Most Common Single-Nucleotide Polymorphisms Associated With Rheumatoid Arthritis in Persons of European Ancestry Confer Risk of Rheumatoid Arthritis in African Americans

Laura B. Hughes,1 Richard J. Reynolds,1 Elizabeth E. Brown,1 James M. Kelley,1 Brian Thomson,2 Doyt L. Conn,3 Beth L. Jonas,4 Andrew O. Westfall,1 Miguel A. Padilla,5 Leigh F. Callahan,4 Edwin A. Smith,6 Richard D. Brasington,7 Jeffrey C. Edberg,1 Robert P. Kimberly,1 Larry W. Moreland,8 Robert M. Plenge,2 and S. Louis Bridges, Jr.1

Objective. Large-scale genetic association studies have identified >20 rheumatoid arthritis (RA) risk alleles among individuals of European ancestry. The influence of these risk alleles has not been comprehensively studied in African Americans. We therefore sought to examine whether these validated RA risk alleles are associated with RA risk in an African American population.

Methods. Twenty-seven candidate single-nucleotide polymorphisms (SNPs) were genotyped in 556 autoantibody-positive African Americans with RA and 791 healthy African American control subjects. Odds ratios (ORs) and 95% confidence intervals (95% CIs) for each SNP were compared with previously published ORs for RA patients of European ancestry. We then calculated a composite genetic risk score (GRS) for each individual based on the sum of all risk alleles.

Results. Overlap of the ORs and 95% CIs between the European and African American populations was observed for 24 of the 27 candidate SNPs. Conversely, 3 of the 27 SNPs (CCR6 rs3093023, TAGAP rs394581, and TNFAIP3 rs6920220) demonstrated ORs in the opposite direction from those reported for RA patients of European ancestry. The GRS analysis indicated a small but highly significant probability that African American patients relative to control subjects were enriched for the risk alleles validated in European RA patients (P = 0.00005).

Conclusion. The majority of RA risk alleles previously validated for RA patients of European ancestry showed similar ORs in our population of African Americans with RA. Furthermore, the aggregate GRS supports the hypothesis that these SNPs are risk alleles for RA in the African American population. Future large-scale genetic studies are needed to validate these risk alleles and identify novel RA risk alleles in African Americans.

Rheumatoid arthritis (RA) is a phenotypically heterogeneous, systemic autoimmune disease characterized by chronic destructive inflammation in synovial joints. The disease can be subdivided into 2 groups...
patent (with overlapping 95% confidence intervals [95% odds ratios (ORs) for individual risk alleles are consistent between 2 methodologic approaches. We first tested whether the risk loci identified in populations of European ancestry, the question about whether these alleles are associated with RA risk in other ethnic groups remains unaddressed. We sought to study the association of these previously identified RA risk loci in a large group of well-characterized African American patients with RA.

Specifically, we hypothesized that many of the risk loci identified in populations of European ancestry will also demonstrate risk for RA in African Americans. Most RA risk alleles outside of the major histocompatibility complex region have moderate effect sizes; thus, large sample sizes have been required for identification and replication. We anticipated that it would be difficult to demonstrate strong statistical support for individual risk alleles in our African American population of 556 autoantibody-positive patients with RA and 791 healthy control subjects. To address this limitation, we used 2 methodologic approaches. We first tested whether the odds ratios (ORs) for individual risk alleles are consistent (with overlapping 95% confidence intervals [95% CIs]) or inconsistent (with 95% CIs that do not overlap) between the European and African American populations. As a second step, we derived an aggregate genetic risk score (GRS) in our population of African American patients with RA and control subjects and analyzed whether the cumulative effects of the validated risk alleles for RA in Europeans also confer risk for RA in African Americans (4).

PATIENTS AND METHODS

Study subjects. We analyzed 27 single-nucleotide polymorphisms (SNPs) in 556 African Americans with autoantibody-positive RA, defined as RF-positive serum or anti-CCP antibody–positive serum. The analysis was limited to autoantibody-positive individuals, because the risk alleles tested here were those previously validated in autoantibody-positive patients of European ancestry (2,3). All patients with RA were participants in the Consortium for the Longitudinal Evaluation of African Americans with Early Rheumatoid Arthritis (CLEAR) registry. The CLEAR registry enrolls self-identified African Americans who meet the American College of Rheumatology (formerly, the American Rheumatism Association) 1987 revised criteria for the classification of RA (8). CLEAR participants are recruited from the University of Alabama at Birmingham (coordinating site), Emory University/Grady Hospital, the University of North Carolina at Chapel Hill, the Medical University of South Carolina, and Washington University.

The CLEAR registry has 2 arms: CLEAR I (enrollment from 2000 to 2005), a longitudinal arm for patients with early RA (duration <2 years) who are followed up longitudinally until 5 years of disease duration, and CLEAR II (enrollment began in 2006 and is ongoing), a cross-sectional arm in which patients with any disease duration (typically longstanding) are seen at one time point. In the current study, we...
analyzed only the autoantibody-positive participants: 228 patients with RA from CLEAR I and 328 patients with RA from CLEAR II. Detailed demographic and clinical data and DNA samples are available for CLEAR participants (see http://medicine.uab.edu/rheum/70918/ for details).

Healthy African American control subjects whose sex, age, and geographic location were similar to those of the patients with RA were recruited through the CLEAR registry. This group comprised 132 control subjects from CLEAR I and 171 control subjects from CLEAR II. The remaining 501 African American control subjects were recruited from the Birmingham, Alabama area. All participants provided informed consent under the approval of each respective institutional review board. Anti-CCP antibodies and RF were determined using previously described methods (9).

**Genotyping.** SNP selection was based on previously identified and validated risk alleles in autoantibody-positive RA patients of European ancestry (5,6,10–22) as well as those identified in a recent genome-wide association study (GWAS) meta-analysis by Stahl et al (7), which included >5,000 autoantibody-positive European patients with RA and 20,000 control subjects (Table 2).

SNP genotyping was performed at the Broad Institute using the Sequenom iPLEX assay, as previously described (15). For quality control, we required that each SNP meet the following criteria: a missing-genotype rate of <10%, a minor allele frequency (MAF) of >1%, and Hardy-Weinberg equilibrium, with $P > 0.001$. We then excluded individuals for whom data were missing for >10% of the SNPs passing quality control.

**Statistical analysis.** We performed SNP associations with RA risk using the software package Plink v1.06 (http://pngu.mgh.harvard.edu/purcell/plink) (23). ORs reflected the differences between patients and control subjects for MAFs in the African American population and were tested with Pearson’s chi-square test. The analysis included 556 autoantibody-positive patients with RA and 791 control subjects, after filtering was done for quality control. We did not perform a Bonferroni adjustment, because each of the SNPs tested is a validated risk allele for RA in the European population. Had we used a Bonferroni adjustment for multiple comparisons, we would not have declared statistical significance unless the $P$ value was less than the nominal alpha level of 0.002 (i.e., 0.05 divided by 27).

We compared ORs for RA patients of European ancestry with those for our population of African American patients with RA. The ORs and 95% CIs for European patients were derived from a meta-analysis of 5,539 autoantibody-positive patients with RA and 20,169 control subjects of European descent (7). Differences in ORs between the European and African American populations were considered to be significant if the 95% CIs of the ORs did not overlap. The statistical power to detect a significant association of SNPs with RA in African Americans was calculated using the Genetic Power Calculator (24). These power calculations were based on the ORs for association with RA in the European population, the sample size of African Americans, and the MAFs in the African American population.

As a final step, we derived an aggregate GRS in our population of autoantibody-positive African American patients with RA and control subjects (4). For each subject, the sum (or count) of the number of risk alleles across the 27 SNPs was calculated (possible range 0–54). We chose a GRS count approach rather than a weighted GRS (in which each SNP in the GRS is weighted by the OR), because we did not want to assume that the OR for African Americans was the same as the OR for individuals of European ancestry. The total number of risk versus non-risk alleles between patients and control subjects was then assessed using contingency table analysis, and the hypothesis that the type of allele was independent of case or control status was assessed with Pearson’s chi-square statistic. Equivalently, the GRS score (probability that the individual had risk alleles) was modeled as a binomial response with case–control status as a dichotomous predictor, using generalized linear models. Both approaches were implemented in the statistical package R (25).

## RESULTS

The characteristics of the patients with RA who are participants in CLEAR I and CLEAR II are shown in Table 1. The demographic characteristics of the 2 groups were similar, with the exception of values for the mean age at RA onset and disease duration, which were significantly different in CLEAR II participants than in CLEAR I participants because of differences in inclusion criteria for the 2 CLEAR arms.

The allele frequencies of each of the candidate SNPs tested for association with RA risk in our African American cohort are shown in Table 2. Notably, some RA risk alleles were present at a substantially lower frequency in African Americans compared with Europeans (e.g., ANKRD55, PTPN22, IRF5, IL2/IL21). We previously reported the allele frequencies and ORs for 5 SNPs in the STAT4 region (rs11889341, rs10931481, rs7574865, rs8179673, and rs10181656) in 723 African American patients with RA (including the 556 autoantibody-positive patients reported here, plus autoantibody-negative patients) and 660 of the control subjects (26). The MAFs for African American patients with RA and control subjects for the STAT4 rs11889341 T allele were determined to be 0.141 and 0.133, respectively, using ABI TaqMan assays. These values are consistent with results obtained using the Sequenom platform, as in the current study.

We compared the ORs and 95% CIs for our African American population with those reported in a recently published meta-analysis of RA patients of European ancestry (7). Of the 27 SNPs tested, 24 demonstrated ORs and 95% CIs that overlapped between the European and African American populations (Table 2 and Figure 1). Only 1 of the 24 SNPs, CTLA4 rs31087243, showed a statistically significant association with RA in African Americans (OR 0.75 [95% CI 0.62–0.91], $P = 0.003$). Seven of the 24 SNPs (TNFAIP3 rs10499194,
CD58 rs11586238, PTPRC rs10919563, PKX rs13315591, \textit{PRKCQ} rs4750316, \textit{IL2/IL21} rs6822844, and \textit{CD40} rs4810485) demonstrated ORs in the direction opposite to that seen in RA patients of European ancestry, but the 95% CIs overlapped between the 2 populations, and there was no statistically significant difference between African American patients with RA and control subjects.

As shown in Table 2 (bottom 3 rows) and Figure 1, evidence for differences in the ORs between the 2 ethnic groups was observed for 3 of the 27 SNPs (\textit{CCR6} rs3093023, \textit{TAGAP} rs394581, and \textit{TNFAIP3} rs6920220). Of these 3 SNPs, only \textit{CCR6} rs3093023 showed a statistically significant association with RA in African Americans (OR 0.79 [95% CI 0.64–0.98], \( P = 0.035 \)). All 3 of these risk alleles had ORs that were in the opposite directions between the African American and European populations.

To overcome the limitation in statistical power due to the relatively small size of our cohort of African American patients with RA, we calculated a GRS based on all RA risk alleles for each person in the study. This approach combines the effect of each risk allele into a single aggregate score, which can then be used to test whether the GRS differentiates between patients with RA and control subjects. The GRS analysis demonstrated that the probability of having risk alleles was higher in patients (mean GRS = 0.43) than in control subjects (mean GRS = 0.41; \( P = 0.00005 \)) (Figure 2). To exclude the possibility that the observed difference was due to a single SNP with a large effect size, we repeated the GRS analysis without the \textit{CTLA4} rs3087243 SNP.
the result remained significant (for patients, GRS/H11005 0.41; for control subjects, GRS/H11005 0.40 \[P/H11005 0.0003\]). This observation supports the hypothesis that many risk alleles observed in European patients with RA are also found in African American patients with RA.

DISCUSSION

To date, the majority of GWAS and subsequent meta-analyses of GWAS data in RA have focused on individuals of European and East Asian ancestry. It has become clear from these and other large-scale genetic studies of complex diseases such as RA that genetic risk loci can differ between these different ethnic groups (18,22,27). There is a paucity of data from large, well-characterized groups of African American patients with RA. In this study, we sought to test the hypothesis that RA risk alleles validated in populations of patients of European ancestry would also be associated with RA in African Americans. The power to test individual risk alleles in our African American sample was limited compared with that in studies in European patients with RA. The results of this study, however, demonstrated (via the GRS analysis) that, cumulatively, risk alleles for RA in Europeans also confer risk for RA in African Americans. Therefore, we conclude that the 2 populations are best characterized as being genetically homogeneous with respect to validated risk alleles for RA.

We observed that the ORs between the European and African American populations were consistent for 24 of the 27 risk loci validated in Europeans with RA. One interpretation of this finding is that the genetic etiology of RA risk in the 2 populations is very similar. However, only 1 of the validated RA risk alleles in Europeans achieved statistically significant association with the risk of RA in the African American population (CTLA4 rs3087243). Although the CLEAR cohort is the largest group of African American patients with RA currently available for analysis, the current study had limited power to detect genetic associations. This limited power, as shown in Table 2, can be attributed to the small effect size of many individual risk alleles and the low frequency of some risk alleles in the African American population. The limited power to detect association signals also affected our ability to demonstrate between-population inconsistencies for the ORs. For example, the low-frequency PTPN22 SNP rs2476601 (MAF = 0.02) showed the largest between-population difference in ORs between the 2 populations.
(Figure 1), with ORs of 1.23 in African Americans and 1.94 in Europeans. However, due to the large 95% CIs for the OR for African Americans, we were unable to conclude that these ORs were different between populations (Table 2).

We observed differences in the direction of the ORs between European and African American populations for 3 loci (CCR6 rs3093023, TAGAP rs394581, and TNFAIP3 rs6920220). There are at least 3 explanations for these differences. First, inconsistent ORs might be attributable to weak correlations between tag SNPs and the causal allele (which has not yet been identified) with actual risk of RA. It is known from the International HapMap Project that the linkage disequilibrium structure between European and African American populations may be different for common SNPs in any given locus. Such an effect may be particularly striking if the causal allele is rare in the population (e.g., frequency <5%) (28). The second possible explanation is that these differences may be explained by genetic heterogeneity; a risk allele in European patients with RA may operate differently in African Americans in terms of its effect on RA risk. Such population-specific differences may reflect gene–gene or gene–environment interactions. The third possible explanation is that the inconsistency may simply be attributable to chance, given the number of hypotheses tested. Further investigation of these alleles among individuals of African ancestry is needed to explore these possibilities.

Although this study had limited statistical power because of the small sample size, when the SNP risk alleles were viewed together in the GRS analysis, there was evidence that the risk alleles in Europeans are also risk alleles in African Americans. The GRS was shown to be significantly different between patients with RA and control subjects, even after excluding the CTLA4 SNP with the largest effect size. However, this approach indicates that even though most of the individual markers were not statistically associated with RA in African Americans, they may contribute to a panel of alleles that collectively confer risk.

Future studies of large, well-characterized cohorts of African American patients with RA and control subjects are needed to definitively determine whether the RA risk alleles in Europeans are associated with RA risk in African Americans. Large-scale GWAS in African American patients with RA are also needed in order to explore novel risk alleles among this genetically admixed ethnic group. We believe that detailed genetic studies of African Americans with RA will lead to important insights into the pathogenesis of this disease.

ACKNOWLEDGMENTS

We gratefully acknowledge all of the study patients for their contributions to this work. We also thank Drs. David Allison, Hemant Tiwari, Maria Danila, Monica Crawford, and Jeffrey Faggard for helpful discussions and review of the manuscript.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Bridges, Jr. had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Hughes, Brown, Jonas, Smith, Brasington, Edberg, Moreland, Plenge, Bridges, Jr.

Acquisition of data. Hughes, Jonas, Callahan, Smith, Brasington, Edberg, Kimberly, Moreland, Plenge, Bridges, Jr.

Analysis and interpretation of data. Hughes, Reynolds, Brown, Kelley, Thomson, Conn, Westfall, Padilla, Smith, Brasington, Plenge, Bridges, Jr.

REFERENCES


