A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci

Ying Liu
Washington University School of Medicine in St. Louis

Cynthia Helms
Washington University School of Medicine in St. Louis

Wilson Liao
University of California - San Francisco

Lisa C. Zaba
Rockefeller University

Shenghui Duan
Washington University School of Medicine in St. Louis

See next page for additional authors

Follow this and additional works at: http://digitalcommons.wustl.edu/open_access_pubs

Part of the Medicine and Health Sciences Commons

Recommended Citation
Liu, Ying; Helms, Cynthia; Liao, Wilson; Zaba, Lisa C.; Duan, Shenghui; Gardner, Jennifer; Wise, Carol; Miner, Andrew; Malloy, M. J.; Pullinger, Clive R.; Kane, John P.; Saccone, Scott; Worthington, Jane; Bruce, Ian; Kwok, Pui-Yan; Menter, Alan; Krueger, James; Barton, Anne; Saccone, Nancy L.; and Bowcock, Anne M., "A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci." PLoS Genetics.4,4. e1000041. (2008).
http://digitalcommons.wustl.edu/open_access_pubs/927

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.
A Genome-Wide Association Study of Psoriasis and Psoriatic Arthritis Identifies New Disease Loci

Ying Liu1, Cynthia Helms1, Wilson Liao2, Lisa C. Zaba3, Shenghui Duan1, Jennifer Gardner1, Carol Wise4, Andrew Miner5, M. J. Malloy5, Clive R. Pullinger2, John P. Kane2, Scott Saccone5, Jane Worthington6, Ian Bruce6, Pui–Yan Kwok2, Alan Menter7, James Krueger3, Anne Barton6, Nancy L. Saccone5, Anne M. Bowcock1*

1 Division of Human Genetics, Department of Genetics, Washington University School of Medicine, St. Louis, Missouri, United States of America, 2 Cardiovascular Research Institute and Center for Human Genetics, University of California San Francisco, San Francisco, California, United States of America, 3 Laboratory for Investigative Dermatology, The Rockefeller University, New York, New York, United States of America, 4 Seay Center for Musculoskeletal Research, Texas Scottish Rite Hospital for Children, Dallas, Texas, United States of America, 5 Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, United States of America, 6 Arc Epidemiology Research Unit, University of Manchester, Manchester, United Kingdom, 7 Department of Internal Medicine, Division of Dermatology, Baylor University Medical Center, Dallas, Texas, United States of America

Abstract

A genome-wide association study was performed to identify genetic factors involved in susceptibility to psoriasis (PS) and psoriatic arthritis (PSA), inflammatory diseases of the skin and joints in humans. 223 PS cases (including 91 with PSA) were genotyped with 311,398 single nucleotide polymorphisms (SNPs), and results were compared with those from 519 Northern European controls. Replications were performed with an independent cohort of 577 PS cases and 737 controls from the U.S., and 576 PSA patients and 480 controls from the U.K. Strongest associations were with the class I region of the major histocompatibility complex (MHC). The most highly associated SNP was rs10484554, which lies 34.7 kb upstream from HLA-C (P = 7.8 x 10^-11, GWA scan; P = 1.8 x 10^-30, replication; P = 1.8 x 10^-39, combined; U.K. PSA: P = 6.9 x 10^-11). However, rs2395029 encoding the G2V polymorphism within the class I gene HCP5 (combined P = 2.13 x 10^-26 in U.S. cases) yielded the highest ORs with both PS and PSA (4.1 and 3.2 respectively). This variant is associated with low viral set point following HIV infection and its effect is independent of rs10484554. We replicated the previously reported association with interleukin 23 receptor and interleukin 12B polymorphisms in PS and PSA cohorts (IL23R: rs11209026, U.S. PS, P = 1.4 x 10^-5; U.K. PSA: P = 8.0 x 10^-5; IL12B:rs6887695, U.S. PS, P = 5 x 10^-5 and U.K. PSA, P = 1.3 x 10^-5) and detected an independent association in the IL23R region with a SNP 4 kb upstream from IL12RB2 (P = 0.001). Novel associations replicated in the U.S. PS cohort included the region harboring lipoma HMGI fusion partner (LHFP) and conserved oligomeric golgi complex component 6 (COG6) genes on chromosome 13q13 (combined P = 2 x 10^-5 for rs7993214; OR = 0.71), the late cornified envelope gene cluster (LCE) from the Epidermal Differentiation Complex (PSORS4) (combined P = 6.2 x 10^-5 for rs6701216; OR 1.45) and a region of LD at 15q21 (combined P = 2.9 x 10^-5 for rs3803369; OR = 1.43). This region is of interest because it harbors ubiquitin-specific protease-8 whose processed pseudogene lies upstream from HLA-C. This region of 15q21 also harbors the gene for SPPL2A (signal peptide peptidase like 2a) which activates tumor necrosis factor alpha by cleavage, triggering the expression of IL12 in human dendritic cells. We also identified a novel PSA (and potentially PS) locus on chromosome 4q27. This region harbors the interleukin 2 (IL2) and interleukin 21 (IL21) genes and was recently shown to be associated with four autoimmune diseases (Celiac disease, Type 1 diabetes, Grave’s disease and Rheumatoid Arthritis).


Editor: Suzanne M. Leal, Baylor College of Medicine, United States of America

Received September 27, 2007; Accepted February 28, 2008; Published April 4, 2008

Copyright: © 2008 Liu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported in part by NIH grant R01 AR050266 (AMB), K01DA015129 (NLS) and TSRHC grant 0305756 (CAW).

Competing Interests: Patent application on findings in process.

* E-mail: bowcock@genetics.wustl.edu

Introduction

Psoriasis (PS) is a chronic inflammatory disease of the skin affecting 2–3% of the population [1]. Approximately 25% of patients also develop psoriatic arthritis (PSA), a common, debilitating auto-immune disease belonging to the family of spondyloarthritides [2,3]. The recurrence risk (λs) of PSA is high, and estimates of 27–47 have been proposed [4,5]. This is much higher than the estimated λs of PS which is estimated to be between 4 and 11 [6].

PS and PSA are interrelated disorders, and the prevalence of PS is 19 times higher among first degree relatives of probands with PSA compared with the general population [7]. The pathogenesis of PS and PSA is complex, involving both genetic and environmental risk factors. Strong association of PS with the MHC class I region (PSORS1 or psoriasis susceptibility locus 1) was demonstrated in the 1970s [8] and has been confirmed in numerous subsequent studies [9–11]. However, the genetics of PSA is not as clear-cut and association with alleles of the HLA class I region is not reported to be as strong with PSA as with PS.
Psoriasis (PS) and psoriatic arthritis (PSA) are common inflammatory diseases of humans affecting the skin and joints. Approximately 2% of Europeans are affected with PS, and ~10–30% of patients develop PSA. Genetic variation in the MHC (multiple histocompatibility locus antigen cluster) increases risk of developing PS. However, only ~10% of individuals with this risk factor develop PS, indicating that other genetic effects and environmental triggers are important. Recent approaches using a case/control approach and genome-wide association studies with DNA markers known as SNPs (single nucleotide polymorphisms) have been fruitful in identifying genetic factors for common diseases. This study describes the first large scale genome-wide scan for additional PS and PSA susceptibility genes using 233 cases and 519 controls. It revealed that the MHC is truly the most important risk factor for PS and that it plays a very major role in PSA, confirmed recently identified associations with interleukin 23 receptor and interleukin 12B in both PS and PSA, and identified new associations. These include a region on chromosome 4q27 that contains genes for interleukin 2 and interleukin 21 that has been recently implicated in other autoimmune diseases, and seven additional regions that include chromosome 13q13 and 15q21.

Author Summary

Psoriasis (PS) and psoriatic arthritis (PSA) are common inflammatory diseases of humans affecting the skin and joints. Approximately 2% of Europeans are affected with PS, and ~10–30% of patients develop PSA. Genetic variation in the MHC (multiple histocompatibility locus antigen cluster) increases risk of developing PS. However, only ~10% of individuals with this risk factor develop PS, indicating that other genetic effects and environmental triggers are important. Recent approaches using a case/control approach and genome-wide association studies with DNA markers known as SNPs (single nucleotide polymorphisms) have been fruitful in identifying genetic factors for common diseases. This study describes the first large scale genome-wide scan for additional PS and PSA susceptibility genes using 233 cases and 519 controls. It revealed that the MHC is truly the most important risk factor for PS and that it plays a very major role in PSA, confirmed recently identified associations with interleukin 23 receptor and interleukin 12B in both PS and PSA, and identified new associations. These include a region on chromosome 4q27 that contains genes for interleukin 2 and interleukin 21 that has been recently implicated in other autoimmune diseases, and seven additional regions that include chromosome 13q13 and 15q21.

Results/Discussion

For our “discovery” phase, 223 PS cases (132 cases with PS without arthritis and 91 PS cases with arthritis (PSA)) were typed on the Illumina HumanHap300 arrays. We compared case data to publicly available genotype data of 519 European controls from the New York Cancer Project[18] collected with the same platform. The number of cases used for this scan is smaller than that used in many recently described genome-wide association scans. However, the 91 cases of PSA had at least one first degree relative with PS and were expected to be enriched for genetic factors. Power calculations based on 223 controls and 519 controls indicated that using a threshold of P<5×10^-5, we had 70% power of detecting a locus with a genotype relative risk (GRR) of 2.0, and over 99% power to detect a locus with a GRR of 3.0 such as the MHC (see below). However, many replicated associations have small GRRs [19] and we had only 10% power of detecting a locus with a GRR of 1.5.

Following the genotyping of samples, stringent quality control measures were implemented. We required that all samples used for the discovery phase passed a 93% genotyping call rate threshold, and that all SNPs passed a 95% call rate threshold. Justification for this threshold is based on the evaluation of empirical distributions (Figures S1 and S2). With sample call rates ≤93%, there was an elevation in observed sample heterozygosity, i.e. deviation from Hardy-Weinberg equilibrium, suggesting possible genotyping errors (e.g., sample contamination or allele drop-out). Likewise, there was a significant discrepancy of missingness between case and control groups when the SNP success rate was <95%.

For the discovery phase, a total of 311,398 SNPs were pruned to 305,983 SNPs after filtering for low call rate, minor allele frequency <0.01 and deviation from Hardy-Weinberg equilibrium (P<0.001). Quality control also led to the removal of 29 samples, leaving 218 cases for further analysis (Table S1). The average genotyping rate in the remaining individuals was 0.995.

To investigate other biases[20] that could be introduced with shared controls, we assessed the median distribution of test statistics using the genomic-control factor λGC[21]. With a set of 463 ancestry informative SNPs (AIMs), λGC = 1 indicating no inflation. We also performed analysis on the same set of AIM SNPs with STRUCTURE software[22]. Under the assumption of two population clusters, there was no association between the most likely inferred cluster and case/control status and the average allele frequency difference between clusters was less than 2.5%. These results showed that population substructure is unlikely to be confounding our results. However, analysis of all markers used in the discovery study yielded λGC = 1.101 before correction (Q-Q plot shown in Figure S3). A similar value was obtained with EIGENSTRAT[23] where λ = 1.107. Examining stratified subsets of cases (PS without arthritis or PSA) with all markers also yielded similar λ values (PS without arthritis: λGC = 1.07; PSA: λGC = 1.05). Following adjustment of P values with the genomic control method, λ = 1. The discovery P values adjusted by the genomic control method as implemented in PLINK[24] are presented in the tables and figures.

To detect associations, we first performed a preliminary analysis with a Cochran-Armitage trend test[25]. Figure 1 illustrates negative logarithm of the P values obtained across the genome, considering all cases and all controls (Figure 1A) and considering only the 91 PSA cases and all controls (Figure 1B). Results were then rank ordered on the basis of P values. 84 SNPs in 35 genomic regions were associated with PS with arthritis or PSA with all markers also yielded similar λ values (PS without arthritis: λGC = 1.07; PSA: λGC = 1.05). Following adjustment of P values with the genomic control method, λ = 1. The discovery P values adjusted by the genomic control method as implemented in PLINK[24] are presented in the tables and figures.

To detect associations, we first performed a preliminary analysis with a Cochran-Armitage trend test[25]. Figure 1 illustrates negative logarithm of the P values obtained across the genome, considering all cases and all controls (Figure 1A) and considering only the 91 PSA cases and all controls (Figure 1B). Results were then rank ordered on the basis of P values. 84 SNPs in 35 genomic regions were associated with PS with arthritis or PSA with all markers also yielded similar λ values (PS without arthritis: λGC = 1.07; PSA: λGC = 1.05). Following adjustment of P values with the genomic control method, λ = 1. The discovery P values adjusted by the genomic control method as implemented in PLINK[24] are presented in the tables and figures.

To detect associations, we first performed a preliminary analysis with a Cochran-Armitage trend test[25]. Figure 1 illustrates negative logarithm of the P values obtained across the genome, considering all cases and all controls (Figure 1A) and considering only the 91 PSA cases and all controls (Figure 1B). Results were then rank ordered on the basis of P values. 84 SNPs in 35 genomic regions were associated with PS with arthritis or PSA with all markers also yielded similar λ values (PS without arthritis: λGC = 1.07; PSA: λGC = 1.05). Following adjustment of P values with the genomic control method, λ = 1. The discovery P values adjusted by the genomic control method as implemented in PLINK[24] are presented in the tables and figures.
of variants upon PSA susceptibility specifically, 576 PSA cases from the UK and 480 controls from the UK were also employed. An alternative genotyping technology (iPlex; Sequenom) was used for the replication phase. The platforms used for the discovery and replication phases gave very similar results: Concordance rates on the basis of 116 samples and 301 SNPs typed with both platforms was 98.74%.

Our 100 top ranked SNPs with any cohort (PS, PS without arthritis, or PSA) are listed in Table S2, to facilitate future attempts to replicate our findings. A total of 289 SNPs, including SNPs from the MHC, and two previously reported associated SNPs within IL12B and IL23R[16] were genotyped in the replication analysis.

MHC

The MHC, and in particular, the HLA class I region, is the only region that has been shown to be consistently associated with PS. The first nine top-ranking SNPs were from the MHC and seven were significant, even when adjusted with the Bonferroni correction for multiple tests (Table S2). Overall, 32 SNPs from the MHC had adjusted P values <5×10^{-5} (Figure 2, Table S2). The most significant association was with rs10484554 (adjusted P = 7.8×10^{-11}, GWA scan; P = 5.61×10^{-20}, replication; P = 9.77×10^{-30}, combined) (Figure 2, Table 1). This SNP lies 34.7 kb upstream from the transcriptional start site of HLA-C. Strongest association with this region is consistent with previous results from our group and others [9–11]. The rs10484554*T allele had frequencies of 0.325 in U.S. cases and 0.15 in U.S. controls (OR: 2.8 (95% CI: 2.4–3.3). To determine the relationship of this allele with the classical HLA-C allele strongly associated with psoriasis (HLA-Cw*0602), we investigated the transmission of this allele with classical HLA-C alleles in 250 nuclear families with psoriasis that we have reported elsewhere [26]. The rs10484554*T allele was found on nearly all haplotypes with HLA-Cw*0602 or HLA-Cw*1203 alleles and was also strongly correlated with the previously described highly associated PSORS1 SNP n9*G[26,27] (rs10456057*G) allele. We have previously shown that SNPs upstream from HLA-C are more strongly correlated with PS than HLA-Cw*0602 is, and that these risk alleles are also correlated with HLA-Cw*1203 [26]. Hence, rs10484554*T may be a good proxy for the PSORS1 variant.

In the case of the U.K. PSA replication samples, rs10484554 was again highly significant (P = 6.86×10^{-11}) (Table 2), although the frequency of the rs10484554*T allele exhibited population differences when frequencies in the U.K. and U.S. were compared. In the U.K., the rs10484554*T allele was found at a lower frequency in cases and controls (0.19 and 0.07 respectively; OR: 2.4 (95% CI: 1.8–3.1)).

A second SNP from the HLA class I region lying between MICA and MICB (rs2395029) was highly associated with PS and PSA. This SNP results in the G2V polymorphism of the class I gene HCP5 (HLA complex P5) which encodes an endogenous retroviral element. For this SNP, PS was associated with a combined P = 2.13×10^{-20} in the U.S. cohort and 1.86×10^{-10} in the U.K. PSA cohort (Table 1, Table 2). The OR of the rs2395029*C allele with both PS and PSA was higher than with any other SNP tested (4.1 and 3.2 with PS and PSA respectively). This allele was found at a frequency of ~0.12 in cases and 0.04 in controls and did not exhibit the population frequency differences of rs10484554. The LD relationship between rs2395029 and rs10484554 is not strong (r^2 = 0.33 in European CEPIH HapMap samples and r^2 = 0.23 in our U.S. case/control cohort). Conditioning upon rs10484554, the P value for rs2395029 was still significant (P = 7×10^{-10}), hence effects from this SNP are likely to be independent.

HCP5 is expressed primarily in cells of the immune system such as spleen, blood and thymus [http://smd-www.stanford.edu/], consistent with a potential role in autoimmunity. This allele was
recently shown to explain 9.6% of the total variation in viral set point following HIV-1 infection[28]. This is of interest, since psoriasis can be triggered by infection with HIV and other viruses. Hence, it is possible that HCP5-C carriers mount a strong immune reaction to viral infection, but that in genetically susceptible individuals, this reaction leads to excessive inflammation in skin and joints. Overall, our observations indicate that MHC class I region SNPs are more highly associated with both PS and PSA than any other SNPs.

IL23R associations
A recent global association scan using a set of pooled PS samples and controls against a set of 25,215 gene-centric SNPs confirmed a previously reported association with IL12B (rs3212227 in its 3' UTR) [15] and identified a second region of association 60 kb upstream from its mRNA start site (rs6887695) [16]. An analysis of additional genes encoding components of the IL12B pathway led to the identification of associations with IL23R (R381Q: rs11209026 and L310P: rs7530511) [16]. These SNPs were
Table 1. Summary of association with previously reported PS susceptibility loci (MHC, IL23R and IL12B) in U.S. PS cohort (810 cases, 1256 controls).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| HapMap CEPH European samples; 0.009 in cases; 0.026 in controls. Conditioning upon rs11209026, the P value for rs11209026 was 0.013. Hence, effects from this SNP may be independent of rs11209026 and its association with PS should be investigated in other cohorts.

SNP rs112131065 lies downstream from IL23R (63 kb from rs11209026) and 4.041 kb upstream from the gene for interleukin 12 receptor B2 (IL12RB2) (Figure 3). IL12RB2 is involved in IL12 dependent signaling, is upregulated by gamma interferon in T$_d_1$ cells and plays a role in T$_d_1$ differentiation[29]. Association with a SNP closer to IL12RB2 than IL23R is of interest since animals where IL12RB2 is inactivated develop autoimmune disease[30].

Association with the previously reported IL23R associated SNP rs11209026 in the discovery cohort was not significant (adjusted $P = 0.081$). Genotyping of rs11209026 and rs12131065 in the U.S. replication cohort yielded combined P values of $1.4 \times 10^{-4}$ and 0.0011 respectively (Table 1) consistent with replication of this locus with respect to previous studies. In the case of these two SNPs, the protective T and A alleles were found at frequencies of $0.04$ and $0.2$ in cases versus $0.07$ and $0.24$ of controls respectively. In the U.K. replication PSA cohort, association with rs11209026 was consistent with replication ($P = 8.3 \times 10^{-4}$), with the rs11209026*T allele being found at frequencies of $0.04$ in cases and $0.08$ in controls (Table 2).

IL12B Associations

Although the associated IL12B SNPs rs3212227 and rs6887695 were not interrogated by the Illumina screening panel of SNPs used here, typing of these SNPs in our replication U.S. case/control cohorts yielded P values of $0.021$ and $5 \times 10^{-5}$ and replicated previously reported associations (Table 1). In the U.K. PSA cohort, association with rs6887695 was also consistent with replication ($P = 0.0013$) (OR: 0.69; 95% CI: 0.56–0.85) (Table 2).

Novel Psoriasis Loci

In the discovery cohort, there were four SNPs from 13q13 where $P<5 \times 10^{-5}$ (adjusted $P<2 \times 10^{-7}$). These were: rs1864668, rs4514547, rs4569131 and rs7999214 (Table S2). These SNPs lie within a region on chromosome13q13 that encodes the conserved oligomeric golgi complex component 6 (COG6) gene and a lipoma HMGC fusion partner (LHFP)[31]. Three of the top associated SNPs were tested in the U.S. replication cohort, and all showed evidence of replication at $P<0.05$ (Table 3, Figure 4). Results were most significant with rs7999214 (adjusted $P = 10^{-4}$, GWA scan; $P = 0.0033$, replication; $P = 2 \times 10^{-6}$, combined), Rs3812888, (adjusted $P = 0.0017$, GWA scan; $P = 4 \times 10^{-4}$, replication; $P = 10^{-5}$, combined) was the only SNP where replication results would remain significant following the stringent Bonferroni correction for multiple tests ($P = 0.048$). The OR of the rs3812888*C allele was 1.38 (95% CI: 1.15–1.66). The rs3812888*C allele was found at frequencies of 0.43 in cases and 0.35 in controls.

COG6 is a component of the conserved oligomeric golgi (COG) complex and is involved in intracellular transport and glycoprotein modification[32]. The glycosylation pathways in the golgi
Table 2. Summary of association with previously reported PS susceptibility loci in U.K. PSA cohort (576 cases, 480 controls).

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene/Region</th>
<th>Trend P value</th>
<th>Minor allele</th>
<th>Freq. UK PSA cases</th>
<th>Freq. UK PSA controls</th>
<th>OR UK PSA (95% C.I.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10484554</td>
<td>MHC</td>
<td>6.86 x 10^-11</td>
<td>T</td>
<td>0.19</td>
<td>0.07</td>
<td>2.4 (1.8, 3.1)</td>
</tr>
<tr>
<td>rs2395029</td>
<td>MHC</td>
<td>1.86 x 10^-10</td>
<td>C</td>
<td>0.12</td>
<td>0.04</td>
<td>3.2 (2.2, 4.6)</td>
</tr>
<tr>
<td>rs11209026</td>
<td>IL23R</td>
<td>0.00083</td>
<td>T</td>
<td>0.043</td>
<td>0.079</td>
<td>0.52 (0.35, 0.77)</td>
</tr>
<tr>
<td>rs12131065</td>
<td>IL23R</td>
<td>0.31</td>
<td>A</td>
<td>0.21</td>
<td>0.23</td>
<td>0.89 (0.72, 1.11)</td>
</tr>
<tr>
<td>rs3212217</td>
<td>IL12B</td>
<td>N.D.</td>
<td>G</td>
<td>N.D.</td>
<td>N.D.</td>
<td>N.D.</td>
</tr>
<tr>
<td>rs6887695</td>
<td>IL12B</td>
<td>0.0013</td>
<td>C</td>
<td>0.213</td>
<td>0.28</td>
<td>0.69 (0.56, 0.85)</td>
</tr>
</tbody>
</table>

UK PSA: U.K. Psoriasis cases with arthritis; OR: Odds Ratio; C.I.: confidence interval; N.D.: Not Done.

doi:10.1371/journal.pgen.1000041.t002

Figure 3. Association localization plots for the ILI23R region on chromosome 1. Symbols are the same as those used in Figure 2. SNPs indicated with an asterisk are rs11465804, rs11209026 (R381Q) and rs12131065.
doi:10.1371/journal.pgen.1000041.g003
apparatus must be intact for protein secretion to continue unabated. In *C. elegans*, a COG complex is required to glycosylate an ADAM protease (a disintegrin and metalloprotease)[39]. In humans, variants within some ADAM genes lead to inflammatory diseases. For example, ADAM33 is an asthma susceptibility gene whose catalytic domain undergoes glycosylation[34]. Recent genetic studies suggest that ADAM33 is a psoriasis susceptibility gene as well [35]. Hence, COG6 could be involved in glycosylation of ADAM33 or other ADAM proteases. Staining of skin sections with a COG6 antibody revealed cytoplasmic staining in the epidermis as well as strong T-cell staining (Figure 5).

When the PSA discovery cohort (*n* = 91) was analyzed separately, four SNPs from a region on chromosome 15q21 between ubiquitin specific protease 8 (USP8) and tumor necrosis factor, alpha-induced protein 8-like 3 (TNFAIP8L3) were associated with *P* < 5 × 10⁻⁵ (Table S2). In the case of the most highly associated SNP (rs4775919), adjusted *P* = 6.7 × 10⁻⁶. Following replication genotyping, this and rs3803369 were associated with PS with *P* values consistent with replication (for Table 3, Figure S4). The SNP lies within the late cornified envelope 1C gene (LCE1C)[40], and is one of a family of genes that are transcribed very late in epidermal differentiation. Our localization of PSORS4 to the LCE region is of interest, and its refinement will be described in Table 3.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1q21</td>
<td>rs6701216</td>
<td>151045150</td>
<td>EDC</td>
<td>0.0034</td>
<td>0.0053</td>
<td>0.0069</td>
<td>0.00005</td>
<td>T</td>
<td>0.174</td>
<td>0.127</td>
<td>1.45 (1.21,1.75)</td>
</tr>
<tr>
<td>2p11</td>
<td>rs2164807</td>
<td>85816062</td>
<td>GNLY-ATOH8</td>
<td>0.0084</td>
<td>0.0115</td>
<td>0.0039</td>
<td>0.000016</td>
<td>G</td>
<td>0.467</td>
<td>0.394</td>
<td>1.35 (1.18,1.54)</td>
</tr>
<tr>
<td>2p12</td>
<td>rs11126740</td>
<td>79754603</td>
<td>CTNNB2</td>
<td>0.0018</td>
<td>0.0018</td>
<td>0.019</td>
<td>0.00014</td>
<td>T</td>
<td>0.311</td>
<td>0.37</td>
<td>0.77 (0.67,0.88)</td>
</tr>
<tr>
<td>3q13</td>
<td>rs6804331</td>
<td>105237077</td>
<td>Gene desert</td>
<td>0.0073</td>
<td>0.013</td>
<td>0.0455</td>
<td>0.0003</td>
<td>C</td>
<td>0.445</td>
<td>0.381</td>
<td>1.30 (1.13,1.50)</td>
</tr>
<tr>
<td>13q13</td>
<td>rs3812888</td>
<td>39126294</td>
<td>COG6</td>
<td>0.0017</td>
<td>4×10⁻⁴</td>
<td>0.00001</td>
<td>0.43</td>
<td>0.35</td>
<td>1.38 (1.15,1.66)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13q13</td>
<td>rs7993214</td>
<td>39248912</td>
<td>COG6</td>
<td>4.7×10⁻⁵</td>
<td>0.0001</td>
<td>0.033</td>
<td>2×10⁻⁵</td>
<td>T</td>
<td>0.279</td>
<td>0.351</td>
<td>0.71 (0.62,0.82)</td>
</tr>
<tr>
<td>14q32</td>
<td>rs2282276</td>
<td>94730882</td>
<td>CLMN</td>
<td>0.0096</td>
<td>0.014</td>
<td>0.047</td>
<td>0.0031</td>
<td>G</td>
<td>0.099</td>
<td>0.073</td>
<td>1.40 (1.12,1.76)</td>
</tr>
<tr>
<td>15q21</td>
<td>rs4775912</td>
<td>49068271</td>
<td>USP8-TNFAIP8L3</td>
<td>0.00034</td>
<td>0.00065</td>
<td>0.0136</td>
<td>5.6×10⁻⁵</td>
<td>G</td>
<td>0.194</td>
<td>0.146</td>
<td>1.41 (1.19,1.67)</td>
</tr>
<tr>
<td>15q21</td>
<td>rs3803369</td>
<td>49163121</td>
<td>USP8-TNFAIP8L3</td>
<td>0.0012</td>
<td>0.00252</td>
<td>0.0138</td>
<td>2.9×10⁻⁵</td>
<td>A</td>
<td>0.195</td>
<td>0.145</td>
<td>1.43 (1.21,1.69)</td>
</tr>
</tbody>
</table>

Trend *P* values from discovery, replication and combined analyses for U.S. PS samples (810 cases, 1256 controls) are shown. The minor allele, its frequency in cases and controls and its odds ratios and 95% C.I.s from the combined data-set are also shown. Abbreviations are described in footnotes to Table 1. doi:10.1371/journal.pgen.1000041.t003

As discussed earlier, this region is one that is most likely to harbor the PSORS1 variant. To ensure that our observations were not due to cross-hybridization of chromosome 15q21 SNPs with PSORS1 SNPs, we investigated alignment of genes from this region of chromosome 15q21 with the remainder of genome. We did not detect any significant identity with any other region, including the MHC. This, and the fact that chromosome 15 associated SNPs are in Hardy-Weinberg equilibrium indicate that our results are unlikely to be artifactual and due to amplification of PSORS1 sequences. The similarity between the PSORS1 and 15q21 variants and their biological consequences need to be investigated further since it may provide important insights into the nature of the PSORS1 variant. However, it is worth noting that our PSA cases which provided strongest evidence for association with 15q21 all had at least one first degree relative with PS, and association with this locus may be harder to detect in “sporadic” cases.

We also observed association of PS with a region of the Epidermal Differentiation Complex (EDC), which harbors a previously established psoriasis locus (PSORS4)[39]. In this instance, rs6701216 yielded a combined *P* = 6.2×10⁻⁵ (OR 1.45) (Table S3, Figure S4). This SNP lies within the late cornified envelope 1C gene (LCE1C)[40], and is one of a family of genes that are transcribed very late in epidermal differentiation. Our localization of PSORS4 to the LCE region is of interest, and its refinement will be described in detail elsewhere (Robarge et al., in preparation).

There were two other regions selected for follow-up where *P* values were <0.05 in the PS replication cohort, and where evidence for association increased in the combined cohort (Table 3). One was an intergenic region located between granulysin (GNLY) and atonal (ATOH) on chromosome 2p11. The most highly associated SNP was rs2164807 (adjusted *P* = 0.0015, GWA scan; *P* = 0.0039; replication; *P* = 1.6×10⁻⁵ combined). For this SNP, the G allele was found at a frequency of ~0.47 in cases and ~0.39 in controls (OR: 1.35, 95% CI: 1.18–1.54). GNLY (Protein NKG5, Lymphokine LAG-2) is of considerable interest with respect to psoriasis. It is present in cytotoxic granules of cytotoxic T lymphocytes and natural killer cells, and is released upon antigen stimulation[41]. It has been shown to have antimicrobial activity against a broad range of microbes including Gram-positive and Gram-negative bacteria, fungi, and parasites including *M. tuberculosis* and other organisms. Priming of granulysin in CD4 is dysregulated in the CD4+ T cells of HIV-infected patients[42].
Other genes that should be evaluated in additional PS cohorts on the basis of replication P values <0.05, and increased significance in combined cohorts (Table 3) are calponin-like transmembrane domain protein (CLMN)[43], the gene for the catenin member, CTNNA2[44] and a gene desert on chromosome 3q13.

Association with Chromosome 4q27: A Recently Identified Autoimmune Locus

In the discovery cohort there were 3 SNPs from chromosome 4q27 with P<5×10⁻³ (adjusted P<10⁻⁴). These were rs13151961, rs6822844 and rs6840978 (Table S2, Figure S6). The most significant of these was rs13151961, where adjusted P = 4×10⁻⁵ (Table 4). Association of this region with PSA was confirmed in an independent cohort of patients from the UK (for rs13151961, P = 0.003, Table 4), where the frequency of the associated T allele was ~0.25 in cases and ~0.31 in controls (OR: 0.64; 95% CI: 0.49–0.84). Association with additional SNPs in high LD with rs13151961 (rs6840978 and rs6822844), was also replicated in this cohort. Although association could not be replicated in our U.S. cohort with a case/control approach, the trend in allele frequencies in cases versus controls was similar to that seen in the U.K. cohort (for rs13151961 the frequency of the T allele was ~0.26 in U.S. cases and ~0.29 in our U.S. controls, Table 4).

A recent study from the Wellcome Trust Case Control Consortium [19] identified this 4q27 region in a search for risk factors for type 1 diabetes (T1D). In a follow-up study [45], some support for this association with T1D was provided. In the latter
To explore association of this region in our cohort of U.S. PS cases, without the possible confounding influence of subtle geographic variability at this locus in Europeans, we performed family based association tests in our 242 psoriasis nuclear families which are described elsewhere[26,49]. This approach provided evidence for replication of association of PS with rs6822844 and rs6840978 (PDT P = 0.029 and 0.007 respectively). For these SNPs, the over-transmitted rs6822844*G and rs6840978*C alleles were also the risk allele from case/control studies. Haplotype studies in families also revealed association with the rs6822844*G/rs6840978*C haplotype (multipoint TDT P = 0.006). Hence, our findings support chromosome 4q27 as harboring a variant/haplotype for PSA and PS.

As reported elsewhere the 4q27 locus that contains these associated SNPs corresponds to two closely correlated \(~439\) kb and \(~40\) kb haplotype blocks[46]. This extensive LD makes it very difficult to determine the predisposing variant. Chromosome 4q is also the location of PSORS3, which is generally placed slightly more distally[49]. However, the locus identified here may contribute in part to the previous observations of linkage. The long region of LD at chromosome 4q27 contains several genes[46]: T-2 nuclear RNA-binding protein (TENR), a gene encoding a protein of unknown function (KIAA1109), and genes encoding the interleukin-2 (IL2) and interleukin-21 (IL21) cytokines. TENR is expressed primarily in testis and KIAA1109 transcripts are ubiquitous, hence their roles in autoimmunity are not particularly compelling. However, IL2 and IL21 are of particular interest with respect to PS. IL2 is considered to be a pathogenic cytokine for PS[50], and blockade of the IL2 receptor with therapeutic antibodies has induced disease resolution in some cases[51]. IL2 is a survival factor for T cells and promotes the differentiation of cytotoxic T-lymphocytes and NK cells. Both of these cell types are present in psoriasis lesions. Moreover, many IL2 receptor (IL2R) positive T-cells that fit the phenotypic definition of regulatory T cells (Treg) are also present in psoriasis lesions. IL-2 may influence how a common precursor T-cell differentiates into either a Treg or a Th17 T-cell, since addition of IL-2 has been shown to suppress the differentiation of Th17 T-cells in mice[52]. IL-2 antibodies elicit normal epidermis and psoriatic epidermis, with generally lower staining in the dermis (Figure 5). The pattern of staining appears to be to dendritic cells (DCs) which are likely to be epidermal Langerhans cells. This pattern of IL-2 staining is probably due to DC activation and upregulation of IL2R. Cells with IL2 receptors include T-cells, B-cells, NK-cells, and dendritic cells.

The epidermal staining for IL-21 is much lower than for IL-2 (Figure 5) and appears to be mainly on dendritic cells in the superficial dermis. IL-21 is a product of activated T cells (under conditions of Th17 polarization). It then acts in an autocrine or paracrine fashion on T-cells to up-regulate expression of the IL23 receptor which has already been implicated in psoriasis pathogenesis. IL23R sensitizes cells to IL-23 which stimulates IL17 synthesis and/or prolongs the survival of Th17 cells[53]. Blocking IL21 reduces the progression of lupus erythematosus[54] which is one of the autoimmune diseases that is now being considered as a “Th17” mediated disease. Therefore IL21 may play a role in Th17 formation in this and other autoimmune diseases where these cells are pathogenic.

Extensive resequencing of IL2 coding and flanking regions has already been performed in TID samples and no coding or obvious regulatory/splice variants were identified[45]. As stated previously, this region needs to be resequenced thoroughly followed by comprehensive genotyping in larger numbers of samples to identify the autoimmune associated variant/s[45].
Conclusions

The observed associations in the current study are of interest for several reasons. It is noteworthy that the strongest association is with the MHC. Even in PSA, where associations are reportedly less than with PS (without PSA), associations with the class I region appear to be more significant than with any other region. We were also able to replicate previously reported associations with IL12B and IL23R and detected a potentially novel association upstream from IL12RB2. Novel associations within COG6 and the region on chromosome 15q21 harboring USP8 and SPPL2A are of interest. These and other regions reported here are worthy of follow up in other cohorts. Moreover, the association with chromosome 4q27 provides further evidence that this region is a common locus for multiple forms of autoimmune disease.

A recent study reported the IL13/IL4 region from chromosome 5q31 as being associated with PS[55]. Overall, the risk contributed by the MHC, the IL13/IL4 region and the IL23R and IL12B variants was estimated to be 3.83. With the COG6 and chromosome 15q21 loci described here, the risk would be increased. However, PS is a complex disease, and overlapping subsets of risk factors may be sufficient for susceptibility, so that risk effects cannot be computed in an additive manner.

The ability to identify low risk variants for common diseases such as PS and PSA will be limited by the cohort size, and larger...
Table 4. Association results at chromosome 4q27 in PSA and PS cohorts (U.S.: 810 PS cases, 1256 U.S. controls; U.K.: 576 PSA cases, 480 controls).

<table>
<thead>
<tr>
<th>SNP</th>
<th>Location (hg18)</th>
<th>Minor allele</th>
<th>UK PSA P value</th>
<th>OR (95% CI)</th>
<th>Rep. US PS P value</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs13151661</td>
<td>123334952</td>
<td>G</td>
<td>2.5×10⁻⁶</td>
<td>0.121</td>
<td>0.64 (0.49, 0.84)</td>
<td>0.17</td>
</tr>
<tr>
<td>rs7684187</td>
<td>123560609</td>
<td>T</td>
<td>2.4×10⁻⁴</td>
<td>0.247</td>
<td>0.72 (0.59, 0.87)</td>
<td>0.52</td>
</tr>
<tr>
<td>rs6822844</td>
<td>123728877</td>
<td>A</td>
<td>2.0×10⁻⁴</td>
<td>0.192</td>
<td>0.155 (0.12, 0.19)</td>
<td>0.43</td>
</tr>
<tr>
<td>rs6849798</td>
<td>123774105</td>
<td>A</td>
<td>6.6×10⁻⁴</td>
<td>0.296</td>
<td>0.192 (0.16, 0.23)</td>
<td>0.43</td>
</tr>
</tbody>
</table>

The minor allele, its frequency in cases and controls and its odds ratios and 95% C.I.s from the combined dataset are also shown. Abbreviations are described in footnotes to Table 1.

Genome-Wide Association Study of Psoriasis

Materials and Methods

Subjects

The cohorts for the discovery and replication phases of this study are all of European descent and are described in Table S1. The discovery cohort consisted of 223 Caucasian individuals with PS or PSA from the US. Cases were ascertained through active follow up by hospital dermatologists or from clinical settings and a genomic DNA was isolated from whole blood by standard procedures.

Replication cohorts were from both the U.K. and the U.S. The U.K. cohort consisted of 576 PSA patients from the UK and are described elsewhere [60]. In brief, PSA patients under active follow up by hospital dermatologists were recruited from throughout the UK although the majority came from the North-West region of England. All patients satisfied the inclusion criteria of having both clinically documented inflammatory synovitis and PS. Each patient was assessed by a trained research nurse, who undertook a standardized clinical history and examination. Detailed demographic and clinical information was obtained and whole blood was taken for DNA extraction and subsequent genetic analysis. Control samples (n = 480) were obtained from blood donors.
Genotyping Methods

DNA was normalized to a concentration of 100 ng/μl (diluted in 10 mM Tris/1 mM EDTA). Samples were quantitated with a Nanodrop Spectrophotometer (ND-1000). For the discovery phase, approximately 1 μg of genomic DNA was used to genotyoe each sample on the Illumina HumanHap300v2A Genotyping BeadChip. This was performed at the Robert S. Boas Center for Genomics and Human Genetics at The Feinstein Institute for Medical Research, Manhasset, NY. This assay relies on allele specific primer extension and the use of a single fluorochrome. Samples were processed according to the standard Illumina Infinium II automated protocol. This involved whole genome amplification, fragmentation, precipitation, resuspension in hybridization buffer and hybridization to the Illumina Bead Chips for a minimum of 16 h at 48°C. After hybridization the BeadChips were processed for the single base extension reaction, followed by staining and imaging on an Illumina Bead Array Reader. Normalized bead intensity data were loaded into the Illumina Beadstudio 2.0 software which converted fluorescence intensities into SNP genotypes.

Genotyping for all the replication studies was performed with the Sequenom MassArray system (iPLEX assay). This involves primer extension chemistry and mass spectrometric analysis described as our web site http://hg.wustl.edu/info/Sequenom_description.html.

Quality Control

Before analysis, we performed quality filtering of both samples and SNPs to ensure robust association tests. Based on previous criteria [61], we required that all samples used for the discovery phase pass a 93% genotyping call rate threshold, and that all SNPs pass a 95% call rate threshold.

In the case of the replication studies, 57 individuals from the total of 2370 individuals in the replication study were removed because of low genotyping (i.e. when over half of the genotypes for a sample were missing). SNPs with <75% call rates were also excluded from analysis to obtain an average genotyping rate of 0.902. Genotypes were also evaluated for departure from HWE in controls. Because of low genotyping (i.e. when over half of the genotypes for a sample were missing), SNPs with <75% call rates were also excluded from analysis to obtain an average genotyping rate of 0.902. Genotypes were also evaluated for departure from HWE in controls. After pruning, 244 SNPs remained.

A total of 463 ancestry informative SNPs (AIM) present on the Illumina HumanHap300v2A Genotype BeadChip were used to check for possible confounding population substructure in the discovery sample with STRUCTURE software ([62]). For this analysis, genotypes at these SNPs were analyzed for all 742 samples (225 PS cases and 519 controls).

To investigate other biases [20] that could be introduced with shared controls we assessed the potential effect of substructure with the genomic-control method [21] and with EIGENSTRAT [23].

Statistical Analysis for Association

The Cochran-Armitage Test for trend [25] was conducted with Purcell’s PLINK program (http://pngu.mgh.harvard.edu/~purcell/plink). However, several SNPs in the current study that exhibited significant differences in cases/controls, were also different when allele frequencies in controls were compared with those from European CEPH typed for SNPs in the HapMap project. NYCP participants are quite diverse with respect to European origin, and many SNPs are reported to show differences among European subgroups [63]. These were identified by a comparison of SNP allele frequencies in European CEPH individuals typed for the HapMap project and were not selected for follow-up studies.

Measures of linkage disequilibrium, D’ and r², and allele frequencies were based on precomputed scores from the International HapMap website or were computed locally from HapMap genotypes or from our own case and control genotypes with Haplovip 3.2 (http://www.broad.mit.edu/mpg/haplovip/).

Power calculations for association were calculated at: http://pngu.mgh.harvard.edu/~purcell/gpc/. Association localization plots (Figures 2–4, 6, S4–S5) were generated with an R code modified from snp.plotter (http://cshl.nimh.nih.gov/~kristin/snp.plotter.html) and Regional Association Plot (http://www.broad.mit.edu/diabetes/scandinavs/figures.html).

Family Based Tests

Family based association tests on 271 nuclear families were performed with the Pedigree Transmission Disequilibrium Test [64] as described elsewhere [11, 12, 65].

Immunohistochemistry

Tissue sections were fixed with acetone and stained with 10 μg/mL purified mouse anti-human monoclonal antibodies to IL-2 (R&D, clone 5334.21), IL-21 (R&D, J148-1134), COG6 (Abnova, H00057511-M01) and SPPL2A (Abgent, AP6312a). Biotin labeled horse anti-mouse antibodies (Vector Laboratories) were amplified with avidin-biotin complex (Vector Laboratories) and developed with chromogen 3-amino-9-ethylcarbazole (Sigma Aldrich).

Supporting Information

Table S1 Summary of cases and controls used in discovery and replication stages. Found at: doi:10.1371/journal.pgen.1000041.s001 (0.03 MB DOC)

Table S2 Top ranking SNPs where P<5×10⁻⁶ in discovery cohorts (also stratified on the basis of the presence of PSA or no PSA). Found at: doi:10.1371/journal.pgen.1000041.s002 (0.25 MB PDF)

Figure S1 Heterozygosity of sample versus genotyping call rate. Found at: doi:10.1371/journal.pgen.1000041.s003 (0.95 MB TIF)

Figure S2 Distribution of SNP success rate in the discovery study. Found at: doi:10.1371/journal.pgen.1000041.s004 (1.36 MB TIF)

Figure S3 Q-Q plot of GWA analyses in unrelated individuals used in the discovery study obtained with PLINK. Black dots
indicate the negative log of unadjusted trend P values ($\lambda_{GC} = 1.101$) and red dots indicate the negative logarithm of adjusted trend P values ($\lambda_{GC} = 1$). Found at: doi:10.1371/journal.pgen.1000041.s005 (6.03 MB TIF)

**Figure S4** Association localization plot for replicated region at chromosome 4q21. Results for SNPs used in the discovery phase (adjusted for GC) are presented as diamonds. Negative LOG P values are provided on the Y axis. The X axis corresponds to the locations of SNPs. The P value for all samples (original GWA scan plus replication samples) are shown as circles. The P value obtained with the most highly associated SNP (from the original GWA scan plus the replication samples) is shown as a red circle. The SNPs shown as orange diamonds are in r2>0.8 (European HapMap CEPH (CEU) samples) with the most significant SNP identified in our study. The recombination rate based on the CEU HapMap is shown in light blue along the x axis (scale on the right). The green arrows indicate the locations of select genes. The LD relationship of Illumina SNPs derived from CEU HapMap genotypes are shown below the graph. The most highly associated SNP, rs6701216, is indicated with an asterisk above the LD plot.

Found at: doi:10.1371/journal.pgen.1000041.s006 (9.70 MB TIF)

**Figure S5** Association localization plots for autoimmune locus on chromosome 4q27 showing P values obtained in discovery sample. Symbols are the same as those used in Figure S4. The asterisk above the LD plot corresponds to rs6849978.

Found at: doi:10.1371/journal.pgen.1000041.s007 (9.39 MB TIF)

**Acknowledgments**

We are indebted to the many patients and their families who participated in this study. We thank Annette Lee and Peter Gregersen for hybridizations to the Illumina arrays. Drs. A. Perlmutter (Texas Dermatology), Terry Miner (St. Louis University), Richard Brasington (Washington University Medical Center), and John Koo/Psoriasis Center Staff (UCSF) provided some of the samples used in this study. We thank Rachel Donaldson for technical help, Weimin Duan for help with analysis and Michael Lovett for critical comments on the manuscript.

**Author Contributions**

Conceived and designed the experiments: A. Bowcock. Performed the experiments: L. Zaha, S. Duan. Analyzed the data: Y. Liu, C. Helms, J. Kane, N. Saccone. Contributed reagents/materials/analysis tools: W. Liao, J. Gardner, C. Wise, A. Miner, M. Malloy, C. Pullinger, J. Kane, J. Worthington, I. Bruce, P. Kwok, A. Menter, A. Barton. Wrote the paper: A. Bowcock. Prepared figures: S. Saccone. Revised the manuscript: A. Bowcock.

**References**