linkage disequilibrium (LD) present across the MHC region of the short arm of chromosome 6. Substantial genetic evidence now points to sequence variation, across the locus harbouring the *Cw*0602* allele, as the main causal determinant of psoriasis susceptibility in this region. For example, a combined sequencing and haplotype mapping analysis and two family-based association studies have found that only *HLA-C* contained coding variants unique to the risk haplotype (4–6). Several genome-wide association studies (GWAS) for psoriasis have now been reported (7–16). These studies have confirmed that by far the most significant association signal across the genome is for SNPs in strong LD with *HLA*Cw0602*. Our own recent observation (13) demonstrated a genetic interaction between *ERAP-1* and a SNP tagging *HLA-Cw*0602* providing an independent link for a role for *HLA-C*, as *ERAP-1* is involved in peptide trimming prior to presentation by class I HLA proteins.

In addition to the *HLA-C* locus, fine mapping of the MHC region has provided evidence of two further independent association signals as well as evidence of the involvement of other HLA classical alleles (17). Here we sought to determine whether further loci could be revealed through the analysis of a data set some 2-fold larger, and with a greatly increased SNP density, than hereto analysed (13). We sought via this approach to provide insight into loci with multiple causal variants within a gene or region, as seen in a number of common complex disorders, and to inform ongoing studies of the pathobiology of psoriasis.

## RESULTS

We used data from our previously reported GWAS, composed of 2178 psoriasis cases and 5175 geographically matched controls, as a discovery analysis set (13). Analysis was performed on all SNPs passing quality control within the extended MHC region, defined as the ~8 Mb interval between *SCGN* and *RPL12P1* (chr6: 25 652 429–33 368 333). The classical MHC region lies between *C6orf40* and *HCG24* (chr6: 29 640 147– 33 115 544) and is split into class I (chr6: 29 640 147– 31 478 898), class II (chr6: 31 478 898–32 191 844) and class III (chr6: 32 191 844–33 115 544) (18). (Chromosome positions are in Genome Reference Consortium Build 37.) After the imputation of classical alleles and SNPs and the removal of redundant SNPs ($r^2 = 1$ with another SNP), the data set contained 58 015 SNPs (of which 2191 had been experimentally genotyped) together with 48 variables representing the classical HLA alleles (14 alleles for both *HLA-A* and *HLA-B* and 20 alleles from *HLA-C*).

Our primary analysis of the discovery sample was a stepwise conditional regression on the full data set consisting of typed SNPs, imputed SNPs and classical alleles. We detected five loci with *P*-values for association of $<1 \times 10^{-6}$. LD between each was low ($r^2 < 0.01$ for all pairwise comparisons in controls), supporting the hypothesis that they represented independent loci (Table 1 and Fig. 1). Only one of these signals represents a classical allele. When all SNPs are included in the models, the *P*-values for each SNP are still $\leq 1 \times 10^{-5}$, also supporting the hypothesis that they are independent loci. We next repeated the stepwise conditional analysis without the imputed classical alleles in our data set and four of the five loci gave concordant results, that is at each locus a SNP or a proxy ($r^2 > 0.98$) had a similar *P*-value (Supplementary Material, Table S1). The SNP rs142291993 has a low minor allele frequency, raising concerns for the reliability of the imputation and was not taken forward for replication. In addition to the confirmation of four loci, another SNP, rs380924, had a *P*-value of $<1 \times 10^{-6}$; as the *P*-value for this SNP from the primary analysis ($P = 6.24 \times 10^{-6}$) was close to the cut-off for significance, we took this SNP forward for replication. We next repeated the stepwise conditional analysis with only the imputed classical alleles, and found five independent hits (including *HLA-C*0602*) at $P < 1 \times 10^{-6}$ (Table 2) . Finally, we conditioned on these five classical alleles and looked for further independent signals in SNP data. We found two SNPs with independent signals (Table 2), indicating that the imputed classical alleles alone
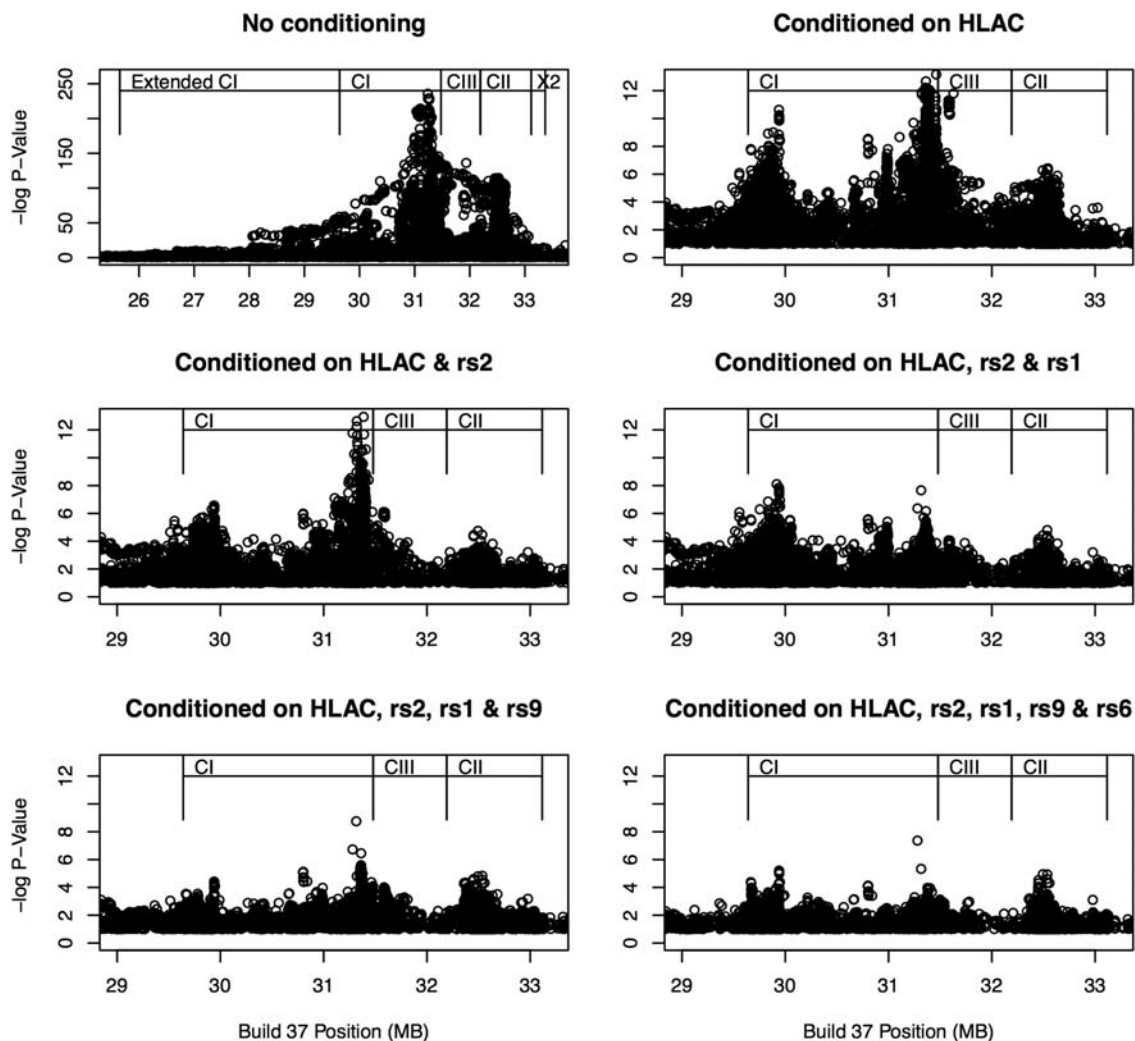
**Figure 1.** Large-scale conditional regional association plots of independent MHC signals. The first panel shows the association signals with no conditioning. Subsequent panels show association signals conditioned on the most significant SNPs from each previous panel (HLAC, *HLA-Cw*0602*; rs2, rs2507971; rs1, rs142291993; rs9, rs9260313 and rs6, rs66609536. For all plots, the *y*-axis is the −log 10 of the *P*-value and the *x*-axis the base pair position of the SNP. CI, CII and CIII, classes I, II and III; Extended C1, the extended class I; and X2, the extended class II region.

cannot account for all the association signals in the data. One of these SNPs appears to represent the same signal as a SNP identified our primary analysis. rs2517670 lies within 1.2 kb of rs380924 and has an $r^2$ of 0.93. Once the classical alleles and the two additional SNPs are included in the model, nothing else emerges with a conditional *P*-value $<1 \times 10^{-6}$, although rs2507971 (also from our primary analysis) has a *P*-value of $4.43 \times 10^{-6}$.

Our replication cohort consisted of 12 107 samples composed of 2913 cases and 9194 controls. All five independent signals were replicated by proxy SNPs contained within the Immunochip data set at values of $P = 0.01$ or below (Table 3). When all *P*-values were included in the model, the *P*-value for SNP rs3130922 became non-significant. However, this is the least effective proxy, so the combination of signal reduction and additional degrees of freedom reduces the signal strength. As expected, the strongest evidence for association appeared with the imputed *HLA-Cw*0602* allele (Fig. 2). The other SNPs are

located near *MICB* (rs2507971), *HLA-A* (rs9260313) and *HCG9* (rs380924) (Fig. 3). The rs66609536/rs13437088 SNP near *MICA* has been reported in a study by Feng *et al.* (17). We looked for interaction between each of the five independent association signals and SNPs from the GWAS. None was significant at the genome-wide level. The previously reported and replicated interaction between *ERAP* and *HLA-C* had a *P*-value of $10^{-5}$.

Finally, we mined publicly available transcriptome data (12) to assess whether any of the genes lying in proximity of association signals showed differential expression in psoriatic skin. The only gene that appeared to be up-regulated was *MICB*. Although the change in gene expression was modest (1.68-fold increase on average), it was consistent and statistically very significant ($P = 8.8 \times 10^{-9}$, using a paired *t*-test). An up-regulation of *MICB* in lesional skin is also apparent in other psoriasis transcriptome datasets, suggesting that the trend we observed is likely to be genuine (19).

**Table 2.** Independent association signals from analysis prioritizing classical alleles

| Classical allele | MAF | Conditional *P*-value | Odds ratio and 95% CI |
| --- | --- | --- | --- |
| HLAC_0602 | 0.15 | $2.20 \times 10^{-236}$ | 4.37 (3.98–4.79) |
| HLAB_3901 | 0.01 | $1.20 \times 10^{-9}$ | 2.00 (1.51–2.62) |
| HLAB_1518 | 0.004 | $1.50 \times 10^{-7}$ | 2.13 (1.3–3.49) |
| HLAB_5701 | 0.08 | $4.09 \times 10^{-7}$ | 4.64 (4.11–5.25) |
| HLAB_3801 | 0.01 | $6.78 \times 10^{-7}$ | 1.73 (1.25–2.40) |
| *rs2517670* | 0.34 | $7.67 \times 10^{-12}$ | 1.51 (1.40–1.64) |
| *6-31367052* | 0.07 | $2.86 \times 10^{-8}$ | 1.11 (0.97–1.27) |

MAF, minor allele frequency.
Top section: independent association signals from stepwise conditional analysis of imputed classical alleles only. Bottom section (italics): further independent association signals from stepwise conditional analysis using genotypes and imputed SNPs and conditioning on all classical alleles nominated in the top section of the table.

None of the novel disease-associated SNPs uncovered here or their proxies appears to be eQTLs according to a list defined in psoriatic skin, uninvolved skin and normal skin (20).

## DISCUSSION

Linkage and association studies have long demonstrated that the MHC region harbours the major genetic determinant for psoriasis susceptibility. Although the association with *HLA-Cw*0602* is well substantiated, the existence of other MHC loci has been investigated in less detail.

Here, we exploited the statistical power of extended case–control data sets to robustly identify further MHC susceptibility determinants. We identify five independent hits in the class I region. rs9260313 (29 917 kb distal to *HLA-A*) is the most telomeric but is within 25 kb of rs380924 (29 940 kb proximal to *HCG9*)—both lie within the *HLA-G* transcript. Over 1 Mb away at intervals of ∼120 and 100 kb, respectively, lie *HLA-C* (31 237–31 240 kb), rs66609536 (31 362 kb proximal to *MICA*) and rs2507971 (31 461 kb proximal to *MICB*). *HLA-Cw*0602* is the well-known risk allele for psoriasis and the signal at rs66609536 was previously described by Feng *et al.* (17).

Our study was not able to replicate a signal detected within the *C6orf10* transcript (16). However, and of note, the *C6orf10* association was not formally validated in the Feng *et al.* study, as the SNPs showing association in the discovery and replication data sets were not genetically correlated, suggesting they may not be representing the same causal variant. The three novel association signals emerging from our study lie in close proximity to *MICB* (rs2507971), *HLA-A* (rs9260313) and the *HCG9* non-coding RNA gene (rs380924). None of the replicated SNPs are in high LD with common class I classical MHC alleles ($r^2 < 0.27$ in controls and $r^2 < 0.33$ in the whole sample).

Recently, there has been renewed interest in the exclusive use of imputed classical alleles in association studies, both in psoriasis and in other diseases. Results from our classical allele-only conditional analyses of the MHC are broadly consistent with others (21). The implicit assumption with these analyses is that classical alleles are a priori more likely to

be truly causal than non-coding SNPs. It is also possible that a SNP in complete LD with a causal classical allele could, by chance, end up with a higher association signal than the classical allele itself, particularly when using data where the number of SNPs greatly outnumber the imputed classical alleles. Although we agree these arguments hold some weight, we note that even when we conditioned on the set of all imputed classical alleles with independent signals in our data, we still found additional SNP signals requiring explanation. Thus, we consider the question of the relative importance of classical alleles versus other genetic mechanisms (such as the regulation of gene expression) in the MHC-related pathobiology of psoriasis to be an unresolved issue. Fine mapping of the MHC locus has also been carried out in ulcerative colitis and rheumatoid arthritis but neither disease shows association with the loci we highlight here (22,23).

The identification of multiple psoriasis risk signals within the class I MHC region is noteworthy, especially as the critical SNPs appear to cluster in proximity to functionally related genes. *HLA-A*, *HLA-C* and *HLA-G* all encode cell-surface proteins that play a fundamental role in the presentation of intracellular antigens to cytotoxic T-lymphocytes. Conversely, *MICA* and *MICB* encode ligands for the natural killer cell NKG2D receptor, which is essential to the immune surveillance of epithelial tissues (24).

The large number of genetic markers and individuals in this study provide excellent tools for the fine mapping of independent signals in the region. Thus, further refinement and functional characterization of these association signals hold the promise to generate significant mechanistic insights into the dysregulation of immune responses in psoriasis.

## MATERIALS AND METHODS

### Discovery samples and genotypes

The 2178 cases were typed on the Illumina Human660W-Quad GWAS platform and the 5175 controls on the Illumina custom Human1.2M-Duo. Further details of the sample, genotyping and quality control are described in the original paper. Analysis was performed on all SNPs passing quality control within the extended MHC region.

### Replication samples and genotypes

For replication, we used the Genetic Analysis of Psoriasis Consortium (GAPC) data set, with additional controls from other collaborators. Subjects were of self-declared European Caucasian ancestry. Informed consent was obtained in adherence to the Declaration of Helsinki principles. A breakdown of sample origins is provided in Supplementary Material, Table S2. All cases were diagnosed by an experienced dermatologist and documented to have clinical features of chronic psoriasis. DNA was isolated from blood or lymphoblastoid cell lines established from B-cells, using standard methods. Samples were genotyped at the Sanger Institute through the Wellcome Trust Case Control Consortium 2, using the Immunochip, which is a custom Illumina Infinium high-density array consisting of 196 524 variants compiled largely from associated regions identified in previous GWAS on 12

**Table 3.** Independent association signals in the replication arm from conditional regression

| Discovery arm SNP | I-chip proxy | MAF | LD# | Conditional $P$-value | Odds ratio (95% CI) | Full-model $P$-value | Full-model odds ratio (95% CI) |
|---|---|---|---|---|---|---|---|
| *HLA-C_0602* | rs12199223 | 0.13 | 0.99 (0.99) | $4.72 \times 10^{-213}$ | 4.3 (3.92–4.71) | $1.56 \times 10^{-234}$ | 1.29 (1.28–1.30) |
| rs2507971 | rs3130922 | 0.31 | 0.77 (0.76) | $1.08 \times 10^{-2}$ | 1.10 (1.19–1.02) | $2.83 \times 10^{-1}$ | 1.01 (1.00–1.01) |
| rs9260313 | rs7745413 | 0.18 | 0.58 (0.59) | $8.02 \times 10^{-4}$ | 0.85 (0.78–0.95) | $3.42 \times 10^{-5}$ | 0.97 (0.97–0.98) |
| rs66609536 | rs13437088 | 0.27 | 0.98 (0.98) | $1.70 \times 10^{-13}$ | 1.32 (1.23–1.44) | $1.33 \times 10^{-14}$ | 1.04 (1.04–1.05) |
| rs380924 | rs2735079 | 0.38 | 0.98 (0.98) | $5.27 \times 10^{-4}$ | 1.14 (1.22–1.05) | $4.42 \times 10^{-4}$ | 1.02 (1.01–1.02) |

MAF, minor allele frequency; LD#, linkage disequilibrium $r^2$ values, calculated from the controls in the discovery arm.
SNPs were added to the model in the order they were selected in the discovery arm.
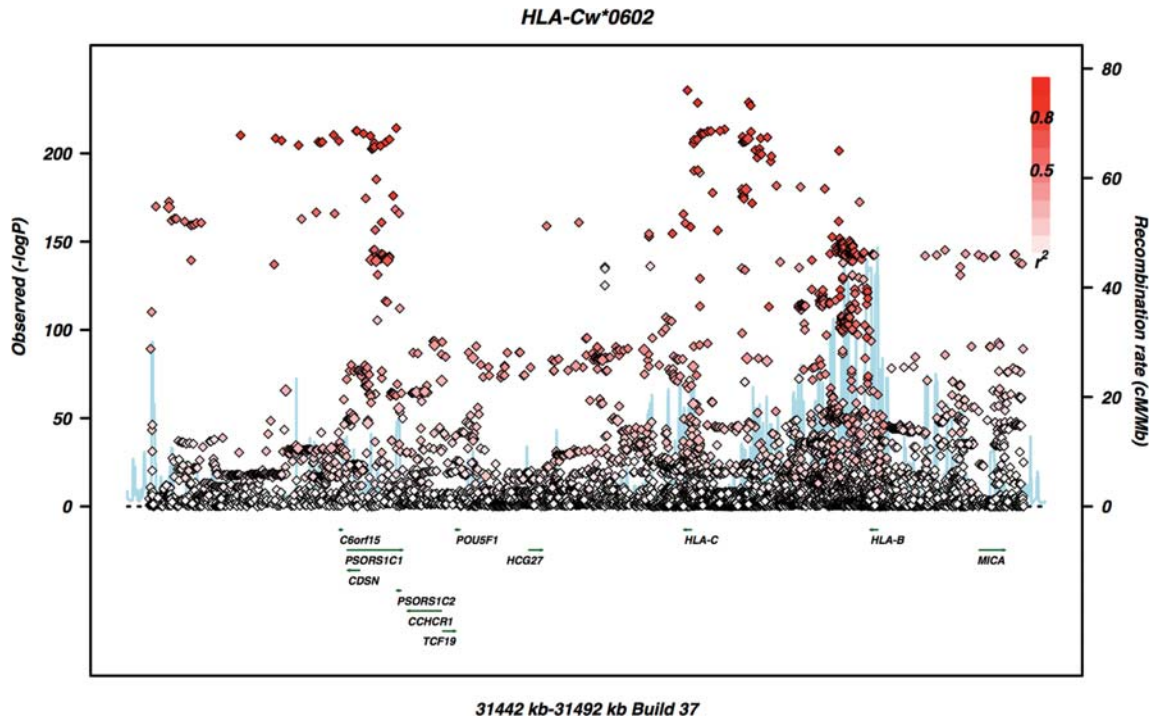


**Figure 2.** Fine-scale conditional regional association plots of the HLA-C signal generated using the SNAP tool (33). Imputed SNPs are shown as diamonds and genotyped SNPs as squares. The base pair position is on the $x$-axis, $-\log 10$ of the $P$-value on the $y$-axis and the recombination fraction on the $z$-axis.

different immune-mediated inflammatory diseases, including psoriasis (25).

SNPs were excluded on the basis of a call rate of <0.95. Samples with call rates <0.98 were excluded as were those that appeared to have outlying ethnicity from a principal component analysis, leaving 12 107 samples (2913 cases and 9194 controls).

## Imputation of SNPs and classical alleles

IMPUTE2 (26) was used to impute additional SNPs in the discovery sample, using the European 1000 Genomes data set (December 2010 data update, from the 20100804 sequence and alignment release) as a reference. Imputed SNPs with an info score <0.5, minor allele frequency <1% and extreme departures from Hardy–Weinberg equilibrium ($P < 10^{-6}$) were excluded. This method produces highly accurate HLA imputations at class I and class II loci with call rates of 95–99% and accuracy between 92 and 98% at the four-digit level. Classical alleles for *HLA-C*, *HLA-B* and *HLA-A* were

imputed using HLA*IMP (27,28). We analysed alleles with >1% frequency. Each allele was recoded as if it were a binary locus, so individuals would have no, one or two copies of the allele in question.

## Identification of independent signals

Analysis was carried out on a data set that included genotyped SNPs, imputed SNPs and imputed classical alleles. Investigation of independent signals was undertaken using stepwise conditional analysis, originally suggested for the analysis of the HLA region and genetic data by Cordell and Clayton (29). Analysis was performed using SNPTEST, a program that performs association tests while accounting for the uncertainty in imputed genotypes (30). First, unconditional association analyses, using additive genetic models, were performed for all SNPs plus imputed classical alleles in the region. After identifying the most significant polymorphism from this scan, we re-ran the analysis, using an additive model for the remaining SNPs but conditioning on both the
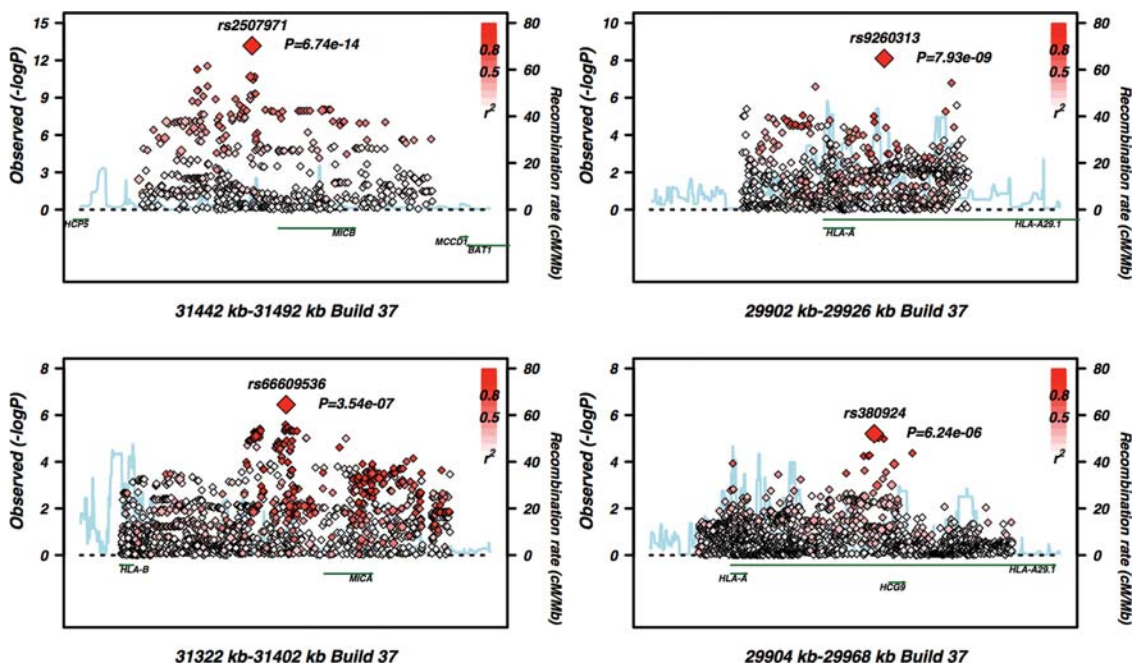
**Figure 3.** Fine-scale conditional regional association plots of the other four independent signals generated using the SNAP tool (33). Imputed SNPs are shown as diamonds and genotyped SNPs as squares. The base pair position is on the *x*-axis, −log 10 of the *P*-value on the *y*-axis and the recombination fraction on the *z*-axis.

additive and dominant effects of the previously identified SNP to ensure that false downstream signals would not be generated by departures from additivity in the signals identified so far. We repeated this procedure conditioning on all previously identified SNPs in each step until no SNP had a conditional *P*-value of $< 1 \times 10^{-6}$. Determination of an appropriate *P*-value is problematic; Bonferroni correction would suggest a *P*-value threshold of $9 \times 10^{-7}$ but would be too extreme given the large number of correlated SNPs. We therefore chose a slightly less stringent *P*-value of $1 \times 10^{-6}$. Having identified five polymorphisms that appeared to have independent signals in the combined analysis of SNPs plus classical alleles, we performed a logistic regression using the R statistical package to fit a model that included all polymorphisms as additive effects and we examined all regression coefficients to ensure they remained significant. All analyses included the first principal component from the genome-wide genetic covariance matrix as a covariate to correct for English/Irish population structure, as described in Strange *et al*. (13).

To further investigate the role of the classical alleles, we re-performed the stepwise procedure using only classical alleles until none remained in the model with a *P*-value of $< 1 \times 10^{-6}$. At this point, we re-introduced the genotyped and imputed SNPs to see whether residual signals remained that could not be explained by the classical alleles already in the model. We also performed stepwise regression on a data set that excluded the imputed classical alleles.

### Replication analysis

We examined conditional association signals by proxy in the Immunochip data set. Tagging $r^2$ values (calculated from controls) are given in Table 3. SNP rs7745413 had a

Hardy–Weinberg equilibrium *P*-value of $3 \times 10^{-8}$ in the controls but was included as such deviations are often seen in the MHC region. We performed conditional analysis to determine both whether the same hits were present and whether they were independent. We entered the SNPs into the model in the order of selection in the discovery data set. We used principal components from the genetic covariance matrix (across all Immunochip SNPs) as covariates to control for the multiple European origins of the samples. The first 10 principal components were used as covariates, as this was deemed to adequately control the inflation in association statistics (lambda = 1.13).

### Hyperlasso

Stepwise conditional analysis is a popular method for model selection because it reflects a straightforward parsimony approach to the selection of new hits. At each step, a specific test is made of the evidence for or against expanding the current set of independent signals. However, since the procedure depends on local optimization at each step, it is possible that the final set of hits arrived at may not be globally optimal. An alternative approach that has been applied to searching for association signals in the MHC region is the Hyperlasso (31,32), a Bayesian regression method that starts with a normal exponential gamma prior for the association coefficient (log odds ratio) assigned to each SNP. We applied the Hyperlasso to our data, but found that the method tended to pick two SNPs in LD in place of a single SNP to represent an association signal. It is possible that further tuning of the parameters would have improved its performance but as the results from the stepwise regression replicated we did not pursue this method any further.

### Interaction

We checked for interaction between each independent locus and SNPs throughout the rest of the genome using logistic regression, with the phenotype as the dependent variable. For each SNP in the GWAS, we investigated five models, each model included the main effect of the GWAS SNP, the main effects of the five independent loci, the principal component to control for ancestry and an interaction term between the GWAS SNP and one of the five independent loci. We coded the *HLA-Cw\*0602* as dominant according to its suggested effect. We followed up results with *P*-values $<1 \times 10^{-8}$, but none replicated in the Immunochip data.

### SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

### REFERENCES

1. Nestle, F.O., Kaplan, D.H. and Barker, J. (2009) Psoriasis. *New Engl. J. Med.*, **361**, 496–509.
2. Capon, F., Burden, A.D., Trembath, R.C. and Barker, J.N. (2012) Psoriasis and other complex trait dermatoses: from loci to functional pathways. *J. Invest. Dermatol.*, **132**, 915–922.
3. Trembath, R.C., Clough, R.L., Rosbotham, J.L., Jones, A.B., Camp, R.D., Frodsham, A., Browne, J., Barber, R., Terwilliger, J., Lathrop, G.M. *et al.* (1997) Identification of a major susceptibility locus on chromosome 6p and evidence for further disease loci revealed by a two stage genome-wide search in psoriasis. *Hum. Mol. Genet.*, **6**, 813–820.
4. Fan, X., Yang, S., Huang, W., Wang, Z.M., Sun, L.D., Liang, Y.H., Gao, M., Ren, Y.Q., Zhang, K.Y., Du, W.H. *et al.* (2008) Fine mapping of the psoriasis susceptibility locus PSORS1 supports HLA-C as the susceptibility gene in the Han Chinese population. *PLoS Genet.*, **4**, e1000038.
5. Helms, C., Saccone, N.L., Cao, L., Daw, J.A., Cao, K., Hsu, T.M., Taillon-Miller, P., Duan, S., Gordon, D., Pierce, B. *et al.* (2005) Localization of PSORS1 to a haplotype block harboring HLA-C and distinct from corneodesmosin and HCR. *Hum. Genet.*, **118**, 466–476.
6. Nair, R.P., Stuart, P.E., Nistor, I., Hiremagalore, R., Chia, N.V., Jenisch, S., Weichenthal, M., Abecasis, G.R., Lim, H.W., Christophers, E. *et al.* (2006) Sequence and haplotype analysis supports HLA-C as the psoriasis susceptibility 1 gene. *Am. J. Hum. Genet.*, **78**, 827–851.
7. Capon, F., Bijlmakers, M.J., Wolf, N., Quaranta, M., Huffmeier, U., Allen, M., Timms, K., Abkevich, V., Gutin, A., Smith, R. *et al.* (2008) Identification of ZNF313/RNF114 as a novel psoriasis susceptibility gene. *Hum. Mol. Genet.*, **17**, 1938–1945.
8. Cargill, M., Schrodi, S.J., Chang, M., Garcia, V.E., Brandon, R., Callis, K.P., Matsunami, N., Ardlie, K.G., Civello, D., Catanese, J.J. *et al.* (2007) A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. *Am. J. Hum. Genet.*, **80**, 273–290.
9. Ellinghaus, E., Ellinghaus, D., Stuart, P.E., Nair, R.P., Debrus, S., Raelson, J.V., Belouchi, M., Fournier, H., Reinhard, C., Ding, J. *et al.* (2010) Genome-wide association study identifies a psoriasis susceptibility locus at TRAF3IP2. *Nat. Genet.*, **42**, 991–995.
10. Huffmeier, U., Uebe, S., Ekici, A.B., Bowes, J., Giardina, E., Korendowych, E., Juneblad, K., Apel, M., McManus, R., Ho, P. *et al.* (2010) Common variants at TRAF3IP2 are associated with susceptibility to psoriatic arthritis and psoriasis. *Nat. Genet.*, **42**, 996–999.
11. Liu, Y., Helms, C., Liao, W., Zaba, L.C., Duan, S., Gardner, J., Wise, C., Miner, A., Malloy, M.J., Pullinger, C.R. *et al.* (2008) A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. *PLoS Genet.*, **4**, e1000041.
12. Nair, R.P., Duffin, K.C., Helms, C., Ding, J., Stuart, P.E., Goldgar, D., Gudjonsson, J.E., Li, Y., Tejasvi, T., Feng, B.J. *et al.* (2009) Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. *Nat. Genet.*, **41**, 199–204.
13. Strange, A., Capon, F., Spencer, C.C., Knight, J., Weale, M.E., Allen, M.H., Barton, A., Band, G., Bellenguez, C., Bergboer, J.G. *et al.* (2010) A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat. Genet.*, **42**, 985–990.
14. Stuart, P.E., Nair, R.P., Ellinghaus, E., Ding, J., Tejasvi, T., Gudjonsson, J.E., Li, Y., Weidinger, S., Eberlein, B., Gieger, C. *et al.* (2010) Genome-wide association analysis identifies three psoriasis susceptibility loci. *Nat. Genet.*, **42**, 1000–1004.
15. Sun, L.D., Cheng, H., Wang, Z.X., Zhang, A.P., Wang, P.G., Xu, J.H., Zhu, Q.X., Zhou, H.S., Ellinghaus, E., Zhang, F.R. *et al.* (2010) Association analyses identify six new psoriasis susceptibility loci in the Chinese population. *Nat. Genet.*, **42**, 1005–1009.
16. Zhang, X.J., Huang, W., Yang, S., Sun, L.D., Zhang, F.Y., Zhu, Q.X., Zhang, F.R., Zhang, C., Du, W.H., Pu, X.M. *et al.* (2009) Psoriasis genome-wide association study identifies susceptibility variants within LCE gene cluster at 1q21. *Nat. Genet.*, **41**, 205–210.
17. Feng, B.J., Sun, L.D., Soltani-Arabshahi, R., Bowcock, A.M., Nair, R.P., Stuart, P., Elder, J.T., Schrodi, S.J., Begovich, A.B., Abecasis, G.R. *et al.* (2009) Multiple Loci within the major histocompatibility complex confer risk of psoriasis. *PLoS Genet.*, **5**, e1000606.
18. Horton, R., Wilming, L., Rand, V., Lovering, R.C., Bruford, E.A., Khodiyar, V.K., Lush, M.J., Povey, S., Talbot, C.C. Jr, Wright, M.W. *et al.* (2004) Gene map of the extended human MHC. *Nat. Rev. Genet.*, **5**, 889–899.
19. Yao, Y., Richman, L., Morehouse, C., de los Reyes, M., Higgs, B.W., Boutrin, A., White, B., Coyle, A., Krueger, J., Kiener, P.A. *et al.* (2008) Type I interferon: potential therapeutic target for psoriasis? *PLoS ONE*, **3**, e2737.
20. Ding, J., Gudjonsson, J.E., Liang, L., Stuart, P.E., Li, Y., Chen, W., Weichenthal, M., Ellinghaus, E., Franke, A., Cookson, W. *et al.* (2010) Gene expression in skin and lymphoblastoid cells: refined statistical method reveals extensive overlap in *cis*-eQTL signals. *Am. J. Hum. Genet.*, **87**, 779–789.
21. Chen, H., Hayashi, G., Lai, O.Y., Dilthey, A., Kuebler, P.J., Wong, T.V., Martin, M.P., Fernandez Vina, M.A., McVean, G., Wabl, M. *et al.* (2012) Psoriasis patients are enriched for genetic variants that protect against HIV-1 disease. *PLoS Genet.*, **8**, e1002514.
22. Achkar, J.P., Klei, L., de Bakker, P.I., Bellone, G., Rebert, N., Scott, R., Lu, Y., Regueiro, M., Brzezinski, A., Kamboh, M.I. *et al.* (2012) Amino acid position 11 of HLA-DRbeta1 is a major determinant of chromosome 6p association with ulcerative colitis. *Genes. Immun.*, **13**, 245–252.
23. Raychaudhuri, S., Sandor, C., Stahl, E.A., Freudenberg, J., Lee, H.S., Jia, X., Alfredsson, L., Padyukov, L., Klareskog, L., Worthington, J. *et al.* (2012) Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat. Genet.*, **44**, 291–296.
24. Hayday, A.C. (2009) Gammadelta T cells and the lymphoid stress-surveillance response. *Immunity*, **31**, 184–196.
25. Cortes, A. and Brown, M.A. (2011) Promise and pitfalls of the Immunochip. *Arthritis Res. Ther.*, **13**, 101.

26. Howie, B.N., Donnelly, P. and Marchini, J. (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.*, **5**, e1000529.

27. Dilthey, A.T., Moutsianas, L., Leslie, S. and McVean, G. (2011) HLA*IMP—an integrated framework for imputing classical HLA alleles from SNP genotypes. *Bioinformatics*, **27**, 968–972.

28. Leslie, S., Donnelly, P. and McVean, G. (2008) A statistical method for predicting classical HLA alleles from SNP data. *Am. J. Hum. Genet.*, **82**, 48–56.

29. Cordell, H.J. and Clayton, D.G. (2002) A unified stepwise regression procedure for evaluating the relative effects of polymorphisms within a gene using case/control or family data: application to HLA in type 1 diabetes. *Am. J. Hum. Genet.*, **70**, 124–141.

30. Marchini, J., Howie, B., Myers, S., McVean, G. and Donnelly, P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.*, **39**, 906–913.

31. Hoggart, C.J., Whittaker, J.C., De Iorio, M. and Balding, D.J. (2008) Simultaneous analysis of all SNPs in genome-wide and re-sequencing association studies. *PLoS Genet.*, **4**, e1000130.

32. Vignal, C.M., Bansal, A.T. and Balding, D.J. (2011) Using penalised logistic regression to fine map HLA variants for rheumatoid arthritis. *Ann. Hum. Genet.*, **75**, 655–664.

33. Johnson, A.D., Handsaker, R.E., Pulit, S.L., Nizzari, M.M., O'Donnell, C.J. and de Bakker, P.I. (2008) SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*, **24**, 2938–2939.