Dense Genotyping of Immune-Related Regions Identifies Loci for Rheumatoid Arthritis Risk and Damage in African Americans

Maria I Danila,1* Vincent A Laufer,1* Richard J Reynolds,1 Qi Yan,2 Nianjun Liu,3 Peter K Gregersen,4 Annette Lee,4 Marlena Kern,4 Carl D Langefeld,5 Donna K Arnett,6 and S Louis Bridges, Jr.1

1Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Birmingham, Alabama, United States of America; 2Division of Pulmonary Medicine, Allergy and Immunology; Department of Pediatrics, University of Pittsburgh, Pittsburgh, Pennsylvania, United States of America; 3Department of Epidemiology and Biostatistics, School of Public Health, Indiana University Bloomington, Bloomington, Indiana, United States of America; 4The Feinstein Institute for Medical Research, Northwell Health, Manhasset, New York, United States of America; 5;Wake Forest School of Medicine, Wake Forest University, Winston-Salem, North Carolina, United States of America; and 6University of Kentucky College of Public Health, Lexington, Kentucky, United States of America

More than 100 risk loci for rheumatoid arthritis (RA) have been identified in individuals of European and Asian descent, but the genetic basis for RA in African Americans is less well understood. We genotyped 610 African Americans with autoantibody-positive RA and 933 African American controls on the Immunochip (iChip) array. Using multivariable regression, we evaluated the association between iChip markers and the risk of RA and radiographic severity. The single nucleotide polymorphism (SNP) rs1964995 (odds ratio = 1.97, \( p = 1.28 \times 10^{-15} \)) near HLA-DRB1 was the most strongly associated risk SNP for RA susceptibility; SNPs in AFF3, TNFSF11 and TNFSF18 loci were suggestively associated (\( 10^{-4} < p < 3.1 \times 10^{-6} \)). Trans-ethnic fine mapping of AFF3 identified a 90% credible set containing previously studied variants, including rs9653442, rs7608424 and rs6712515, as well as the novel candidate variant rs11681966; several of these likely influence AFF3 gene expression level. Variants in TNFRSF9, CTLA4, IL2RA, C5/TRAF1 and ETS1 – but no variants within the major histocompatibility complex – were associated with RA radiographic severity. Conditional regression and pairwise linkage disequilibrium (LD) analyses suggest that additional pathogenic variants may be found in ETS1 and IL2RA beyond those found in other ethnicities. In summary, we used the dense genotyping of the iChip array and the unique LD structure of African Americans to validate known risk loci for RA susceptibility and radiographic severity, and to better characterize the associations of AFF3, ETS1 and IL2RA.

INTRODUCTION

Rheumatoid arthritis (RA) is a systemic autoimmune disorder characterized by synovial joint inflammation, with disease phenotype ranging from mild joint involvement to severe joint destruction and permanent disability (1). The factors responsible for RA heterogeneity are poorly understood, but both genetic and environmental factors contribute to its pathogenesis and clinical expression. Most RA patients have serum autoantibodies, such as rheumatoid factor (RF) or anticyclic citrullinated peptide antibody (ACPA), which can be present before the onset of clinically relevant disease (2) and are associated with radiographic severity (3).

To replicate and fine map risk loci identified in genome-wide association studies (GWAS) of autoimmune and inflammatory disorders such as RA, the Immunochip Consortium designed the Immunochip (iChip), a custom Illumina Infinium high-density array that has been used to study RA in patients of several racial and ethnic backgrounds (4–8). Using the iChip and many other arrays, 100 RA risk loci of genome-wide significance (\( p < 5 \times 10^{-8} \)) have been identified in individuals of European and
Asian ancestry, including HLA-DRB1, PAD4, PTPN22 and CTLA4 (9,10). However, there is a paucity of genetic association data on RA in African Americans. Many of the genetic influences on RA are similar among those of European ancestry and African Americans (11). However, there are important differences; for instance, polymorphisms in CCR6, TAGAP and TNFAIP3 have discordant odds ratios (ORs) compared with those reported in European RA patients (11). Furthermore, the PTPN22 risk allele containing rs2476601, which has the highest effect size on RA susceptibility of any locus outside the major histocompatibility complex (MHC) in European populations, is essentially absent from the Yoruban population and is present in Asian and African American populations in low frequency (12).

HLA-DRB1 alleles encoding the shared epitope (SE) (13) have the strongest association with RA in Europeans and Asians. In addition to their role in susceptibility of RA, HLA-DRB1 SE alleles are associated with erosive disease (14) and mortality in Europeans (15,16). Our group has shown that 43% of African Americans with RA have at least one HLA-DRB1 SE allele compared with ~60–70% of Europeans with RA (17). At the level of HLA-DRB1 amino acid residues and their association with RA susceptibility, there are both similarities and major differences between Europeans and African Americans. The valine residue at position 11, as found in Europeans, is most strongly associated with RA in African Americans (18). However, an aspartic acid residue at position 11, indicative of the classical allele *09:01, confers a two-fold increased risk of RA in African Americans and is also associated with RA in Koreans (19), but not in individuals of European ancestry. After conditioning on residue substitutions at position 11, amino acid positions 71 and 74 are not significantly associated with RA in African Americans, as they are in Europeans (18).

Subphenotypes (ie, disease subtraits) (20) are frequently more heritable than the complex disease traits of which they are a part (21). In addition, genetic association studies on specific subphenotypes tend to focus on less heterogeneous patients than studies on overall disease susceptibility. Radiographic severity is a characteristic RA subphenotype, with an estimated heritability between 45% and 58% (22). Genetic influences on radiographic severity have been examined in several ethnic groups (23–31), and ~30 risk loci have been identified in Europeans and Asian populations, including CXCR5, AFF3, C5-TRAF1, IL2RA, IL6, IL10 and FCRL3. Typically, these studies have included smaller numbers of participants than studies on susceptibility of RA, and the definition of radiographic severity has not been uniform, which may limit statistical power and complicate replication. Finally, few studies have addressed radiographic severity of RA in African Americans.

The objective of our study is to investigate the associations of known autoimmune disease risk loci with RA and its radiographic severity in African Americans. In view of the heterogeneity of disease associations among ethnic groups, shorter haplotype blocks and differences in allele frequency in African Americans, we hypothesize that fine mapping will identify differences in the genetic architecture of RA in African Americans compared with other ethnicities. Strengths of this study include analysis of the largest group of African Americans with RA currently available in the world, with accompanying high-quality radiographic outcomes data (34). A major goal of research into the genetics of complex diseases is to identify pathogenic variants that produce disease associations. In view of this, we draw on available data from association testing of >100,000 Asians and Europeans and use cutting-edge algorithms to prioritize genetic risk variants in African Americans with RA in the AFF3 locus. This study represents an important addition to the literature on the genetics of RA, which has primarily involved participants of Asian and European descent.

**MATERIALS AND METHODS**

**Study Population**

The CLEAR (Consortium for the Longitudinal Evaluation of African Americans with Early Rheumatoid Arthritis) study enrolled African Americans with RA of <2 years’ disease duration (CLEAR I), African Americans with RA irrespective of disease duration (CLEAR II) and African American healthy controls, as previously described (17). Participants were enrolled at five academic sites: University of Alabama at Birmingham (coordinating center); Grady Hospital/Emory University, Atlanta, Georgia; University of North Carolina, Chapel Hill; Washington University, St. Louis, Missouri; and Medical University of South Carolina, Charleston. CLEAR controls were African Americans without rheumatic diseases who were matched (as a group) by age, sex and geographic location to CLEAR RA patients. The Institutional Review Boards of the participating institutions approved human subject research protocols. Biologic specimens and patient information, including sociodemographic characteristics, medical history, medications and disease activity measures, were collected (17). The majority of CLEAR participants were ACPA-positive, as previously reported (32). Of the 837 African American healthy controls included for analysis in the current study, 404 were from CLEAR and 433 were from the Birmingham, Alabama area (33).

Radiographs of hands/wrists and feet were obtained at the CLEAR enrollment visit for participants with RA and assigned a modified total Sharp score (mTSS) (range 0–448) using the modified Sharp/van der Heijde method (34). Scoring was performed using state-of-the-art methods under the auspices of Désirée van der Heijde, a world expert in quantitative assessment of radiographs in rheumatic diseases (34). Furthermore, mTSS scores for participants from CLEAR I and CLEAR II have been validated extensively (35,36).
Sample Genotyping

Genotyping was carried out using the iChip array at the Feinstein Institute for Medical Research in Manhasset, New York. Genotype clustering was performed using the GenTrain2 clustering algorithm. Genotype calling was performed with the genotyping module of the GenomeStudio data analysis software package.

Quality Control

Rigorous quality control procedures were employed, including checks for gender inconsistency, relatedness (duplicates and first- or second-degree relatives) and ethnic outliers. The sample call rate threshold was 95%. The marker call rate was >98.5% across all SNPs, after removing low-quality SNPs and rare SNPs, those with minor allele frequency (MAF) <5% and SNPs out of Hardy-Weinberg equilibrium (using control samples only, using $p$ value >1 x 10^-6).

Association Testing of iChip Markers with RA Susceptibility

Of 610 RA cases, 593 (97%) were autoantibody-positive (defined as positive for either RF or anti-CCP antibody tests) and were included in the analysis of RA susceptibility. Multivariable logistic regression was used to evaluate the association between iChip markers and autoantibody-positive RA. Sex and European admixture proportion (calculated using Eigenstrat v6.0) (17) were included as covariates. Two-sided $p$ values are reported, except as noted for trans-ethnic fine mapping of the AFF3 locus (see Trans-ethnic fine mapping of the AFF3 locus section). To adjust for variability due to HLA-DRB1 in the extended MHC locus (Chr6:26,000,000–34,000,000), we fit a model accounting for the variability of all four-digit HLA-DRB1 SE alleles. locusZoom plots were used to display the fine mapping results graphically (37).

Association Testing of iChip Markers with RA Severity

The modified total radiographic scores (mTSS) were overdispersed in the CLEAR cohort, with a high proportion of individuals having no erosions or joint space narrowing (mTSS = 0); 156 of 230 CLEAR I participants (67.8%) and 150 of 365 CLEAR II participants (41.1%) (see Supplementary Figure S1). We assessed several count regression models and found that the zero-inflated negative binomial model had the best fit for the data, likely due to the high proportion of participants without damage (mTSS = 0). Thus, we used this method to evaluate the association of genetic markers with radiographic severity (under an additive genetic model). Association testing was carried out using the PSCL package in R (38) after adjusting for body mass index, sex, smoking status, percent European admixture (see [17] for details) and disease duration (in months) as covariates. Due to the inclusion criteria, disease duration was much shorter in CLEAR I (early RA) (median 1.01 years; interquartile range 0.57–1.52 years) than in CLEAR II (any disease duration) (median 9.25 years; interquartile range 3.42–17.75 years). Using a square root transformation for disease duration improved the model fit and reduced genomic inflation ($\lambda_{GC} = 1.10$) compared with a model using untransformed disease duration. After removal of the SNPs in the extended MHC and other associated loci, the $\lambda_{GC}$ value was further reduced to 1.04 (See quantile-quantile plot in Supplementary Figure S2).

Trans-Ethnic Fine Mapping of the AFF3 Locus

We conducted trans-ethnic fine mapping of the AFF3 locus, combining our RA susceptibility data with those from a previously published large trans-ethnic meta-analysis (10). To accomplish this, we: (1) aligned reference and alternate alleles from all Asian, European and CLEAR populations to match those from the 1000 Genomes project (39); (2) generated LD matrices either from our genotyping data (African Americans) or from the 1000 Genomes project (Asian and European populations; data from Okada et al.) (10); (3) annotated SNPs from all three ethnicities using 8,138 genomic annotations (for example, DNAse hypersensitivity, enhancer markings and so on) provided with the PAINTOR3 algorithm; and (4) trimmed these to the top five uncorrelated annotations (correlation coefficient <0.10), excluding the annotations with lower Bayes factors.

We then confirmed the algorithm was working properly by examining the results it produced in RA loci in which the causal variant was known. For instance, we generated a posterior probability of 1.0 for rs2476601 in PTPN22 in Europeans with RA. Following this, we calculated the posterior probability that each variant in the AFF3 locus was pathogenic using PAINTOR3 (39), which assigns a probability ranging from 0 (very unlikely) to 1 (highly likely). We ran the algorithm using genetic data from all three populations, and defined a “90% credible set” for candidate pathogenic variants as previously reported (40) (see Table 3). Although PAINTOR3 is capable of modeling more than one causal variant per locus, in this study we conducted trans-ethnic fine mapping under the assumption of one causal variant.

Calculation of Number of Effective Markers for the iChip Array

Because the iChip contains many variants concentrated in specific loci and in LD with one another, the number of independent tests is much smaller than the actual number of variants genotyped. Estimates of the number of effective markers for custom genotyping arrays such as the iChip vary widely, between 2,800 and 60,000 LD-independent markers (41–43). To find independent SNPs, we used Plink (44), as previously utilized for iChip data (45), and found 16,154 LD-independent SNPs. We thus defined an iChip-wide statistical significance threshold as 0.05 divided by 16,154 LD-independent SNPs, or $p = 3.1 \times 10^{-6}$, similar to previous reports. We report any variants having $p < 1 \times 10^{-4}$ as showing suggestive statistical associations (for both susceptibility and severity).
**AFRICAN AMERICAN RA: MAPPING AUTOIMMUNE LOCI**

Table 1. Demographic, clinical, genetic and radiographic characteristics of African American participants with RA and healthy controls from the CLEAR registry.

<table>
<thead>
<tr>
<th>Baseline characteristics</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, mean (SE)</td>
<td>50.0 (13.0)</td>
<td>48.1 (12.4)</td>
</tr>
<tr>
<td>Sex (female), %</td>
<td>82.7</td>
<td>75.5</td>
</tr>
<tr>
<td>Disease duration in months, mean (SE)</td>
<td>12.9 (7.1)</td>
<td>31.4 (7.8)</td>
</tr>
<tr>
<td>Body mass index, mean (SE)</td>
<td>0.17 (0.09)</td>
<td>0.16 (0.09)</td>
</tr>
<tr>
<td>Global European admixture estimate, mean (SE)</td>
<td>6.32 × 10^{-4}</td>
<td>0.70</td>
</tr>
<tr>
<td>Number of tender joints, median (IQR 25–75)</td>
<td>4.0 (1.0–12.0)</td>
<td>4.0 (1.0–10.0)</td>
</tr>
<tr>
<td>Number of swollen joints, median (IQR 25–75)</td>
<td>3.0 (1.0–7.0)</td>
<td>4.0 (1.0–9.5)</td>
</tr>
</tbody>
</table>

Medications

- Biologics ever used (%): 4.4, 19.5
- Methotrexate, current use (%): 63.9, 60.3

Radiographic score, mean (SE)

| Joint-erosion score                           | 1.6 (4.3)              | 10.7 (18.5)            |
| Joint-narrowing score                        | 2.1 (5.7)              | 18.1 (27.7)            |
| Total score                                   | 3.7 (9.4)              | 28.8 (44.1)            |

SE: standard error; IQR: interquartile range; DMARD: disease-modifying anti-rheumatic drug.

**RESULTS**

Following quality control procedures, 100,268 SNPs with MAF >0.05 were available for analysis in 610 RA cases and 837 healthy controls (as stated in Materials and Methods, 593 [97%] were autoantibody-positive and were included in subsequent analyses). The demographic characteristics of African Americans from the CLEAR registry included in this study are presented in Table 1. Characteristics of Birmingham controls did not differ significantly from the CLEAR registry with respect to sex, European admixture proportion or other variables (17).

**Association between iChip Markers and RA Susceptibility**

We evaluated the association between iChip markers and RA using logistic regression and adjusting for the proportion of overall European admixture. We observed seven non-HLA loci suggestively associated with RA (defined as p < 10^{-5}); the lead SNP (ie, the most strongly associated) at each locus is shown in Table 2. As expected, the markers with the strongest association with RA were found in the MHC region (Figure 1A). We identified rs1964995 in HLA-DRB1 as the variant with strongest association with RA (OR = 1.97, p = 1.28 × 10^{-15}).

We performed a conditional analysis of the variation contained within the extended MHC region as previously described by Raychaudhuri et al. (46). As shown in Figure 1B, conditioning on the HLA-DRB1 alleles substantially attenuated the strength of association of other variants within the extended MHC region. rs3134792 near HLA-B displayed an OR of 2.01 (95% confidence interval [CI] = 1.42–2.88; p = 9.92 × 10^{-5}) for the association with RA susceptibility after conditioning. This effect size and direction of effect are consistent with those reported for amino acid position 9 of HLA-B in Europeans (OR = 2.12, CI = 1.89–2.38). No variants in HLA-DRB1 were associated with RA after controlling for the HLA-DRB1 alleles. However, the direction of effect and ORs measured in this locus were similar to those found in studies of European populations. Specifically, although above the threshold for statistical significance, rs9277357 had an OR of 1.34 (95% CI = 1.14–1.59; p = 5.39 × 10^{-4}), which is consistent with that previously reported for amino acid position 9 in HLA-DRB1 (OR = 1.40, CI = 1.31–1.50).

**Table 2. Variants outside the HLA region associated with autoantibody-positive RA in African Americans at p < 10^{-5}**

<table>
<thead>
<tr>
<th>rsID</th>
<th>Chr</th>
<th>Position</th>
<th>A1(^a)</th>
<th>OR</th>
<th>95% CI</th>
<th>P value</th>
<th>Nearest genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs61828386</td>
<td>1</td>
<td>17286364</td>
<td>G</td>
<td>0.69</td>
<td>0.58–0.82</td>
<td>1.79 × 10^{-5}</td>
<td>TNFSF18, FASLG</td>
</tr>
<tr>
<td>rs67164098</td>
<td>2</td>
<td>6855613</td>
<td>A</td>
<td>1.67</td>
<td>1.31–2.13</td>
<td>4.65 × 10^{-5}</td>
<td>CNRIP1</td>
</tr>
<tr>
<td>rs11681966</td>
<td>2</td>
<td>100759457</td>
<td>C</td>
<td>1.50</td>
<td>1.23–1.78</td>
<td>4.04 × 10^{-5}</td>
<td>AFF3</td>
</tr>
<tr>
<td>rs10758368</td>
<td>9</td>
<td>36310778</td>
<td>A</td>
<td>0.70</td>
<td>0.58–0.83</td>
<td>5.48 × 10^{-5}</td>
<td>RNF38</td>
</tr>
<tr>
<td>rs9533119</td>
<td>13</td>
<td>43049426</td>
<td>A</td>
<td>1.36</td>
<td>1.17–1.59</td>
<td>6.32 × 10^{-5}</td>
<td>TNFSF11</td>
</tr>
<tr>
<td>rs2934178</td>
<td>15</td>
<td>48218221</td>
<td>C</td>
<td>0.70</td>
<td>0.59–0.82</td>
<td>2.96 × 10^{-5}</td>
<td>SEMA6D</td>
</tr>
</tbody>
</table>

Chr: chromosome.

\(^a\)The test allele and minor allele for this the study.

\(^b\)Indicates a validated risk locus for RA.

All supplementary materials are available online at www.molmed.org.
We performed more detailed analysis on AFF3 (see Materials and Methods and Figure 2), a validated RA risk locus among Europeans and Asians (10,47,48). We found that rs11681966 was suggestively associated with RA in African Americans (OR = 1.5, 95% CI 1.23–1.78, \( p = 4.04 \times 10^{-5} \)). The lead AFF3 SNP (rs9653442) associated with RA in European ancestry (\( p = 3.6 \times 10^{-12} \)) was not strongly associated with RA susceptibility (\( p = 0.015 \)) in African Americans. Similarly, rs10209110, the index variant in AFF3 in an African American reference set (\( p = 0.84 \)). Therefore, due to differing association strengths and LD patterns in the locus, we conducted trans-ethnic fine mapping of this locus using data from African Americans, Asians and European RA patients and controls using PAINTOR3 (39) (see Methods).

Most association studies on AFF3 have examined roughly the region from chr2:100,800,000 to 100,850,000, which contains index variants identified by multiple prior GWAS. However, our index variant (rs11681966, at chr2:100,759,457) is outside this region, located near the 5’ end of AFF3 (>1 kb). Thus, we defined the risk locus as a broader region (from chr2:100,709,000 to 100,875,000). We then conducted trans-ethnic fine mapping using PAINTOR3. Figure 2 shows the 90% credible set for variants in the AFF3 locus, enriched genomic annotations used to help construct the credible set and zoom plots in Figure S3. Consistent with previous reports of autoimmune disease in other ethnicities, our trans-ethnic fine mapping analysis of the AFF3 locus in combined African American, European and Asian RA identified rs6712515 and rs7608424 as likely candidates to be pathogenic variants (see Figure 2 and Table 3). Several of these SNPs are listed as index variants in the National Human Genome Research Institute (NHGRI) GWAS catalog (49) and have been noted in prior studies (50).

Figure 1. (A) Manhattan plot of the association of iChip variants with autoantibody-positive RA in African Americans. The x-axis indicates chromosome and position, the y-axis indicates association strength \(-\log(p)\). The blue line illustrates suggestive statistical association (\( p = 1 \times 10^{-5} \)). The red line illustrates iChip-wide level of statistical significance (\( p = 3.1 \times 10^{-5} \)). (B) Conditional analysis in the extended MHC region (chr2:26,000,000–34,000,000). Axes have the same meanings as in (A). Red, green and blue dots are SNPs in the regions immediately surrounding HLA-DRB1, HLA-DPB1 and HLA-B, respectively. The left panel shows the association summary statistics before conditioning on HLA-DRB1 alleles. The strongest association maps to HLA-DRB1. The right panel shows the locus after conditioning on HLA-DRB1 4-digit alleles.

Association between iChip Markers and Radiographic Severity of RA

After quality control procedures, 100,169 SNPs with MAF >0.05 were available for analysis in 548 autoantibody-positive RA patients who had radiographic scores. A Manhattan plot illustrating the genetic variants associated with severity is shown in Figure 3. In contrast to studies in individuals of European ancestry, we did not find a statistically significant association between SNPs tagging the HLA region and radiographic severity (Table 4).

We detected several suggestive associations, including variants in or near AFF3 (Supplementary Figure S3A), TNFRSF9 (Supplementary Figure S3B), CTLA4 (Supplementary Figure S3C), IL2RA (Supplementary Figure S3D), C5/ITGB1 (Supplementary Figure S3E) and NALCN/ITGB1 (Supplementary Figure S3F). rs506746 (near NALCN/ITGB1) was the most strongly associated variant with radiographic severity (\( p = 4.33 \times 10^{-7} \)), but we could not evaluate support from LD for this association due to low marker density for this region on the iChip array, so no further analysis was performed.

We chose to examine two loci (IL2RA and ETS1) in more detail. Multiple IL2RA variants have been associated with autoimmune conditions (juvenile idiopathic arthritis, type 1 diabetes, systemic lupus erythematosus [SLE], multiple sclerosis, Graves’ disease and so on) (49). Similar to previous reports, we observed a suggestive association in the IL2RA locus (rs7077067) with radiographic severity (\( p = 5.16 \times 10^{-5} \)) (51). rs7077067 was the lead SNP in this study, which differs from that in Europeans, rs2104286 (51). In our study, rs2104286 had MAF = 0.05 and was only weakly associated with RA...
The association of rs7077067 with severity ($p = 8.15 \times 10^{-5}$). Previous studies in other ancestries have noted substantial LD between variants in $IL2RA$ and the analysis, pairwise LD estimates and prior studies (not trans-ethnic fine mapping) to further understand these loci. Adjusting for the effect of rs2104286 did not eliminate the association of rs7077067 with severity ($p = 8.15 \times 10^{-5}$). Previous studies in other ancestries have noted substantial LD between variants in $IL2RA$ and the
ETS1 expression in persons of Asian ancestry only. As expected, this SNP was not associated with RA radiographic severity \((p = 0.11)\) in our dataset, nor was the lead SNP in our study in LD with rs6590330 \(\left(r^2 = 0.03\right)\), or with other previously described variants reported in the NHGRI GWAS catalog \((49)\), for example, rs1128334 \(\left(p = 0.27; r^2 = 0.01\right)\) \((53)\).

**DISCUSSION**

Our analyses led to several important findings regarding RA in African Americans. First, SNPs tagging HLA-DRB1 were significantly associated with RA susceptibility, but not radiographic severity. Second, AFF3, TNFSF11 and TNFSF18 (all previously validated loci for RA susceptibility) were associated suggestively with RA susceptibility \(\left(1.0 \times 10^{-4} < p < 3.1 \times 10^{-6}\right)\). Third, TNFRSF9, CTLA4, IL2RA, C5/TRAF1 and CXCR5 were associated suggestively with radiographic severity. Finally, leveraging the differential LD pattern between Europeans and African Americans, we defined suggestive novel lead SNPs for the associations of AFF3 with susceptibility and IL2RA with severity.

As expected from previous studies and our prior work \((18)\), we found that the strongest association with RA susceptibility lies in the MHC region near HLA-DRB1. When the SNPs in the extended

**Table 3.** Ninety percent credible set for pathogenic variants in trans-ethnic fine mapping of AFF3 locus in RA susceptibility.

<table>
<thead>
<tr>
<th>rsID</th>
<th>Chr</th>
<th>Position</th>
<th>Effect allele</th>
<th>Alternate allele</th>
<th>Z AA</th>
<th>Z EAS</th>
<th>Z EUR</th>
<th>Posterior probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs13003982</td>
<td>chr2</td>
<td>100759078</td>
<td>T</td>
<td>C</td>
<td>-3.80</td>
<td>-4.09</td>
<td>-5.65</td>
<td>0.021</td>
</tr>
<tr>
<td>rs11681966</td>
<td>chr2</td>
<td>100759457</td>
<td>A</td>
<td>C</td>
<td>-4.22</td>
<td>-4.09</td>
<td>-5.65</td>
<td>0.105</td>
</tr>
<tr>
<td>rs12712067</td>
<td>chr2</td>
<td>100763900</td>
<td>T</td>
<td>G</td>
<td>-3.82</td>
<td>-4.08</td>
<td>-5.69</td>
<td>0.029</td>
</tr>
<tr>
<td>rs4851257</td>
<td>chr2</td>
<td>100775297</td>
<td>T</td>
<td>C</td>
<td>-3.86</td>
<td>-4.14</td>
<td>-5.59</td>
<td>0.025</td>
</tr>
<tr>
<td>rs4851258</td>
<td>chr2</td>
<td>100780830</td>
<td>T</td>
<td>C</td>
<td>-3.68</td>
<td>-4.17</td>
<td>-5.60</td>
<td>0.053</td>
</tr>
<tr>
<td>rs4851261</td>
<td>chr2</td>
<td>100786717</td>
<td>A</td>
<td>G</td>
<td>-3.69</td>
<td>-4.17</td>
<td>-5.60</td>
<td>0.024</td>
</tr>
<tr>
<td>rs10185059</td>
<td>chr2</td>
<td>100790172</td>
<td>T</td>
<td>C</td>
<td>-3.70</td>
<td>-4.16</td>
<td>-5.63</td>
<td>0.027</td>
</tr>
<tr>
<td>rs10185510</td>
<td>chr2</td>
<td>100790581</td>
<td>T</td>
<td>C</td>
<td>-3.69</td>
<td>-4.16</td>
<td>-5.63</td>
<td>0.027</td>
</tr>
<tr>
<td>rs12712071</td>
<td>chr2</td>
<td>100793876</td>
<td>A</td>
<td>G</td>
<td>-3.61</td>
<td>-4.17</td>
<td>-5.69</td>
<td>0.028</td>
</tr>
<tr>
<td>rs6712515</td>
<td>chr2</td>
<td>100806514</td>
<td>T</td>
<td>G</td>
<td>-2.31</td>
<td>-3.66</td>
<td>-6.91</td>
<td>0.119</td>
</tr>
<tr>
<td>rs9653442</td>
<td>chr2</td>
<td>100825367</td>
<td>T</td>
<td>C</td>
<td>-2.43</td>
<td>-3.50</td>
<td>-6.95</td>
<td>0.188</td>
</tr>
</tbody>
</table>

Chr: chromosome.

\(\text{a} \) Z-score from our study of African Americans.

\(\text{b} \) Z-score for East Asians from the trans-ethnic meta-analysis of Okada et al.

\(\text{c} \) Z-score for Europeans from Okada et al.

\(\text{d} \) Denotes that the variant was the index variant reported in this study, or in another genome-wide association study.

surrounding region, including RBM17 \((51,52)\). In this locus, we found shorter haplotype blocks and lower LD between genetic variants, so we sought to localize an association signal in this locus. We found that the most strongly associated SNPs in our dataset are in the first intron of RBM17, specifically in a ~5kb section of the genome displaying the H3K27Ac histone marks and DNAse hypersensitivity (Figure 4).

![Figure 3](image-url)
MHC were conditioned on the classical HLA-DRB1 alleles, the association signal elsewhere in the MHC region is lost. This finding illustrates that the genome-wide significant SNPs in our study are tagging HLA-DRB1 classical alleles. This was first noted in Europeans (46), but in that study residual significant association signal remained near HLA-B and HLA-DPB1 after the conditioning analysis. We observed residual signals near HLA-B and HLA-DPB1 having the same effect size and direction of effect, but they were not significantly associated with RA. Considering the consistency of effect size, this likely reflects statistical power, but could reflect biological differences as well.

Despite our study being well powered to detect an effect of similar magnitude, we found no association between MHC region SNPs and radiographic severity comparable to other reports. Viatte et al. reported an association between haplotypes defined by amino acid residues at positions 11, 71 and 74 of HLA-DRB1 and radiographic damage (15). There are several possible explanations for this discrepancy. First, the differences could result from cohort inclusion criteria. Viatte et al. included autoantibody-positive RA, autoantibody-negative RA and inflammatory polyarthritis (not meeting ACR criteria for classification of RA), while our study focused exclusively on autoantibody-positive RA. Second, because they did not stratify based on autoantibody positivity, it is possible that their findings reflect the known association between radiographic severity and autoantibody positivity. Finally, biological differences between ethnicities cannot be ruled out.

**AFF3** encodes LAF4, a transcriptional activator with suspected roles in lymphoid tissue development and onconeogenesis (54). The locus has been associated with RA susceptibility in Europeans (4,10) as well as SLE and juvenile idiopathic arthritis (15,56). rs9653442 in particular has been the subject of several investigations as an autoimmune risk variant, and it was the index variant for RA risk in a trans-ethnic meta-analysis (10). In this study, it was found in the 90% credible set for pathogenic variants. However, the **AFF3** locus has been identified as containing multiple independent effects for common complex diseases (50). Consistent with this, our results further suggest several promising candidate pathogenic variants in addition to rs9653442. rs6712515 is an index variant reported in the NHGRI GWAS catalog (49), but has previously been associated with cognitive phenotypes rather than autoimmunity. These two variants as well as rs7608424 are known to be expression quantitative trait loci for **AFF3** expression (57). The index variant in our study, rs11681966, is found only ~400 bases from the transcription start site of **AFF3** in a conserved region capable of binding numerous transcription factors. Another variant in tight linkage with rs11681966, rs13003982, is located only ~40 bp from the transcription start site of **AFF3**. Thus, our data not only suggest candidate pathogenic variants, but an initial finding for functional studies of the contribution of RA genetic variants to **AFF3** to test.

We also detected several suggestive associations with RA radiographic severity. **CTLA4** is associated with RA in several populations (10,58), and the importance of CD28/CTLA4 co-stimulation in RA is highlighted by the efficacy of CTLA4Ig (abatacept) (59). We detected a suggestive association of **TNFRSF9** with RA radiographic severity. **TNFRSF9** (CD137) is a member of the TNF receptor family known for its role in T cell co-stimulation. In RA, a soluble form of CD137 is released by activated lymphocytes and is present in the serum (60). In collagen-induced arthritic mice, treatment with an anti-CD137 antibody protects against disease progression, possibly by amplifying antigen-specific CD11c+ / CD8+ T lymphocytes and suppressing the pathogenic CD4+ T lymphocyte subset (61). While rs506746 (chr13:101981771, near **NALCN** and **ITGBL1**) showed the strongest association with radiographic severity, this finding should be interpreted cautiously because of the low coverage of this region on the iChip array.

We subjected two loci to additional analyses based on context provided by prior studies. We observed a suggestive association between RA severity and rs7077067, a variant near **IL2RA**. This locus has previously been linked to RA susceptibility (10), radiographic severity (31,51) and decreased likelihood of disease remission (62). Interleukin 2 receptor a (**IL2RA** or **CD25**) gene, together

---

**Table 4.** Variants genotyped on the iChip associated with RA radiographic severity in African Americans. Zero and count refer to the model coefficients for the portions of the zero inflated negative binomial model.

<table>
<thead>
<tr>
<th>rsID</th>
<th>Chr</th>
<th>Position</th>
<th>A1</th>
<th>Statistic</th>
<th>Effect size</th>
<th>95% CI</th>
<th>P value</th>
<th>Gene locus</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs228702</td>
<td>1</td>
<td>7945520</td>
<td>G</td>
<td>IRR</td>
<td>0.58</td>
<td>0.44-0.76</td>
<td>6.33 × 10^-5</td>
<td>TNFRSF9^p</td>
</tr>
<tr>
<td>rs13014054</td>
<td>2</td>
<td>33678924</td>
<td>A</td>
<td>IRR</td>
<td>0.53</td>
<td>0.39-0.71</td>
<td>2.92 × 10^-5</td>
<td>RASGRP3^p</td>
</tr>
<tr>
<td>rs73055463</td>
<td>2</td>
<td>204712807</td>
<td>C</td>
<td>IRR</td>
<td>1.99</td>
<td>1.38-2.86</td>
<td>4.72 × 10^-5</td>
<td>CTLA4^p</td>
</tr>
<tr>
<td>rs7304499</td>
<td>9</td>
<td>123687231</td>
<td>C</td>
<td>IRR</td>
<td>2.27</td>
<td>1.49-3.46</td>
<td>9.60 × 10^-5</td>
<td>TRAF1-C5^p</td>
</tr>
<tr>
<td>rs7077067</td>
<td>10</td>
<td>6132692</td>
<td>A</td>
<td>IRR</td>
<td>1.48</td>
<td>1.22-1.78</td>
<td>5.16 × 10^-5</td>
<td>IL2RA^p</td>
</tr>
<tr>
<td>rs7101785</td>
<td>11</td>
<td>696437</td>
<td>G</td>
<td>IRR</td>
<td>1.52</td>
<td>1.24-1.86</td>
<td>5.17 × 10^-5</td>
<td>PHLD81/CXCR5</td>
</tr>
<tr>
<td>rs7127742</td>
<td>11</td>
<td>118521637</td>
<td>G</td>
<td>IRR</td>
<td>0.40</td>
<td>0.25-0.62</td>
<td>5.66 × 10^-5</td>
<td>PHLDB1/CXCR5</td>
</tr>
<tr>
<td>rs4362159</td>
<td>11</td>
<td>128305571</td>
<td>A</td>
<td>IRR</td>
<td>0.48</td>
<td>0.34-0.69</td>
<td>6.26 × 10^-5</td>
<td>ETS^p</td>
</tr>
<tr>
<td>rs506746</td>
<td>13</td>
<td>101981771</td>
<td>A</td>
<td>IRR</td>
<td>0.53</td>
<td>0.41-0.68</td>
<td>4.19 × 10^-7</td>
<td>NALCN/ITGBL1</td>
</tr>
<tr>
<td>rs7193451</td>
<td>16</td>
<td>11050356</td>
<td>G</td>
<td>IRR</td>
<td>1.65</td>
<td>1.30-2.09</td>
<td>7.09 × 10^-5</td>
<td>CLEC16A</td>
</tr>
</tbody>
</table>

Chr: chromosome; IRR: incident rate ratio. aIndicates a validated risk locus for RA.
IL2RA locus may preclude tagging the same pathogenic variant.

We also found an association of radiographic severity of RA with ETS1, a highly conserved transcription factor whose expression in B cells, T cells and natural killer cells strongly affects immune cell function. Ets1 knockout mice display aberrant T cell differentiation, altered cytokine expression and increased differentiation into memory and effector T cells (65). Downregulation of Ets1 increases formation of plasma cells in part by upregulating Pax-5 and inhibiting Blimp1 activity (66). In humans, lupus risk alleles are associated with lower ETS1 mRNA expression, and the genetic basis of these findings differs in an ethnic-specific fashion (67). Specifically, increased binding of pSTAT1 to oligonucleotides containing the rs6590330 risk allele correlates with decreased ETS1 expression in Asian SLE patients, but not in other populations, including African Americans (67). Consistent with this finding, we found that rs7108537, but not rs6590330, was associated with radiographic severity. There are transcription factor binding sites in the immediate vicinity of rs7108537, and the genotype of this SNP appears to affect ETS1 expression in persons of Yoruban ancestry, but not in other populations (68). Therefore, our study of radiographic severity provides additional evidence that population-specific variants may contribute to risk of autoimmunity by decreasing ETS1 expression.

Our association testing results should be interpreted cautiously, as the sample size may result in inflated effect size estimates. In addition, our study was not well powered to detect association of common SNPs (MAF 0.15-0.50) with effect sizes <1.3. Nevertheless, we used the largest registry of African Americans with RA for whom clinical, radiographic and genetic data are available. We were unable to attempt to replicate our findings, because no other cohorts of African Americans with RA are available.

Finally, it should be noted that the markers

**Figure 4.** (Top) Zoom plot indicating the strength of association in the region of chromosome 10 surrounding the IL2RA locus. Circles represent single nucleotide polymorphisms (SNPs); the y-axis measures negative log association p value and the x-axis represents genomic position. (Middle) Diagram indicating histone markings and gene diagrams corresponding to the genomic region in the zoom plot. (Bottom) LD heatmap for the region.
selected during the design of the custom iChip array were derived from the 1000 Genomes project from European individuals, which might be suboptimal for analysis of disease-associated variants in African Americans, and thus additional novel risk variants RA may yet exist in this ethnic group. Limitations of the fine mapping study include exclusion of some genotyped variants due to absence from the reference dataset used for LD and inability to confirm uniform alignment of some variants to reference genomes based on LD and Z-score information. This may lead to the exclusion of potentially interesting variants. For instance, rs11676922, an RA index SNP previously studied in a meta-analysis of RA in Han Chinese and Europeans (69), was not examined in this study. This SNP is in near-perfect LD with rs9653442 as well as rs6712515. As such, investigators who wish to carry out functional studies on variants in AFF3 should note that including this variant might alter the posterior probabilities attributed to other variants.

CONCLUSION

In contrast to other reports, we find that SNPs in the MHC region do not appear to be associated with radiographic severity of RA in African Americans. Our study also demonstrates the utility of ethnic-specific analysis of genetic data. We confirm the association of AFF3 with RA susceptibility and IL2RA and ETS1 with radiographic severity, and our analysis of these loci suggests several candidate variants for functional validation. Our analysis of the AFF3 locus suggests that rs11681966, rs9653442, rs7608424 and rs6712515 are high-priority targets for functional studies, but may exert effects in population-specific contexts. Our data add to evidence that ETS1 autoimmune risk is mediated by ethnic-specific variants decreasing expression. In the IL2RA locus, our data suggest that trans-ethnic fine mapping studies could be valuable for RA susceptibility and radiographic severity due to different LD patterns. Overall, our study suggests that trans-ethnic genetic analysis is likely to be an important step in bringing precision medicine to complex autoimmune diseases, including RA.

ACKNOWLEDGMENTS

We thank the CLEAR investigators for recruiting participants with rheumatoid arthritis and healthy controls: Doyt Conn, MD (Grady Hospital and Emory University, Atlanta, Georgia); Beth Jonas, MD; Leigh Callahan, PhD (University of North Carolina, Chapel Hill); Edwin Smith, MD (Medical University of South Carolina, Charleston); Richard Brasington, MD (Washington University, St. Louis, Missouri); and Larry W Moreland, MD (University of Pittsburgh). We thank Drs. Robert Kimberly and Jeffrey Edberg for providing data on healthy controls from the Birmingham, Alabama, area. We would also like to thank Gleb Kichaev (David Geffen School of Medicine, University of California, Los Angeles) for helpful discussions regarding trans-ethnic fine mapping.

DISCLOSURES

The authors declare they have no competing interests as defined by Molecular Medicine, or other interests that might be perceived to influence the results and discussion reported in this paper.

REFERENCES

21. Weidinger S, Baurecht H, Naumann A, Novak N. (2010) Genome-wide association studies on IgE regulation: are genetics of IgE also genetics of


