Published in final edited form as:

Hum Genet. 2011 December ; 130(6): 807–815. doi:10.1007/s00439-011-1045-2.

MHC region and risk of systemic lupus erythematosus in African-American women

Edward A. Ruiz-Narvaez^{1,2}, Patricia A. Fraser³, Julie R. Palmer^{1,2}, L. Adrienne Cupples⁴, David Reich^{5,6}, Ying A. Wang⁴, John D. Rioux^{7,8}, and Lynn Rosenberg^{1,2}

¹Slone Epidemiology Center at Boston University, Boston, MA

²Department of Epidemiology, Boston University School of Public Health, Boston, MA

³Genzyme Corporation, Cambridge, MA

⁴Department of Biostatistics, Boston University School of Public Health, Boston, MA

⁵Department of Genetics, Harvard Medical School, Boston, MA

⁶Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA

⁷Research Center, Montreal Heart Institute, Montreal, Quebec, Canada

⁸Faculty of Medicine, University of Montreal, Quebec, Canada

Abstract

The major histocompatibility complex (MHC) on chromosome 6p21 is a key contributor to the genetic basis of systemic lupus erythematosus (SLE). Although SLE affects African Americans disproportionately compared to European Americans, there has been no comprehensive analysis of the MHC region in relationship to SLE in African Americans. We conducted a screening of the MHC region for 1,536 single nucleotide polymorphisms (SNPs) and the deletion of the *C4A* gene in a SLE case-control study (380 cases, 765 age-matched controls) nested within the prospective Black Women's Health Study. We also genotyped 1,509 ancestral informative markers throughout the genome to estimate European ancestry in order to control for population stratification due to population admixture. The most strongly associated SNP with SLE was the rs9271366 (odds ratio, OR = 1.70, $p = 5.6 \times 10^{-5}$) near the *HLA-DRB1* gene. Conditional haplotype analysis revealed three other SNPs, rs204890 (OR = 1.86, $p = 1.2 \times 10^{-4}$), rs2071349 (OR = 1.53, $p = 1.0 \times 10^{-3}$), and rs2844580 (OR = 1.43, $p = 1.3 \times 10^{-3}$) to be associated with SLE independent of the rs9271366 SNP. In univariate analysis, the OR for the *C4A* deletion was 1.38, $p = 0.075$, but after simultaneous adjustment for the other four SNPs the odds ratio was 1.01, $p = 0.98$. A genotype score combining the four newly identified SNPs showed an additive risk according to the number of high-risk alleles (OR = 1.67 per high-risk allele, $p < 0.0001$). Our strongest signal, the rs9271366 SNP, was also associated with higher risk of SLE in a previous Chinese genome-wide association study (GWAS). In addition, two SNPs found in a GWAS of European ancestry women were confirmed in our study, indicating that African Americans share some genetic risk factors for SLE with European and Chinese subjects. In summary, we found four independent signals in the MHC region associated with risk of SLE in African American women.

Keywords

systemic lupus erythematosus; African Americans; major histocompatibility complex; single nucleotide polymorphisms

INTRODUCTION

Systemic lupus erythematosus (SLE) is a debilitating chronic autoimmune disease characterized by autoantibody production and the involvement of multiple organs. The existence of a heritable basis for SLE has long been recognized; the risk of SLE to siblings of an SLE proband (sibling risk ratio, λ_S) has been estimated to be as high as 29-fold the risk of the background population (Alarcon-Segovia et al. 2005), and monozygotic twins show a higher concordance rate (24%) than dizygotic twins (2%) (Deapen et al. 1992).

To date, the strongest associations of genetic factors with SLE are with variants located in the major histocompatibility complex (MHC) on the short arm of chromosome 6 (6p21.3) that contains hundred of genes, many of them with immune-related functions (The MHC sequencing consortium 1999; Campbell and Trowsdale 1993; Horton et al. 2004). Before the advent of genome-wide association studies (GWAS), serologic types of the human leukocyte antigen (*HLA*) genes were found to be associated with risk of SLE. In particular, the *HLA* class II *HLA-DRB1*1501* (serologic type *HLA-DR2*) and *HLA-DRB*0301* (serologic type *HLA-DR3*) alleles have been consistently found to be associated with risk of SLE in populations of European ancestry and these alleles confer up to a three-fold increase in the risk of SLE (Criswell 2008; Harley et al. 1998; Tan and Arnett 1998).

Variations of complement genes within the MHC region, have also been associated with risk of SLE. Numerous studies have reported that the null allele of the *C4* gene (*C4A*Q0*) confers higher risk of SLE (Moser et al. 2009; Walport 1993). It is the strongest genetic risk factor for SLE shared by African Americans, European Americans, and northern Europeans. The association with SLE in African Americans is most specific for a deletion of the *C4A* gene that extends from the 5' region of the *C4A* gene through the entire flanking *CYP21A2* gene and terminates in the 5' region of the *C4B* gene.

Recent GWAS (Graham et al. 2008; Han et al. 2009; Harley et al. 2008; Hom et al. 2008) and high-density screenings (Barcellos et al. 2009; Rioux et al. 2009b) have identified several independent single nucleotide polymorphisms (SNPs) in the MHC region associated with risk of SLE. However, those studies have been performed in European or East Asian ancestry populations. The relationship between genetic variants in the MHC region and risk of SLE in African American women, a population at high risk for SLE compared with European Americans (Helmick et al. 2008), is still unknown.

We conducted a screening of the MHC region in a nested case-control study in African American women to identify genetic variants associated with risk of SLE. In addition, we typed the deletion of the *C4A* gene to assess its combined effect with MHC genetic variants.

MATERIAL AND METHODS

Study population

We conducted a nested case-control study within the ongoing prospective Black Women's Health Study (BWHS) that has been described elsewhere (Palmer et al. 2003). Briefly, the study began in 1995 when women 21–69 years of age from across the United States completed a 14-page postal health questionnaire. The initial cohort comprises 59,000

women who self-identified as “black” and had a valid address. Follow-up questionnaires are sent every two years. Follow-up of the baseline cohort through six completed 2-year cycles to date is greater than 80%.

We obtained DNA samples from BWHS participants using the mouthwash-swish method (Cozier et al. 2004). Approximately 50% of participants, 27,200 women, provided a sample. Women who provided samples were slightly older than women who did not, but the two groups were similar with regard to educational level, geographic region of residence, smoking status, body mass index, and a wide variety of other factors.

We used medical records and information from the physician to confirm self-reported cases of SLE (McAlindon et al. 2003). Physicians were asked to complete a checklist of the American College of Rheumatology (ACR) criteria for SLE. We have defined “definite” SLE as cases with 4 or more ACR criteria; “probable” cases as having 3 criteria; and “clinical” cases as those confirmed by the physician as having SLE without listing criteria or for which SLE medications were being used and the woman did not report discoid lupus or rheumatoid arthritis. The present study includes 400 validated SLE cases (220 definite, 49 probable, 131 clinical). We matched each case with 2 controls randomly selected from among BWHS participants who did not report SLE and who had provided saliva samples; matching factors were birth year (± 1 year) and region of residence in the U.S. (Northeast, South, Midwest, West). We matched a total of 400 cases with 800 controls.

Selection of MHC SNPs and ancestral informative markers

We used the Illumina GoldenGate MHC panel developed by the International MHC and Autoimmunity Genetics Network (IMAGEN) consortium (de Bakker et al. 2006; Rioux et al. 2009a). This panel consists of 1,536 SNPs that tag both classic HLA types and common genetic variation through the 3.44 Mb of the classic MHC region (de Bakker et al. 2006; Rioux et al. 2009a). Because this MHC panel was designed to tag common genetic variation in European ancestry populations, for the purpose of assessing African Americans 397 SNPs were replaced by SNPs that better tag genetic variation in the HapMap YRI population. We also assessed the 30 kb deletion of the *C4A* gene.

We used the phase 3 admixture panel to estimate and control for population stratification due to European admixture. The phase 3 panel consists of 1,509 ancestral informative markers (AIMs) that have high allele frequency differences between African and European continental populations. The admixture panel is based on original sets described by Smith et al. (Smith et al. 2004) and Reich et al. (Reich et al. 2005), and further improved by mining of AIMs from Hinds et al. (Hinds et al. 2005) and from the phase 2 International HapMap project (Frazer et al. 2007).

Genotyping methods and quality control

DNA was isolated from mouthwash swish samples provided by SLE cases and controls at the Boston University Molecular Core Genetics Laboratory using the QIAAMP DNA Mini Kit (Qiagen). Whole genome amplification was performed with the Qiagen RePLI-g Kits using the method of multiple displacement amplification. Amplified samples underwent purification and PicoGreen quantification at the Broad Institute Center for Genotyping and Analysis (Cambridge, MA) before being plated for genotyping.

Genotyping of the MHC panel and the phase 3 admixture panel was carried out at the Broad Institute Center for Genotyping and Analysis using Illumina GoldenGate technology. We included 98 blinded duplicate samples to assess reproducibility of the genotypes. An average reproducibility of 99% was obtained among the blinded duplicates. We excluded all SNPs with calling rate < 90%, a deviation from Hardy-Weinberg equilibrium in the control

samples at $p < 0.001$, or a minor allele frequency < 0.01 . We also excluded samples with calling rates $< 80\%$. Genotyping of the deletion of the *C4A* gene was performed in the laboratory of Dr. Fraser using long PCR with a specific primer design to detect the 30 kb deletion of most of the *C4A* gene (Grant et al. 2000). We genotyped by restriction fragment length polymorphism methodology (Fraser et al. 2000) a panel of DNA samples known to possess the *C4A* gene deletion as positive controls. Using these quality control criteria, we successfully genotyped 1,286 SNPs in the MHC region, 1,174 AIMs, and the *C4A* deletion in 380 SLE cases and 765 controls. The mean call rate in the final data set for both SNPs and samples was 99.5%.

Statistical analysis

We used PLINK (Purcell et al. 2007) version 1.06 to calculate summary statistics for the genotype data. We tested for association between MHC SNPs and SLE using the 1-df Cochran-Armitage trend test of an additive genetic model with 100,000 permutations to estimate empirical p-values and to correct for multiple testing; a MHC-wide significant result was declared at p-permuted < 0.05 .

We used the conditional haplotype method (Valdes et al. 1997; Valdes and Thomson 1997) to detect secondary signals that are independent of the MHC-wide significant SNPs. The conditional haplotype method stratifies by haplotypic background and tests the null hypothesis that one or more SNPs have no independent haplotypic effect once we condition for such background. We first selected SNPs with $p < 0.01$ based on univariate analysis; we then applied the conditional haplotype method conditioning on MHC-wide significant SNPs to test for independent effects of secondary SNPs. A significant secondary signal was declared if the p-value from the conditional haplotype method was < 0.001 in agreement with previous use of the method in the MHC region (Barcellos et al. 2009).

We used PROC LOGISTIC of the SAS statistical software version 9.1.3 (SAS Institute Inc., Cary, NC, USA) to estimate odds ratios (ORs) and 95% confidence intervals (95% CI) for MHC-wide significant SNPs as well as for secondary SNPs identified through the conditional haplotype method and the *C4A* deletion. First, we ran individual models for each SNP and the *C4A* deletion. Then, we ran a combined model that contained all the identified SNPs plus the *C4A* deletion to test whether they have independent effects on risk of SLE after simultaneously adjusting for each other genetic variant. We adjusted the ORs for age, geographical region of residence (Northeast, South, Midwest, West), place of birth (US, foreign country), and European admixture proportion. We estimated individual European admixture proportions using a Bayesian approach as implemented in the Admixmap software (Hoggart et al. 2003; McKeigue et al. 2000). We used previously published data (Freedman et al. 2006; Reich et al. 2005; Smith et al. 2004) and data from the International HapMap Project (Frazer et al. 2007) to estimate allele frequencies of each AIM in the parental populations.

We calculated a multi-loci genotypic score to estimate the joint effects of the identified polymorphisms that have independent effects. Each individual SNP was coded as 0, 1, and 2 based on the number of SLE alleles that increased risk; the genotype score consisted of the sum of SLE risk alleles of the identified SNPs. We used logistic regression to estimate ORs and 95% CI of the different categories of the genotype score, using as reference the category with the lowest number of SLE risk alleles.

RESULTS

Table 1 presents general characteristics of SLE cases and controls. SLE cases had a slightly lower percentage of European ancestry as compared to controls (19.7% vs. 20.9%, $p = 0.10$).

Figure 1 displays results of the single marker associations with SLE in the MHC region. We found SNP rs9271366 in the intergenic region between the *HLA-DRB1* and *HLA-DQA1* genes to be associated with SLE at an MHC-wide level (p for trend = 5.6×10^{-5} , p -permutated = 0.05). Conditional haplotype analysis revealed three secondary signals that were independent of the SNP rs9271366: the SNP rs2844580 (p -conditional = 0.0009) about 9 kb upstream of the *HLA-B* gene, the SNP rs204890 (p -conditional = 0.0007) inside the *ATF6B* gene, and the SNP rs2071349 (p -conditional = 0.0004) inside the *HLA-DPB1* gene. The deletion of the *C4A* gene was more frequent in SLE cases than controls (6.5% vs. 4.8%, p = 0.075). We found that the four newly identified SNPs remained significantly associated with risk of SLE even after simultaneous adjustment in a combined model that contained the four SNPs plus the *C4A* deletion. The deletion of the *C4A* gene was not associated with risk of SLE after simultaneous adjustment for the other four SNPs (OR = 1.01). A summary of the results and odds ratios are shown in Table 2.

Linkage disequilibrium values (D' and r^2) among the four SNPs and the deletion of the *C4A* gene, shown in Table 3, indicate that these variants are not in linkage disequilibrium with each other and the four newly identified SNPs most likely represent independent signals in the MHC region.

We explored the joint effect of the four newly identified SNPs on the risk of SLE by means of a genotype score based on the number of increasing risk alleles (Table 4). The *C4A* deletion was not included in the genotype score because the combined analysis showed no independent effect of the *C4A* deletion after adjustment for the other four SNPs. SLE cases had a higher average genotype score compared to controls (4.6 vs. 4.2 risk alleles respectively, p for differences of means < 0.0001). We found that risk of SLE increased in an additive way according to the number of risk alleles; for each additional risk allele, risk of SLE increased by 67%; OR (95% CI) = 1.67 (1.46–1.92). Subjects in the highest category of genotype score (≥ 6 risk alleles) had almost five times the risk of SLE as compared to subjects in the reference category (≤ 3 risk alleles) OR (95% CI) = 4.65 (2.82–7.69). No major difference in results were observed when we restricted our analysis to SLE cases diagnosed during follow up or within the five year period before baseline (N = 207), cases meeting at least 3 ACR criteria (N = 250), cases meeting at least 4 ACR criteria (N = 202), or to clinical subgroups with different manifestations of SLE (Table 4).

Replication of findings from other ethnic groups

We examined whether previously reported MHC SNPs associated with SLE are also associated with SLE in the BWHS. Reported MHC SNPs were taken from a GWAS by the International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEN) in women of European ancestry (Harley et al. 2008), a GWAS in a Chinese Han population (Han et al. 2009), and a high-density screening of the MHC region in European ancestry subjects (Barcellos et al. 2009) (Table 5). If the reported SNP was not part of our MHC panel, we determined the best proxy in the HapMap YRI population using the SNAP software (Johnson et al. 2008). The SLEGEN GWAS reported the rs3131379 and rs120942 SNPs (Harley et al. 2008), and the Chinese GWAS found 13 SNPs in the MHC region that represented two independent signals in the MHC region at rs9271100 and at rs3997854 (Han et al. 2009). It is noteworthy that in the Chinese GWAS one of the 13 SNPs found in the MHC region was the rs9271366 SNP that was the most significant finding in the present study. Both rs9271100 (the most significant MHC SNP in the Chinese GWAS) and rs9271366 (the most significant MHC SNP in the present study) are just 10 kb apart from each other and they are in the same haplotype block in both HapMap CHB and YRI populations. The high density MHC screening in European ancestry subjects reported 10 independent signals along the MHC region (Barcellos et al. 2009); we were able to find proxies for eight of the reported SNPs. We found that two SNPs from the SLEGEN GWAS,

rs3131379 and rs1270942 (tested by the perfect proxy rs440454), and two SNPs from the Chinese GWAS, rs9271100 (tested by the rs9271366, $r^2 = 0.73$) and rs3997854, were associated with increased risk of SLE in the BWHS. One of the SNPs from the Chinese GWAS, the C-allele of the rs3997854 SNP, which was significantly positively associated with SLE in the BWHS, was inversely associated in the Chinese sample. None of the SNPs or their proxies from the MHC high-density screening in European ancestry subjects were associated with risk of SLE in the BWHS.

We also assessed whether the *HLA-DRB1*0301* (*HLA-DR3*) and *HLA-DRB1*1501* (*HLA-DR2*) alleles, which have been associated with higher risk of SLE in European ancestry populations (Hartung et al. 1992; Yao et al. 1993), are also associated with risk of SLE in the BWHS. Although we did not determine HLA serotypes, our MHC panel included tagSNPs that capture variation of the classical HLA types (de Bakker et al. 2006). The haplotype rs3129763_A/rs2647012_T/rs3128968_T tagged the *HLA-DRB1*0301* allele ($r^2=0.71$, $D'=0.89$ in HapMap YRI population), and the rs2116264_G allele tagged the *HLA-DRB1*1501* allele ($r^2=1.0$, $D'=1.0$ in HapMap YRI population). We did not find significant associations of the *HLA-DRB1*0301* allele (6.4% in SLE cases vs. 6.4% in controls, p for difference = 0.99) or the *HLA-DRB1*1501* allele (1.7% in SLE cases vs. 2.8% in controls, p for difference = 0.11) with SLE in the BWHS.

DISCUSSION

In the present study we identified four independent SNPs (rs9271366, rs204890, rs2071349, and rs2844580) in the MHC region associated with risk of SLE in a population of African American women. The four identified SNPs extend over a region of 1.7 Mb and they point to different signals that may contribute to SLE risk in African American women.

SNPs rs9271366 (the most significant signal in the present study) and rs2071349 are both located inside the HLA class II region that has been consistently associated with SLE in European, East Asian and African American populations (Fernando et al. 2007; Han et al. 2009; Hom et al. 2008; Reveille et al. 1998; Rioux et al. 2009b; Uribe et al. 2004). It is noteworthy that a recent GWAS found the rs9271366 SNP to be associated with SLE in Chinese Han subjects at genome-wide significance (Han et al. 2009). The rs9271366 was part of a group of genome-wide significant SNPs near the *HLA-DRB1* gene that belong to the same haplotype block and seem to be pointing to the same causal signal (Han et al. 2009).

In populations of European ancestry, a major part of the signal from the HLA class II region comes from the presence of haplotypes carrying the *HLA-DRB1*0301* and *HLA-DRB1*1501* alleles (Graham et al. 2007; Graham et al. 2002). In the present study, both HLA class II alleles (as determined by tagging SNPs) were in low frequency and not associated with SLE, suggesting that the observed associations in the HLA class II region in African American women are not explained by classic class II alleles previously identified as risk factors in populations of European ancestry.

The rs204890 SNP is located in intron 6 of the *ATF6B* gene that codes for a transcription factor involved in the unfolded protein response pathway during endoplasmic reticulum stress. It is noteworthy that the rs8283 SNP identified in a previous high-density screening of the MHC region (Barcellos et al. 2009) is located in the 3'UTR region of the *ATF6B* gene. Although our best proxy for rs8283, SNP rs429150, $r^2=0.94$, was not associated with SLE in the present study, both rs204890 and rs8283 are located in the same LD block in the HapMap CEU population suggesting that a single causal variant may be responsible for the

rs204890 association in the present study and the rs8283 association in European ancestry subjects.

The rs2844580 SNP is located ~9 kb from the *HLA-B* gene and ~38 kb from the MHC class I chain-related A (*MICA*) gene. Genetic variation in both genes has been found to be associated with SLE in European ancestry populations (Gambelunghe et al. 2005; Price et al. 1999; Sanchez et al. 2006; Smerdel-Ramoya et al. 2005), the most consistent being the association with 8.1 ancestral haplotype (AH) (Candore et al. 2002; Price et al. 1999). The 8.1 AH consists of the *HLA-A*01*, *HLA-Cw7*, *HLA-B*08*, *MICA-5.1*, *DRB1*0301*, and *DQA1*05-DQB1*02* alleles (Ide et al. 2005) and has been associated with different immune diseases in European populations (see review in (Candore et al. 2002)). Because the 8.1 AH is specific to European populations, our findings show that other genetic variants in the *HLA-B* and *MICA* neighborhood may contribute to SLE risk.

Our combined analysis suggests that the four newly identified SNPs (rs9271366, rs204890, rs2071349, and rs2844580) are independent signals of SLE risk in African American women. This conclusion is supported by the low linkage disequilibrium values among the four newly identified SNPs. In an analysis summing the risk of the four newly identified SNPs into a genetic risk score, the risk of SLE increased in a linear way according to the number of risk alleles. It is noteworthy that the estimated OR was robust to the use of different sets of SLE cases, suggesting that our results are not being distorted by bias related to survival from SLE or to specificity of case definition.

Comparison with previously published results suggests that some genetic risk factors for SLE may be shared among African Americans and other ethnic groups such as East Asian and European populations. Our most significant result, the rs9271366, was also found to be associated with SLE risk in Chinese subjects (Han et al. 2009); the second SNP from the Chinese GWAS (rs3997854) was also associated with SLE risk in the BWHS although in an opposite direction: the C-allele was protective in Chinese subjects but harmful in the BWHS. Differences in allele frequencies and haplotype background may explain the opposite results for the rs3997854 SNP. The C-allele of the rs3997854 SNP has a frequency of 26% in the BWHS as compared to just 8% in Chinese Han subjects (Han et al. 2009), suggesting that different C-allele carrying haplotypes may be present in the BWHS and they may have different effects on risk of SLE. We also found that the two reported SNPs from the SLEGEN GWAS (Harley et al. 2008), rs3131379 and rs1270942 (tested with the perfect proxy rs440454), were also associated with increased risk of SLE in the present study. However, these two SNPs did not show independent effects in the BWHS after conditioning for the rs9271366 SNP. Although none of the SNPs from the high-density screening of the MHC region (Barcellos et al. 2009) was replicated in the present study, some of the genetic regions were shared with BWHS subjects. For example, the rs204890 SNP reported in the present study and the rs8283 from the high-density screening are located both in the *ATF6B* gene. Thus, genetic variation in the *ATF6B* gene seems to affect risk of SLE in African American and European ancestry subjects.

To our knowledge this is the first screening of the MHC region in relationship to SLE in a population-based cohort of African American subjects. Our case-control study was derived from a well-characterized cohort, adding to the strength of the reported results. Control for degree of European ancestry and the case-control matching by region of residence make it unlikely that our results are due to population stratification. The classification of cases of SLE was based on the judgment of the women's physicians as to the presence of ACR criteria for the illness. While these decisions may have varied across physicians, the fact that our most significant result (the rs9271366 SNP) was also reported in a previous GWAS in Chinese subjects and that we reproduced findings in significant SNPs from the SLEGEN

study supports the validity of the BWHS case definition and results. Our study had 80% statistical power to identify ORs of 1.4 or greater for SNPs with minor allele frequencies >10% at MHC-wide significance level. Therefore, it is likely that rare variants or SNPs with smaller effects on the risk of SLE in the MHC region remain to be found.

In summary, our results suggest the presence of at least four independent signals in the MHC region associated with SLE in African American women, and they may be shared in part with other ethnic groups. Taken together, our results and previous GWAS in European and East Asian ancestry populations show that women of different ancestral origins may share some genetic components for the risk of SLE. The identity of the causal variants that are being tagged by the reported SNPs is still unknown. Further studies are needed to narrow the position of the potential causal variants.

Acknowledgments

This work was supported by grant U01 AI067146 from the National Institute of Allergy and Infectious Diseases, and by grants R01CA058420 and R01CA098663 from the National Cancer Institute. The genotyping for the present project was subsidized by a grant to the Broad Institute Center for Genotyping and Analysis grant U54 RR020278 from the National Center for Research Resources. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Allergy and Infectious Diseases, the National Cancer Institute, or the National Institutes of Health. We thank Dr. Timothy McAlindon and Dr. Elizaveta Vaysbrot for their medical record reviews.

References

- Alarcon-Segovia D, Alarcon-Riquelme ME, Cardiel MH, Caeiro F, Massardo L, Villa AR, et al. Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort. *Arthritis Rheum.* 2005; 52:1138–47. [PubMed: 15818688]
- Barcellos LF, May SL, Ramsay PP, Quach HL, Lane JA, Nititham J, et al. High-density SNP screening of the major histocompatibility complex in systemic lupus erythematosus demonstrates strong evidence for independent susceptibility regions. *PLoS Genet.* 2009; 5:e1000696. [PubMed: 19851445]
- Campbell RD, Trowsdale J. Map of the human MHC. *Immunol Today.* 1993; 14:349–52. [PubMed: 8363724]
- Candore G, Lio D, Colonna Romano G, Caruso C. Pathogenesis of autoimmune diseases associated with 8.1 ancestral haplotype: effect of multiple gene interactions. *Autoimmun Rev.* 2002; 1:29–35. [PubMed: 12849055]
- Cozier YC, Palmer JR, Rosenberg L. Comparison of methods for collection of DNA samples by mail in the Black Women's Health Study. *Ann Epidemiol.* 2004; 14:117–22. [PubMed: 15018884]
- Criswell LA. The genetic contribution to systemic lupus erythematosus. *Bull NYU Hosp Jt Dis.* 2008; 66:176–83. [PubMed: 18937628]
- de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, et al. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet.* 2006; 38:1166–72. [PubMed: 16998491]
- Deapen D, Escalante A, Weinrib L, Horwitz D, Bachman B, Roy-Burman P, et al. A revised estimate of twin concordance in systemic lupus erythematosus. *Arthritis Rheum.* 1992; 35:311–8. [PubMed: 1536669]
- Fernando MM, Stevens CR, Sabeti PC, Walsh EC, McWhinnie AJ, Shah A, et al. Identification of two independent risk factors for lupus within the MHC in United Kingdom families. *PLoS Genet.* 2007; 3:e192. [PubMed: 17997607]
- Fraser PA, Lu LY, Ding WZ, Najundaswamy SN, Chen DF, Uko G, et al. HLA-B*44031; DRB1*1503 and other sub-Saharan African major histocompatibility complex haplotypes in African Americans and Afro-Caribbeans carry C4A gene deletions: implications for ethnicity-specific lupus susceptibility genes. *Arthritis Rheum.* 2000; 43:2378–9. [PubMed: 11037899]

- Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature*. 2007; 449:851–61. [PubMed: 17943122]
- Freedman ML, Haiman CA, Patterson N, McDonald GJ, Tandon A, Waliszewska A, et al. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. *Proc Natl Acad Sci U S A*. 2006; 103:14068–73. [PubMed: 16945910]
- Gambelunghe G, Gerli R, Bocci EB, Del Sindaco P, Ghaderi M, Sanjeevi CB, et al. Contribution of MHC class I chain-related A (MICA) gene polymorphism to genetic susceptibility for systemic lupus erythematosus. *Rheumatology (Oxford)*. 2005; 44:287–92. [PubMed: 15522921]
- Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, et al. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. *Nat Genet*. 2008; 40:1059–61. [PubMed: 19165918]
- Graham RR, Ortmann W, Rodine P, Espe K, Langefeld C, Lange E, et al. Specific combinations of HLA-DR2 and DR3 class II haplotypes contribute graded risk for disease susceptibility and autoantibodies in human SLE. *Eur J Hum Genet*. 2007; 15:823–30. [PubMed: 17406641]
- Graham RR, Ortmann WA, Langefeld CD, Jawaheer D, Selby SA, Rodine PR, et al. Visualizing human leukocyte antigen class II risk haplotypes in human systemic lupus erythematosus. *Am J Hum Genet*. 2002; 71:543–53. [PubMed: 12145745]
- Grant SF, Kristjansdottir H, Steinsson K, Blondal T, Yuryev A, Stefansson K, et al. Long PCR detection of the C4A null allele in B8-C4AQ0-C4B1-DR3. *J Immunol Methods*. 2000; 244:41–7. [PubMed: 11033017]
- Han JW, Zheng HF, Cui Y, Sun LD, Ye DQ, Hu Z, et al. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat Genet*. 2009; 41:1234–7. [PubMed: 19838193]
- Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXX, KIAA1542 and other loci. *Nat Genet*. 2008; 40:204–10. [PubMed: 18204446]
- Harley JB, Moser KL, Gaffney PM, Behrens TW. The genetics of human systemic lupus erythematosus. *Curr Opin Immunol*. 1998; 10:690–6. [PubMed: 9914226]
- Hartung K, Baur MP, Coldewey R, Fricke M, Kalden JR, Lakomek HJ, et al. Major histocompatibility complex haplotypes and complement C4 alleles in systemic lupus erythematosus. Results of a multicenter study. *J Clin Invest*. 1992; 90:1346–51. [PubMed: 1401069]
- Helmick CG, Felson DT, Lawrence RC, Gabriel S, Hirsch R, Kwoh CK, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. *Arthritis Rheum*. 2008; 58:15–25. [PubMed: 18163481]
- Hinds DA, Stuve LL, Nilsen GB, Halperin E, Eskin E, Ballinger DG, et al. Whole-genome patterns of common DNA variation in three human populations. *Science*. 2005; 307:1072–9. [PubMed: 15718463]
- Hoggart CJ, Parra EJ, Shriver MD, Bonilla C, Kittles RA, Clayton DG, et al. Control of confounding of genetic associations in stratified populations. *Am J Hum Genet*. 2003; 72:1492–1504. [PubMed: 12817591]
- Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, et al. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. *N Engl J Med*. 2008; 358:900–9. [PubMed: 18204098]
- Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, et al. Gene map of the extended human MHC. *Nat Rev Genet*. 2004; 5:889–99. [PubMed: 15573121]
- Ide A, Babu SR, Robles DT, Wang T, Erlich HA, Bugawan TL, et al. “Extended” A1, B8, DR3 haplotype shows remarkable linkage disequilibrium but is similar to nonextended haplotypes in terms of diabetes risk. *Diabetes*. 2005; 54:1879–83. [PubMed: 15919812]
- Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*. 2008; 24:2938–9. [PubMed: 18974171]

- McAlindon TE, Formica M, Palmer JR, Lafyatis R, Rosenberg L. Assessment of strategies for identifying diagnosed cases of systemic lupus erythematosus through self-report. *Lupus*. 2003; 12:754–9. [PubMed: 14596424]
- McKeigue PM, Carpenter JR, Parra EJ, Shriver MD. Estimation of admixture and detection of linkage in admixed populations by a Bayesian approach: application to African-American populations. *Ann Hum Genet*. 2000; 64:171–86. [PubMed: 11246470]
- Moser KL, Kelly JA, Lessard CJ, Harley JB. Recent insights into the genetic basis of systemic lupus erythematosus. *Genes Immun*. 2009; 10:373–9. [PubMed: 19440199]
- Palmer JR, Wise LA, Horton NJ, Adams-Campbell LL, Rosenberg L. Dual effect of parity on breast cancer risk in African-American women. *J Natl Cancer Inst*. 2003; 95:478–83. [PubMed: 12644541]
- Price P, Witt C, Allcock R, Sayer D, Garlepp M, Kok CC, et al. The genetic basis for the association of the 8.1 ancestral haplotype (A1, B8, DR3) with multiple immunopathological diseases. *Immunol Rev*. 1999; 167:257–74. [PubMed: 10319267]
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81:559–75. [PubMed: 17701901]
- Reich D, Patterson N, De Jager PL, McDonald GJ, Waliszewska A, Tandon A, et al. A whole-genome admixture scan finds a candidate locus for multiple sclerosis susceptibility. *Nat Genet*. 2005; 37:1113–8. [PubMed: 16186815]
- Reveille JD, Moulds JM, Ahn C, Friedman AW, Baethge B, Roseman J, et al. Systemic lupus erythematosus in three ethnic groups: I. The effects of HLA class II, C4, and CR1 alleles, socioeconomic factors, and ethnicity at disease onset. LUMINA Study Group. Lupus in minority populations, nature versus nurture. *Arthritis Rheum*. 1998; 41:1161–72. [PubMed: 9663471]
- Rioux JD, Goyette P, Vyse TJ, Hammarstrom L, Fernando MM, Green T, et al. Mapping of multiple susceptibility variants within the MHC region for 7 immune-mediated diseases. *Proc Natl Acad Sci U S A*. 2009a
- Rioux JD, Goyette P, Vyse TJ, Hammarstrom L, Fernando MM, Green T, et al. Mapping of multiple susceptibility variants within the MHC region for 7 immune-mediated diseases. *Proc Natl Acad Sci U S A*. 2009b; 106:18680–5. [PubMed: 19846760]
- Sanchez E, Torres B, Vilches JR, Lopez-Nevot MA, Ortego-Centeno N, Jimenez-Alonso J, et al. No primary association of MICA polymorphism with systemic lupus erythematosus. *Rheumatology (Oxford)*. 2006; 45:1096–100. [PubMed: 16537577]
- Smerdel-Ramoya A, Finholt C, Lilleby V, Gilboe IM, Harbo HF, Maslinski S, et al. Systemic lupus erythematosus and the extended major histocompatibility complex--evidence for several predisposing loci. *Rheumatology (Oxford)*. 2005; 44:1368–73. [PubMed: 16174649]
- Smith MW, Patterson N, Lautenberger JA, Truelove AL, McDonald GJ, Waliszewska A, et al. A high-density admixture map for disease gene discovery in african americans. *Am J Hum Genet*. 2004; 74:1001–13. [PubMed: 15088270]
- Tan FK, Arnett FC. The genetics of lupus. *Curr Opin Rheumatol*. 1998; 10:399–408. [PubMed: 9746854]
- The MHC sequencing consortium. Complete sequence and gene map of a human major histocompatibility complex. *Nature*. 1999; 401:921–3. [PubMed: 10553908]
- Uribe AG, McGwin G Jr, Reveille JD, Alarcon GS. What have we learned from a 10-year experience with the LUMINA (Lupus in Minorities; Nature vs. nurture) cohort? Where are we heading? *Autoimmun Rev*. 2004; 3:321–9. [PubMed: 15246029]
- Valdes AM, McWeeney S, Thomson G. HLA class II DR-DQ amino acids and insulin-dependent diabetes mellitus: application of the haplotype method. *Am J Hum Genet*. 1997; 60:717–28. [PubMed: 9042932]
- Valdes AM, Thomson G. Detecting disease-predisposing variants: the haplotype method. *Am J Hum Genet*. 1997; 60:703–16. [PubMed: 9042931]
- Walport MJ. The Roche Rheumatology Prize Lecture. Complement deficiency and disease. *Br J Rheumatol*. 1993; 32:269–73. [PubMed: 8461918]

Yao Z, Hartung K, Deicher HG, Brunnler G, Bettinotti MP, Keller E, et al. DNA typing for HLA-DPB1-alleles in German patients with systemic lupus erythematosus using the polymerase chain reaction and DIG-ddUTP-labelled oligonucleotide probes. Members of SLE Study Group. Eur J Immunogenet. 1993; 20:259–66. [PubMed: 8399121]

Figure 1. Scatterplot of the genotyped SNPs along the MHC region

The rs9271366 SNP was associated with risk of SLE at MHC-wide level. The other three SNPs: rs2044580, rs204890, and rs2071349 were associated with risk of SLE after conditioning for the rs9271366 using the conditional haplotypic method.

Table 1

Characteristics of systemic lupus erythematosus (SLE) cases and controls in the Black Women's Health Study

Characteristics	Cases(n=380)	Controls(n=765)
Mean age at baseline(1995)	39	39
Region of residence(%)		
Northeast	24.6	21.2
South	30.7	36.0
Midwest	27.6	21.9
West	17.1	20.9
%European mean ancestry ^a	19.7	20.9

^aFrom the genotyping of 1,174 ancestral informative markers

Table 2
Odds ratios (ORs)^a and 95% confidence intervals (CIs) of the four newly identified SNPs and the C4A deletion

SNP	Risk allele	Risk allele frequency(%)		P-value ^b	Per-allele OR(95% CI)		Nearest gene(s)
		Cases	Controls		Individual models ^c	Combined model ^d	
rs2844580	G	28.8	22.7	1.3×10 ⁻³	1.43(1.17–1.75)	1.56(1.25–1.94)	HLA-B/MICA
rs204890	G	94.1	89.2	1.2×10 ⁻⁴	1.86(1.31–2.65)	1.96(1.35–2.85)	ATF6B
rs9271366	C	19.7	13.3	5.6×10 ⁻⁵	1.70(1.33–2.18)	1.79(1.38–2.33)	HLA-DRB1/HLA-DQA1
rs2071349	C	89.2	84.3	1.0×10 ⁻³	1.53(1.16–2.02)	1.65(1.23–2.22)	HLA-DPB1
C4A del	Deletion	6.5	4.8	0.075	1.38(0.93–2.04)	1.01(0.66–1.54)	C4A

^a Adjusted for age, region of residence, and percent of European admixture
^b P-value of the 1-df Cochran-Armitage test
^c ORs were estimated for each SNP and the C4A deletion in individual models
^d ORs were estimated using a single combined model that contained the four newly identified SNPs plus the C4A deletion

D' and r² values in the BWHs among the four identified SNPs rs2844580, rs204890, rs9271366, rs2071349, and the deletion of the *C4A* gene (*C4A*del)^a

Table 3

SNPs	rs2844580	rs204890	rs9271366	rs2071349	C4A
rs2844580		0.17	0.11	0.02	0.50
rs204890	0.00		0.37	0.00	0.50
rs9271366	0.00	0.00		0.10	0.26
rs2071349	0.00	0.00	0.01		0.02
C4A 	0.04	0.00	0.02	0.00	

^aD' values are above the diagonal; r² values are below the diagonal

Table 4

Odds ratios (ORs) and 95% confidence intervals (CIs) of the genotype score based on the four identified SNPs, rs2844580, rs204890, rs9271366, and rs2071349

Number of risk alleles	Frequency (%)		OR (95% CI)
	Cases (n=375)	Controls (n=756)	
≤3	10.4	20.9	1.00 (reference)
4	36.3	45.0	1.52 (1.01–2.29)
5	35.2	26.5	2.62 (1.72–3.98)
≥6	18.1	7.7	4.65 (2.82–7.69)

Case group	Per risk allele OR (95% CI)
All cases (n=375)	1.67 (1.46–1.92)
Incident cases or diagnosed <5 yrs before baseline (n=207)	1.54 (1.30–1.83)
Cases with 4+ symptoms (n=202)	1.70 (1.43–2.02)
Cases with 3+ symptoms (n=250)	1.69 (1.44–1.99)
Cases with abnormal antinuclear antibody titers (N=171)	1.58 (1.32–1.90)
Cases with arthritis (N=145)	1.37 (1.13–1.67)
Cases with immunologic disorder (N=109)	1.45 (1.17–1.81)
Cases with skin manifestations (N=103)	1.44 (1.15–1.81)
Cases with hematologic disorder (N=89)	1.79 (1.40–2.28)
Cases with serositis (N=65)	1.61 (1.22–2.12)
Cases with renal disorder(N=58)	1.59 (1.19–2.12)

Table 5

Comparison with results from previous GWAS and scan of the MHC region

Index SNP	Tested in BWHS	r ² in YRI	Reported OR ¹	MAF%, BWHS		BWHS OR (95% CI) ¹
				Cases	Controls	
SLEGEN GWAS ²						
rs3131379	rs3131379	1.00	2.36	9.4	6.0	1.68 (1.21–2.34)
rs1270942	rs440454	1.00	2.35	14.0	9.4	1.70 (1.28–2.26)
Chinese GWAS ³						
rs9271100	rs9271366	0.73	1.90	19.7	13.3	1.70 (1.33–2.18)
rs3997854	rs3997854	1.00	0.44	30.8	26.3	1.24 (1.02–1.50)
High-density MHC screening ⁴						
rs3828903	rs6916394	1.00	0.67	30.0	32.0	0.90 (0.75–1.09)
rs2246626	rs2246626	1.00	0.77	39.5	41.7	0.91 (0.76–1.10)
rs8283	rs429150	0.94	0.63	45.4	44.8	1.01 (0.84–1.21)
rs7746019	rs3129944	1.00	1.77	41.3	37.9	1.15 (0.96–1.38)
rs3117103	rs1980493	1.00	2.35	14.0	13.8	1.00 (0.77–1.29)
rs7769979	rs6918223	1.00	0.58	38.6	40.2	0.92 (0.77–1.11)
rs10947345	rs2621421	1.00	1.29	32.8	33.9	0.96 (0.80–1.16)
rs383711	rs9368758	0.90	0.65	8.2	9.9	0.82 (0.60–1.13)

¹ Reported and BWHS ORs are using the major allele as reference

² Harley et al. (Harley et al. 2008)

³ Han et al. (Han et al. 2009)

⁴ Barcellos et al. (Barcellos et al. 2009)