ARTICLE

Identification of *IRF8, TMEM39A,* and *IKZF3-ZPBP2* as Susceptibility Loci for Systemic Lupus Erythematosus in a Large-Scale Multiracial Replication Study

Christopher J. Lessard,[1,2] Indra Adrianto,[1] John A. Ice,[1] Graham B. Wiley,[1] Jennifer A. Kelly,[1] Stuart B. Glenn,[1] Adam J. Adler,[1] He Li,[1,2] Astrid Rasmussen,[1] Adrienne H. Williams,[3] Julie Ziegler,[3] Mary E. Comeau,[3] Miranda Marion,[3] Benjamin E. Wakeland,[4] Chaoying Liang,[4] Paula S. Ramos,[5] Kiely M. Grundahl,[1] Caroline J. Gallant,[6] Marta E. Alarcón-Riquelme for the BIOLUPUS and GENLES Networks,[1,7] Graciela S. Alarcón,[8] Juan-Manuel Anaya,[9] Sang-Cheol Bae,[10] Susan A. Boackle,[11] Elizabeth E. Brown,[8] Deh-Ming Chang,[12] Soo-Kyung Cho,[10] Lindsey A. Criswell,[13] Jeffrey C. Edberg,[8] Barry I. Freedman,[14] Gary S. Gilkeson,[5] Chaim O. Jacob,[15] Judith A. James,[1,2,16] Diane L. Kamen,[5] Robert P. Kimberly,[8] Jae-Hoon Kim,[10] Javier Martin,[17] Joan T. Merrill,[18] Timothy B. Niewold,[19] So-Yeon Park,[10] Michelle A. Petri,[20] Bernardo A. Pons-Estel,[21] Rosalind Ramsey-Goldman,[22] John D. Reveille,[23] R. Hal Scofield,[1,2,16,24] Yeong Wook Song,[25] Anne M. Stevens,[26,27] Betty P. Tsao,[28] Luis M. Vila,[29] Timothy J. Vyse,[30] Chack-Yung Yu,[31,32] Joel M. Guthridge,[1] Kenneth M. Kaufman,[1,33,34] John B. Harley,[33,34] Edward K. Wakeland,[4] Carl D. Langefeld,[3] Patrick M. Gaffney,[1,2] Courtney G. Montgomery,[1] and Kathy L. Moser[1,2,*]

Systemic lupus erythematosus (SLE) is a chronic heterogeneous autoimmune disorder characterized by the loss of tolerance to self-antigens and dysregulated interferon responses. The etiology of SLE is complex, involving both heritable and environmental factors. Candidate-gene studies and genome-wide association (GWA) scans have been successful in identifying new loci that contribute to disease susceptibility; however, much of the heritable risk has yet to be identified. In this study, we sought to replicate 1,580 variants showing suggestive association with SLE in a previously published GWA scan of European Americans; we tested a multiethnic population consisting of 7,998 SLE cases and 7,492 controls of European, African American, Asian, Hispanic, Gullah, and Amerindian ancestry to find association with the disease. Several genes relevant to immunological pathways showed association with SLE. Three loci exceeded the genome-wide significance threshold: interferon regulatory factor 8 (*IRF8*; rs11644034; $p_{meta\text{-}Euro} = 2.08 \times 10^{-10}$), transmembrane protein 39A (*TMEM39A*; rs1132200; $p_{meta\text{-}all} = 8.62 \times 10^{-9}$), and 17q21 (rs1453560; $p_{meta\text{-}all} = 3.48 \times 10^{-10}$) between IKAROS family of zinc finger 3 (*AIOLOS*; *IKZF3*) and zona pellucida binding protein 2 (*ZPBP2*). Fine mapping, resequencing, imputation, and haplotype analysis of *IRF8* indicated that three independent effects tagged by rs8046526, rs450443, and rs4843869, respectively, were required for risk in individuals of European ancestry. Eleven additional replicated effects ($5 \times 10^{-8} < p_{meta\text{-}Euro} < 9.99 \times 10^{-5}$) were observed with *CFHR1, CADM2, LOC730109/IL12A, LPP, LOC63920, SLU7, ADAMTSL1, C10orf64, OR8D4, FAM19A2,* and *STXBP6*. The results of this study increase the number of confirmed SLE risk loci and identify others warranting further investigation.

[1]Arthritis and Clinical Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA; [2]Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA; [3]Department of Biostatistical Sciences, Wake Forest University Health Sciences, Winston-Salem, NC 27157, USA; [4]Department of Immunology, University of Texas Southwestern Medical Center at Dallas, Dallas, TX 75390, USA; [5]Division of Rheumatology and Immunology, Department of Medicine, Medical University of South Carolina, Charleston, SC 29425, USA; [6]Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Uppsala 75105, Sweden; [7]Centro de Genómica e Investigaciones Oncológicas, Pfizer-Universidad de Granada-Junta de Andalucía, Granada 18100, Spain; [8]Division of Clinical Immunology and Rheumatology, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL 35294, USA; [9]Center for Autoimmune Diseases Research, Universidad del Rosario, Bogotá, Colombia; [10]Department of Rheumatology, Hanyang University Hospital for Rheumatic Diseases, Seoul 133-792, Korea; [11]Division of Rheumatology, University of Colorado Denver, Aurora, CO 80045, USA; [12]National Defense Medical Center, Taipei 114, Taiwan; [13]Rosalind Russell Medical Research Center for Arthritis, University of California, San Francisco, San Francisco, CA 94143, USA; [14]Section on Nephrology, Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA; [15]Department of Medicine, University of Southern California, Los Angeles, CA 90089, USA; [16]Department of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA; [17]Instituto de Parasitología y Biomedicina López-Neyra, Consejo Superior de Investigaciones Cientificas, Granada 18100, Spain; [18]Clinical Pharmacology, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA; [19]Section of Rheumatology and Gwen Knapp Center for Lupus and Immunology Research, University of Chicago, Chicago, IL 60637, USA; [20]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA; [21]Sanatorio Parque, Rosario 2000, Argentina; [22]Division of Rheumatology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA; [23]Rheumatology and Clinical Immunogenetics, University of Texas Health Science Center at Houston, Houston, TX 77030, USA; [24]US Department of Veterans Affairs Medical Center, Oklahoma City, OK 73104, USA; [25]Division of Rheumatology, Seoul National University, Seoul 110-799, Korea; [26]Division of Rheumatology, Department of Pediatrics, University of Washington, Seattle, WA 98105, USA; [27]Center for Immunity and Immunotherapies, Seattle Children's Research Institute, Seattle, WA 98105, USA; [28]Division of Rheumatology, Department of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA; [29]Division of Rheumatology, Department of Medicine, University of Puerto Rico Medical Sciences Campus, San Juan 00936-5067, Puerto Rico; [30]Division of Genetics and Molecular Medicine and Division of Immunology, Infection, and Inflammatory Disease, King's College London, London SE1 9RT, UK; [31]Center for Molecular and Human Genetics, The Research Institute, Nationwide Children's Hospital, Columbus, OH 43205, USA; [32]Department of Pediatrics, Ohio State University, Columbus, OH 43205, USA; [33]Division of Rheumatology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA; [34]US Department of Veterans Affairs Medical Center, Cincinnati, OH 45220, USA
*Correspondence: moserk@omrf.org
DOI 10.1016/j.ajhg.2012.02.023.

Introduction

Systemic lupus erythematosus (SLE [MIM 152700]) is a chronic autoimmune disease that is classically characterized by inflammation, dysregulated type 1 interferon responses, and autoantibodies directed to the nuclear compartment. Women of childbearing age are preferentially affected at a rate nine times that of men, and those of African American and Asian ancestries are affected more frequently and manifest more severe disease than those of European ancestry.[1] Although the etiology of SLE is largely unknown, its pathogenesis most likely involves a complex interplay between environmental (e.g., UV light, Epstein-Barr virus infection, etc.) and genetic (e.g., *MHC*, *IRF5* [MIM 607218], etc.) components.[2] A sibling risk ratio (λ_s) of approximately 30 in SLE illustrates a strong genetic component,[3] and the fact that observational studies have identified many families with multiple cases of SLE and other autoimmune conditions suggests the potential for shared genetic predisposition.[4–6]

Candidate-gene studies and, more recently, genome-wide association (GWA) scans have been highly successful in identifying multiple susceptibility loci.[2,7] The histocompatibility leukocyte antigen (HLA) region has been known to contribute to the risk of SLE and other related autoimmune diseases since the 1970s.[8–11] In the early 2000s, gene expression studies determined that, compared to healthy controls, individuals with SLE overexpress genes in the interferon pathway.[12–14] Association between SLE and variants in the region of *IRF5* was first reported in 2005 and has since been replicated in most GWA scans of SLE.[15–19] In 2008, four GWA scans of SLE cases of European descent were published, and the first GWA scan of Asian descent was published in 2009.[16,17,19–21] Collectively, these studies have identified and confirmed ~35 loci that contribute to the pathogenesis of SLE. These data highlight the importance of several pathways, including those involving lymphocyte activation and function, immune-complex clearance, innate immune response, and adaptive immune responses.[2] However, a substantial portion of the heritable risk has yet to be identified.[17,22] The lack of causal variants, rare variants, and/or other loci yet to be discovered might account for the missing heritability.

In an effort to identify regions contributing to SLE risk, we sought to replicate suggestive association signals in our previously published European American SLE GWA scan.[19] We evaluated 1,580 single-nucleotide polymorphisms (SNPs) in an independent population of 7,998 SLE cases and 7,492 controls of European, African American, Asian, Hispanic, Gullah, and Amerindian ancestry (Tables S1–S3, available online). Three loci, interferon regulatory factor 8 (*IRF8* [MIM 601565]), transmembrane protein 39A (*TMEM39A*), and the region between IKAROS family of zinc finger 3 (*AIOLOS; IKZF3* [MIM 606221]) and zona pellucida binding protein 2 (*ZPBP2* [MIM 608449]) exceeded the genome-wide significance threshold ($p < 5 \times 10^{-8}$). Through fine mapping, resequencing, and imputation of the *IRF8* region, we identified three independent effects required for risk. Moreover, we replicated 11 other loci, several of which had been previously reported in related conditions.

Subjects and Methods

GWA Scan

Genotyping, quality control, procedures for data analysis, and summary statistics for the GWA scan were described previously in Graham et al., 2008.[19]

Study Design

The genotype data used in this study were generated as a part of a joint effort of more than 40 investigators from around the world. These investigators contributed samples, funding, and hypotheses on a combined array containing ~35,000 SNPs (Figure S1). The Oklahoma Medical Research Foundation (OMRF) served as the coordinating center, ran the arrays, and sent the data to a central facility for quality control at Wake Forest Medical Center. These data were then distributed back to the investigators, who requested the SNPs for final analysis and publication.[23–28]

Subjects

The multiracial replication study consisted of 17,003 total samples (8,922 SLE cases and 8,077 controls) and included individuals of self-reported African American, Asian, European, Gullah, Hispanic, and Amerindian ancestry (Table S1). A total of 374 samples were common between the GWA scan and the replication study so that genotypes generated by the two platforms could be confirmed and so that genotypes of SNPs not present on the Affymetrix 5.0 array could be obtained. These data were only used as observed data for the imputation analysis of specific genomic regions, as described below; to maintain independence between the GWA scan and replication samples, we did not include the data generated on these shared samples in the replication or fine-mapping analyses. The OMRF gathered the samples from consenting subjects (according to the guidelines of the ethics committees at the respective institutions where the samples were collected) and prepared them for genotyping. All cases used in this study fulfilled at least 4 of the 11 American College of Rheumatology criteria for SLE, whereas the healthy, population-based controls did not have any family history of SLE or any other autoimmune disease.[29]

Genotyping and Sample Quality Control

A total of 1,580 SNPs that attained $p < 0.05$ in the previously published GWA scan were selected for replication. In addition, 287 SNPs (chosen to capture all variation with a minimum r^2 threshold of 0.8 via the TAGGER algorithm in HAPLOVIEW[30]) within the *IRF8* region and 347 ancestral-informative markers (AIMs) spanning the genome were genotyped. Genotyping of SNPs was performed at OMRF with Infinium chemistry on an Illumina iSelect custom array according to the manufacturer's protocol. The following quality-control procedures were implemented prior to the analysis (Table S2): there were well-defined clusters within the scatter plots, the SNP call rate was >90% across all samples genotyped, the minor allele frequency was >1%, the sample call rate was >90%, $p > 0.05$ for differential missingness between cases and controls, the total proportion missing was <5%, and the

Hardy-Weinberg proportions were $p > 0.01$ in controls and $p > 0.0001$ in cases.

Samples exhibiting excess heterozygosity (>5 standard deviations [SDs] from the mean) or a <90% call rate were excluded from the analysis. The remaining individuals were examined for excessive allele sharing as estimated by identity-by-descent (IBD). In sample pairs with excess relatedness (IBD > 0.4), one individual was removed from the analysis on the basis of the following criteria: (1) remove the sample with the lower call rate, (2) remove the control and retain the case, (3) remove the male sample before the female sample, (4) remove the younger control before the older control, and (5) in a situation with two cases, remove the case with fewer phenotype data available. Discrepancies between self-reported and genetically determined gender were evaluated. Males were required to be heterozygous at rs2557523 (given that the G allele for this SNP is only observed on the Y chromosome and the A allele appears only on the X chromosome) and to have ≤10% chromosome X heterozygosity. Females were required to be homozygous for the A allele at rs2557523 and to have >10% chromosome X heterozygosity.

Ascertainment of Population Stratification

Genetic outliers from each ethnic and/or racial group were removed from further analysis as determined by principal-component analysis and admixture estimates (Figure S6).[31,32] Using the 163 AIMs that passed quality control in both EIGENSTRAT[31] and ADMIXMAP[33,34] to distinguish the four continental ancestral populations—Africans, Europeans, Amerindians, and East Asians—allowed identification of the substructure within the sample set (Figure S6A).[35,36] We utilized principal components from EIGENSTRAT outputs to identify outliers >4 SDs from the mean of each of the first three principal components (PC) for the individual population clusters. After quality control, a total of 1,139 samples were excluded (Figure S6B and Table S3). Overall, 2,586 subjects were included in the GWA scan, and 15,490 subjects were included in the replication study, resulting in a total of 18,076 subjects.

Statistical Analysis

Testing for SNP-SLE association in the replication study was carried out through the computation of logistic regression as implemented in PLINK v.1.07.[37] The calculation of the additive genetic model included adjustments for the first three PCs and gender. Models were also adjusted for ancestry with estimates provided by ADMIXMAP and resulted in no observable difference in association as compared with PC adjustment. We conducted conditional likelihood-ratio tests by using the extended WHAP functionality in PLINK v.1.07. The genome-wide p value threshold for all data replicating GWA-scan results was $p < 5 \times 10^{-8}$ after meta-analysis. For the fine mapping and imputation of *IRF8*, we utilized a Bonferroni-corrected p value threshold of $p < 1.09 \times 10^{-4}$ on the basis of the maximum number of tests across all populations (460 independent variants with $r^2 < 0.8$). Using METAL,[38] we performed meta-analyses of the SNPs observed in both the GWA scan and the multiracial replication study with a weighted Z score. Each racial group was weighted by the square root of its sample size so that sample-size differences could be controlled for between studies. We combined all data generated unless the variant failed quality control in a given racial group.

To test for meta-analysis heterogeneity, we utilized both Cochran's Q test and the I^2 index. Cochran's Q test is a classical method that calculates the weighted sum of the squared deviations between individual study effects and the overall effect across studies.[39] It follows a chi-square distribution with k-1 degrees of freedom (k is the number of studies). A value of $p < 0.05$ was considered significant evidence of heterogeneity. The I^2 index measures the degree or percentage of inconsistency—due to heterogeneity rather than random chance—across studies.[40] The I^2 index ranges from 0% to 100% and can equal 0%–25% (low heterogeneity), 26% –50% (moderate heterogeneity), 51%–75% (high heterogeneity), or 76%–100% (very high heterogeneity).

Linkage disequilibrium (LD) and probable haplotypes were determined with HAPLOVIEW v.4.2.[30] We calculated haplotype blocks for those haplotypes present at >3% frequency by using the solid-spine of LD algorithms with minimum r^2 values of 0.8.[30]

Resequencing

We resequenced the *IRF8* region (chromosome 16: 84,488,150–84,539,352 bp) in 206 (92 SLE cases and 114 healthy controls) European and 46 (25 SLE cases and 21 healthy controls) African American subjects. For each sample, 3–5 μg of whole genomic DNA were sheared and prepared for sequencing with an Illumina Paired-End Genomic DNA Sample Prep Kit. Targeted regions of interest from each sample were then enriched with a SureSelect Target Enrichment System utilizing a custom-designed bait pool (Agilent Technologies). Resequencing was undertaken with an Illumina GAIIx platform according to standard procedures. Post-sequence data were processed with Illumina's Pipeline software v.1.7. All samples were sequenced to minimum average fold coverage of 25×.

Variant Detection and Quality Control

Unique sequences corresponding to an individual nucleotide molecule were aligned to the human genome reference (hg18) with the Burrows-Wheeler Aligner (BWA)[41] alignment tool. Reads were locally realigned around known and suspected insertion and deletion sites with the Genome Analysis Toolkit (GATK) analysis suite so that the best possible read alignment could be generated.[42] Recalculation of the correct quality score for each base within the alignment was then performed empirically with the GATK suite. This process served to correct overestimated high-quality scores initially reported by the sequencer itself.

After local realignment around deletion-insertion polymorphism (DIP) sites and base quality-score recalibration, SNP and DIP genotypes were generated for each sample individually as well as for the samples as a whole. Finally, SNP and DIP genotypes were hard-filtered against a set of criteria designed to remove any remaining low-quality calls. For a variant to be included in the call list, we required a Phred quality score >30, a quality-by-depth ratio of >5.0, a strand bias score of <−0.10, and a homopolymer run of <5 bases. The variant phase was determined with the program BEAGLE.[43] Variants meeting call parameters were output to files compatible with PLINK and other genotyping tools via the VCFtools analysis suite.

To assess the accuracy of sequence-based SNP calling, we cross-referenced the sequenced and genotyped allele calls. We observed ~99% concordance between genotypes and sequence-based variant detection, suggesting high-quality sequence data. We manually inspected the samples with ≥5% of variants differing between sequencing and genotyping to determine where sequence quality was poor. As an additional quality-control measure, we confirmed

each variant identified by our automated workflow by manual inspection of the assembled contig by using the Integrative Genomics Viewer (IGV) program.[44]

Imputation

To increase the informativeness of the *IRF8* region, we conducted imputation in subjects of European, African American, and Asian ancestry over a 100 kb interval spanning the *IRF8* locus. Imputation of the replication data across chromosome 16 (84.45–84.46 Mb) was performed with IMPUTE2 and the reference panels provided in Table S7.[45–47] Imputed genotypes were required to meet or exceed a probability threshold of 0.8, an information measure of >0.4, and the same quality-control-criteria thresholds described above for inclusion in the analyses.

Results

Two SNPs, rs11648084 and rs11644034, telomeric to *IRF8* at 16q24.1 were suggestive of association with SLE in our published GWA scan ($p = 5.99 \times 10^{-4}$ and 2.29×10^{-3}, respectively; odds ratio [OR] = 0.76 and 0.66, respectively; Table 1). Both rs11648084 and rs11644034 were replicated in the current, independent population of SLE cases and controls of European ancestry and exceeded the genome-wide threshold ($p_{meta\text{-}Euro} = 2.34 \times 10^{-9}$ and $p_{meta\text{-}Euro} = 2.08 \times 10^{-10}$, respectively; Table 1). However, neither SNP was significantly associated with SLE in any other population studied, perhaps as a result of the reduced sample size, clinical and/or genetic heterogeneity, decreased minor allele frequency, and/or reduced correlation with the causal variants (Table 1 and Table S4).

To better refine the association signal, 287 additional SNPs covering ~100 kb encompassing the *IRF8* coding region were genotyped (see Subjects and Methods; Figures 1A and 1D, Table 2, and Table S4). The most significant association in the European population was with rs9936079 ($p = 3.96 \times 10^{-9}$, OR = 0.77; Figures 1A and 1D and Table 2), located ~11 kb telomeric to *IRF8* and found to be in strong LD with rs11644034 ($r^2 = 0.92$; Figures 2B and 2C). The Asian population also exhibited association with rs9936079 ($p = 2.95 \times 10^{-3}$, OR = 0.73); however, rs9936079 failed to pass quality-control measures (see Subjects and Methods; differential missingness $p = 10^{-7}$) in individuals of African ancestry and was not associated with disease in patients of Hispanic or Amerindian ancestry. Meta-analysis yielded $p_{meta\text{-}all} = 9.28 \times 10^{-11}$ (Table 2 and Table S4).

We observed a modest association in the African Americans at rs2934498 ($p = 3.92 \times 10^{-4}$, OR = 0.83), which was also significant in those individuals of European ancestry ($p = 5.96 \times 10^{-6}$, OR = 1.19) but not in those of Asian ancestry (Figure 1B,D, Table 2, and Table S4). The strongest Asian association was observed in a region (rs11117427, $p = 1.99 \times 10^{-5}$, OR = 0.64) ~34 kb telomeric to *IRF8* (Figures 1C and 1D, Table 2, and Table S4). Association was also observed with rs11117427 ($p = 3.46 \times 10^{-4}$, OR = 0.84) in the Europeans (Figures 1A and 1D and Table 2). Interestingly, this SNP is only ~2 kb away from rs12444486, which has been reported by Gateva et al.[22] as being suggestive of association with SLE in Europeans (Figure 1D and Table S4).

We resequenced the *IRF8* region in 206 subjects of European ancestry and in 46 subjects of African American ancestry to identify variants not previously evaluated within the *IRF8* region and to assess their association with SLE (see Subjects and Methods). Thirty-eight and 85 variants not present in dbSNP 130 were identified in European and African American individuals, respectively. After imputing these data into our larger European and African American datasets (see Subjects and Methods), the most significantly associated region within the European population was ~19 kb telomeric to *IRF8* (rs4843869, $p = 7.61 \times 10^{-10}$, OR = 0.76; Table 2 and Figures 1A and 1D). Ultimately, three strongly correlated ($r^2 > 0.90$) SNPs emerged as the most significantly associated with SLE in the Europeans: rs11644034 (identified via GWA scan), rs9936079 (identified by fine mapping), and rs4843869 (imputed on the basis of resequencing). Interestingly, our targeted resequencing revealed a DIP (rs11347703, $p = 1.11 \times 10^{-8}$, OR = 0.78) that was located less than 100 bp from a genotyped SNP, rs8052690 ($p = 5.69 \times 10^{-8}$, OR = 0.79, Table 2). This DIP, which has high biological plausibility, was in strong LD with the peak European SNP (rs4843869) and rs8052690 ($r^2/D' > 0.9$; Figures 2B and 2C). Of note, some of the African Americans resequenced in this study did harbor the DIP (rs11347703) identified in the Europeans. However, neither rs11347703 nor any SNP correlated with it was found to be significantly associated with SLE in African Americans, which could be due to the decrease in power and/or a decrease in the minor allele frequency.

The peak association in African Americans after imputation was at rs450443 ($p = 1.41 \times 10^{-4}$, OR = 0.82; Table 2 and Figures 1B and 1D) and was in strong LD ($r^2 = 0.88$) with rs2934498. Patients of European but not Asian ancestry showed association with rs450443 ($p = 9.73 \times 10^{-6}$, OR = 1.18; Table 2 and Figures 1A, 1C, and 1D). We conducted imputation in the Asian population by using 1,000 Genomes phased haplotypes;[48] however, rs11117427 remained the peak signal (Table 2 and Figures 1C and 1D).

To assess the independence of variants in the European population, we used logistic regression models that adjusted for the best tagging SNP at each signal. When we adjusted for rs4843869 in Europeans, the association persisted at rs450443, and variants correlated to it. However, adjusting for rs4843869 negated the association with rs11117427 and its correlated variants (Figures 2B, 2C, and 3, Table S5, and Figure S2A). Adjusting for either rs450443 or rs11117427 was only able to negate the associations of the polymorphisms that were correlated with each of these SNPs (Figures 2B, 2C, and 3, Table S5, and

Table 1. SLE Risk Loci Surpassing the Genome-wide Significance Threshold[a]

				European (3,562 Cases/3,491 Controls)					African American (1,527 Cases/1,811 Controls)		Asian (1,265 Cases/1,260 Controls)		Meta	Test of Heterogeneity	
Chr	SNP	Locus	Alleles[b]	$p_{GWA\ scan}$[c]	$OR_{GWA\ scan}$ (95% CI)	p_{REP}	OR_{REP} (95% CI)	$p_{META\text{-}Euro}$	p	OR (95% CI)	p	OR (95% CI)	$p_{META\text{-}ALL}$[d]	p_Q[e]	I^2
3	rs1132200	*TMEM39A*	G/A	1.65×10^{-3}	0.72 (0.59–0.88)	2.37×10^{-4}	0.83 (0.76–0.92)	1.81×10^{-6}	6.92×10^{-2}	0.75 (0.56–1.02)	1.66×10^{-3}	0.73 (0.59–0.89)	8.62×10^{-9}	0.450	0.0%
16	rs11644034	*IRF8*	G/A	2.29×10^{-3}	0.66 (0.54–0.79)	2.36×10^{-8}	0.78 (0.71–0.85)	2.08×10^{-10}	5.10×10^{-1}	0.95 (0.81–1.11)	2.63×10^{-2}	0.79 (0.65–0.97)	2.72×10^{-9}	0.016	61.5%
16	rs11648084	*IRF8*	G/A	5.99×10^{-4}	0.76 (0.65–0.89)	9.35×10^{-7}	0.83 (0.77–0.89)	2.34×10^{-9}	6.33×10^{-2}	0.90 (0.81–1.01)	7.52×10^{-1}	1.02 (0.91–1.14)	7.00×10^{-7}	0.001	74.0%
17	rs9913957	*IKZF3*	A/G	7.87×10^{-3}	1.75 (1.27–2.41)	5.14×10^{-4}	1.38 (1.15–1.66)	1.38×10^{-5}	1.07×10^{-2}	1.22 (1.05–1.41)	–	–	1.39×10^{-8}	0.105	45.1%
17	rs8076347	*IKZF3*	C/A	3.07×10^{-3}	1.93 (1.41–2.62)	3.04×10^{-3}	1.32 (1.10–1.58)	4.75×10^{-5}	2.19×10^{-3}	1.20 (1.07–1.34)	–	–	3.01×10^{-8}	0.047	55.4%
17	rs8079075	*IKZF3*	A/G	1.47×10^{-3}	1.90 (1.39–2.59)	5.08×10^{-4}	1.39 (1.16–1.68)	3.81×10^{-6}	2.62×10^{-3}	1.26 (1.08–1.46)	–	–	4.83×10^{-9}	0.201	31.3%
17	rs1453560	*ZPBP2*	A/C	7.81×10^{-4}	1.92 (1.41–2.61)	6.42×10^{-4}	1.37 (1.14–1.64)	3.21×10^{-6}	4.86×10^{-4}	1.23 (1.09–1.37)	–	–	3.48×10^{-10}	0.097	46.4%

The following abbreviations are used: Chr, chromosome; OR, odds ratio; GWA, genome-wide association; REP, replication; and CI, confidence interval.
[a]Table S7 contains results for all populations evaluated within this study.
[b]Major/minor alleles.
[c]Result of GWA scan was previously reported in Graham et al.[1]
[d]Data were combined for all racial groups genotyped within our study that passed quality control.
[e]Cochran's Q test statistic.

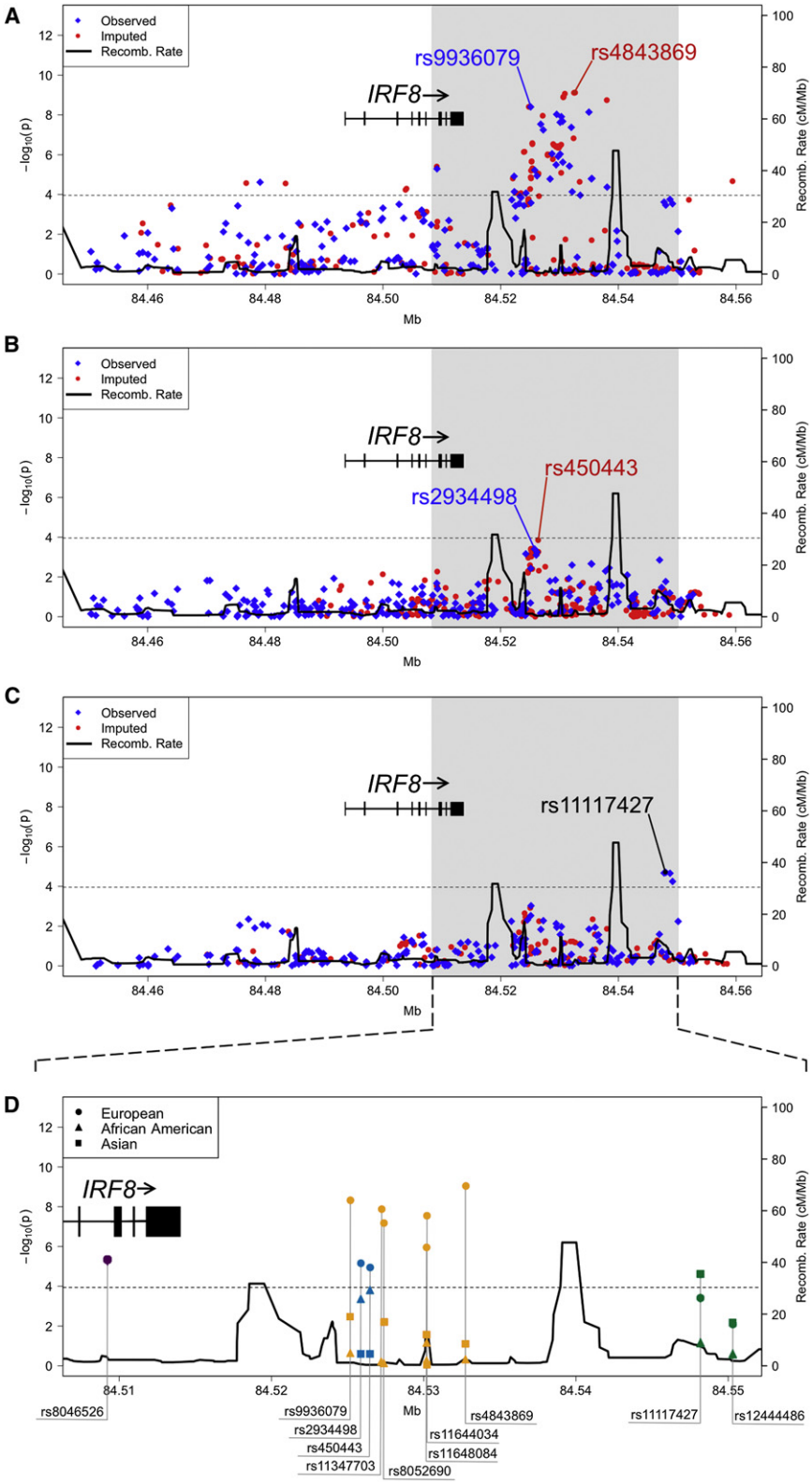

Figure 1. Variants in the Region of *IRF8* Tested for Association with SLE

The association between *IRF8* and SLE in European (A), African American (B), and Asian (C) ancestral populations is given with observed (blue diamonds) and imputed (red circles) variants. The dotted line represents the Bonferroni-corrected threshold ($p = 1.09 \times 10^{-4}$) for the fine-mapping study. The solid black line represents the recombination rate. The variants labeled with blue text represent the most significant observed SNPs, whereas the variants labeled in red represent the most significant SNPs after imputation. In the Asians, rs11117427 was both the most significant observed and imputed variant.

(D) Shown is an expanded view of the most statistically significant region in Europeans (circles), African Americans (triangles), and Asians (squares) for selected variants tagged by rs8046526 (purple), rs450443 (turquoise), rs4843869 (yellow-orange), and rs11117427 (green). The following abbreviation is used: Recomb., recombination.

Figures S2B and S2C). A SNP (rs8046526) in the sixth intron of *IRF8* was also associated with SLE risk in Europeans ($p = 3.96 \times 10^{-6}$, OR = 0.80) and remained significant after adjusting for the other SNPs (Table 2, Figures 1A, 1D, 2B, 2C, and 3, Table S5, and Figure S2D). Adjusting for rs8046526 in the Europeans only negated associations for itself and its correlated variants. However, adjusting the logistic regression model for rs8046526, rs450443, and rs4843869 negated all associations present in the European population (Figures 2B, 2C, and 3, Table S5, and Figure S2E), demonstrating the importance of these three *IRF8* variants for SLE risk.

Haplotype analysis identified a single risk haplotype (H2) ($p = 6.42 \times 10^{-8}$) with a frequency of 18.4% in the European individuals (Figure 2). Two significant protective haplotypes, H6 and H7, were also identified (Figure 2). The risk-associated alleles within the region bounded by SNPs rs11117426–rs34912238 (the peak Asian effect) were also present in the most significant protective haplotype, H7, suggesting that this region might not impact disease risk in Europeans (Figure 2). The only differences between H3 and H6 as well as between H4 and H7 are rs8046526 and rs8058904 in the minor form, suggesting that these SNPs are important in conferring protection from disease (Figure 2). The only differences between H2 and H5 (which are not statistically significant) are the major alleles (for SNPs rs8046526 and rs8058904) residing on the H2

Table 2. *IRF8* Variants Associated with SLE[a]

SNP	Genotyped or Imputed	Position (bp)	European				$p_{African\text{-}American}$	p_{Asian}
			Alleles[b]	MAF[c]	p	OR (95% CI)		
rs8046526	I	84,509,136	C/T	0.14/0.16	3.96×10^{-6}	0.80 (0.73–0.88)	–	–
rs8058904	G	84,509,183	A/G	0.14/0.16	5.14×10^{-6}	0.80 (0.73–0.88)	1.96×10^{-1}	–
rs9936079	G	84,525,095	G/A	0.17/0.22	3.96×10^{-9}	0.77 (0.70–0.84)	–	2.95×10^{-3}
rs385344	I	84,525,105	C/G	0.30/0.27	1.37×10^{-5}	1.18 (1.10–1.27)	–	2.43×10^{-1}
rs34337659	I	84,525,158	T/C	0.31/0.27	1.55×10^{-5}	1.18 (1.10–1.27)	3.97×10^{-4}	1.14×10^{-1}
rs66509440	I	84,525,182	C/T	0.28/0.25	6.36×10^{-6}	1.20 (1.11–1.30)	3.97×10^{-4}	–
rs66804793	I	84,525,190	G/A	0.28/0.25	6.16×10^{-6}	1.20 (1.11–1.30)	3.84×10^{-4}	–
rs74032085	I	84,525,245	T/C	0.28/0.24	2.25×10^{-6}	1.22 (1.12-1.32)	4.72×10^{-4}	–
rs16940044	I	84,525,266	A/G	0.27/0.24	2.16×10^{-6}	1.22 (1.12–1.32)	4.73×10^{-4}	–
rs2934497	I	84,525,379	C/T	0.27/0.23	2.62×10^{-7}	1.24 (1.14–1.35)	5.12×10^{-4}	1.95×10^{-1}
rs2970091	I	84,525,387	G/A	0.28/0.24	2.96×10^{-7}	1.24 (1.14–1.34)	5.31×10^{-4}	1.93×10^{-1}
rs2934498	G	84,525,783	A/G	0.31/0.27	5.96×10^{-6}	1.19 (1.11–1.29)	3.92×10^{-4}	2.09×10^{-1}
rs439885	G	84,526,175	G/A	0.31/0.27	1.16×10^{-5}	1.19 (1.10–1.28)	5.59×10^{-4}	1.98×10^{-1}
rs450443	I	84,526,392	T/G	0.30/0.27	9.73×10^{-6}	1.18 (1.10–1.28)	1.41×10^{-4}	2.04×10^{-1}
rs396987	I	84,526,435	A/G	0.30/0.27	8.89×10^{-6}	1.19 (1.10–1.28)	5.44×10^{-4}	2.04×10^{-1}
rs4843865	G	84,526,806	T/A	0.17/0.21	2.93×10^{-8}	0.78 (0.72–0.85)	6.64×10^{-1}	1.46×10^{-2}
rs11347703	I	84,527,141	G/–	0.18/0.21	1.11×10^{-8}	0.78 (0.72–0.85)	5.73×10^{-1}	–
rs8052690	G	84,527,239	A/G	0.18/0.21	5.69×10^{-8}	0.79 (0.72–0.86)	6.58×10^{-1}	5.28×10^{-3}
rs186249	G	84,528,397	G/C	0.30/0.26	1.66×10^{-5}	1.19 (1.10–1.28)	1.63×10^{-2}	9.37×10^{-1}
rs11117422	G	84,529,514	G/C	0.17/0.21	9.37×10^{-9}	0.77 (0.71–0.84)	6.76×10^{-1}	1.13×10^{-2}
rs11644034	G	84,530,113	G/A	0.17/0.20	2.36×10^{-8}	0.78 (0.71–0.85)	5.10×10^{-1}	2.63×10^{-2}
rs305066	I	84,530,277	C/T	0.33/0.29	8.18×10^{-6}	1.18 (1.10–1.27)	–	3.62×10^{-1}
rs13335265	G	84,530,311	C/G	0.16/0.20	1.23×10^{-8}	0.77 (0.70–0.84)	2.86×10^{-1}	1.06×10^{-2}
rs12711490	G	84,530,529	A/G	0.17/0.20	2.11×10^{-8}	0.78 (0.71–0.85)	4.51×10^{-1}	7.55×10^{-2}
rs11641153	I	84,530,641	A/G	0.16/0.20	1.31×10^{-9}	0.76 (0.70–0.83)	4.85×10^{-1}	7.02×10^{-2}
rs11641155	I	84,530,653	A/G	0.16/0.20	1.23×10^{-9}	0.76 (0.70–0.83)	–	7.02×10^{-2}
rs7205434	I	84,530,696	C/G	0.16/0.20	1.23×10^{-9}	0.76 (0.70–0.83)	4.85×10^{-1}	7.02×10^{-2}
rs4843868	I	84,530,902	C/T	0.16/0.20	8.57×10^{-10}	0.76 (0.70–0.83)	5.82×10^{-1}	7.02×10^{-2}
rs305063	G	84,532,158	C/A	0.32/0.29	7.34×10^{-5}	1.17 (1.08–1.26)	–	9.66×10^{-1}
rs4843323	I	84,532,462	C/T	0.16/0.20	7.71×10^{-10}	0.76 (0.70–0.83)	–	–
rs4843869	I	84,532,642	G/A	0.16/0.20	7.61×10^{-10}	0.76 (0.70–0.83)	4.19×10^{-1}	6.59×10^{-2}
rs7202472	G	84,535,003	C/A	0.15/0.19	7.25×10^{-9}	0.77 (0.70–0.84)	3.94×10^{-1}	1.41×10^{-2}
rs11117426	G	84,547,768	A/G	0.16/0.19	2.42×10^{-4}	0.84 (0.77–0.92)	1.07×10^{-1}	2.12×10^{-5}
rs11117427	G	84,548,058	G/A	0.16/0.18	3.46×10^{-4}	0.84 (0.77–0.93)	–	1.99×10^{-5}
rs12445476	G	84,548,770	A/C	0.16/0.18	1.76×10^{-4}	0.84 (0.76–0.92)	6.49×10^{-2}	2.19×10^{-5}
rs11642873	G	84,549,206	A/C	0.15/0.18	2.92×10^{-4}	0.84 (0.77–0.92)	8.82×10^{-1}	5.63×10^{-5}
rs34912238	I	84,559,404	C/T	0.16/0.19	2.15×10^{-5}	0.82 (0.75–0.90)	–	–

The following abbreviations are used: G, genotyped; I, imputed; MAF, minor-allele frequency; OR, odds ratio; and CI, confidence interval.
[a]All subjects, including the 374 that were removed so that the replication study was independent from the GWA scan, were imputed. Tables S4 and S5 contain results for all populations evaluated within this study.
[b]Major/minor alleles.
[c]Case/control.

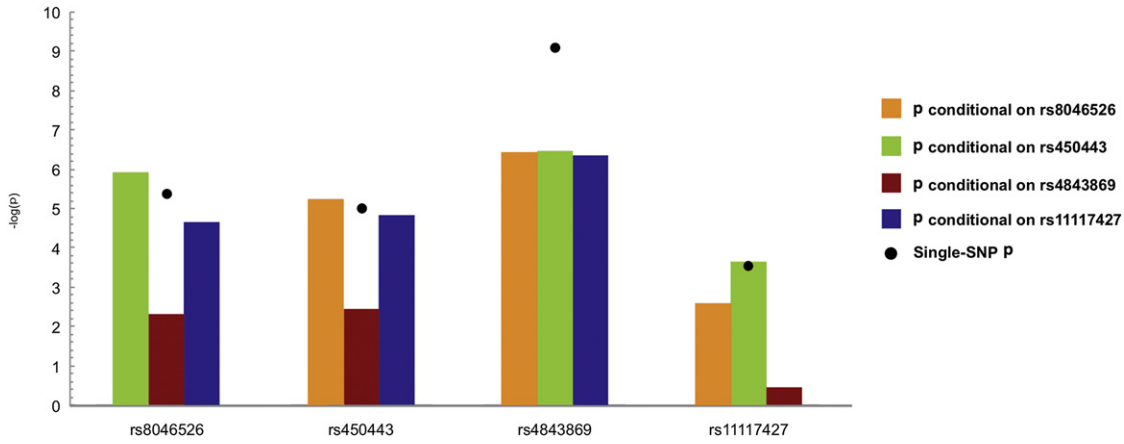

Figure 2. Conditional-Analysis Results Conducted in Individuals of European Ancestry
The results of the conditional analysis with the four SNPs show peak association in Europeans (rs8046526 and rs4843869), African Americans (rs450443), and Asians (rs11117427). The black dot represents the unadjusted single-marker association with SLE.

haplotype and the minor alleles on the neutral H5 haplotype. Thus, it appears that all three regions (tagged by rs8046526, rs450443, and rs4843869) are required for risk. Many variants residing on the risk haplotype are within regions known to bind multiple transcription factors in the ENCODE ChIP-Seq project dataset in immunologic cell types (Figures S3 and S4).[49] Thus, we hypothesize that the risk haplotype has the potential to affect the regulation of *IRF8* expression and/or the expression of other genes in the region.

Within *TMEM39A* in region 3q13.33, a coding SNP (rs1132200) that demonstrated suggestive evidence of association in our previous GWA scan ($p = 1.65 \times 10^{-3}$) was also confirmed in the European replication study ($p = 2.37 \times 10^{-4}$, OR = 0.83; Table 1 and Table S6). This nonsynonymous SNP showed association with SLE in Asian patients ($p = 1.66 \times 10^{-3}$, OR = 0.73) but not in African Americans, Hispanics, Gullah, or Amerindians (Table 1 and Table S6). When analyzing this SNP in all populations that passed quality control, a meta-analysis produced $p_{meta\text{-}all} = 8.62 \times 10^{-9}$, and no evidence of heterogeneity was observed between these datasets (Table 1 and Table S6).

Finally, we replicated several SNPs in the 17q12 region between *IKZF3* and *ZPBP2* (Table 1 and Table S6). Three SNPs within *IKZF3* replicated, and rs8079075 was the most significant SNP in both the samples of European ($p = 5.08 \times 10^{-4}$, OR = 1.39) and African American ($p = 2.62 \times 10^{-3}$, OR = 1.26) ancestry ($p_{meta\text{-}all} = 4.83 \times 10^{-9}$). The most significant SNP in this region (rs1453560) is located between *IKZF3* and *ZPBP2* and was replicated in European ($p = 6.42 \times 10^{-4}$, OR = 1.37) and African American ($p = 4.86 \times 10^{-4}$, OR = 1.23) ancestral populations; the replication resulted in $p_{meta\text{-}all} = 3.48 \times 10^{-10}$ (Table 1 and Table S6). All four SNPs are highly correlated ($r^2 > 0.95$). Even though Cochran's Q test of heterogeneity was not statistically significant, we observed moderate heterogeneity by the I^2 index, perhaps as a result of the differences in allele frequency between the racial groups (Table 1). *IKZF3* and *ZPBP2* are transcribed in opposite directions of one another but share the same promoter region (Figure S5). The ENCODE ChIP-Seq project has identified multiple transcription-factor binding sites for chromatin in the chromosomal region surrounding rs1453560 (Figure S5).[49]

In addition to the three regions (described above) that now exceed genome-wide significance, 11 loci were replicated in the European SLE cases but did not exceed genome-wide significance ($5 \times 10^{-8} < p_{meta\text{-}Euro} < 9.99 \times 10^{-5}$). These 11 loci include the following: *CFHR1* (MIM 134371), *CADM2, LOC730109/IL12A* (MIM 161560), *LPP* (MIM 600700), *LOC63920, SLU7* (MIM 605974), *ADAMTSL1* (MIM 609198), *C10orf64, OR8D4, FAM19A2,* and *STXBP6* (MIM 607958) (Table 3 and Table S7).

Discussion

The interferon regulatory factors are a family of transcription factors that play a critical role in the regulation of several pathways, including the response to pathogens, apoptosis, the cell cycle, and hematopoietic differentiation.[50] IRF8 is expressed in the nucleus (but partially in the cytoplasm) of B cells, macrophages, and CD11b dendritic cells (DCs).[50] IRF8 can be induced by interferon-γ in macrophages and antigen stimulation within T cells. It also plays an important role in the development of B cells and macrophages.[50] In the nucleus, IRF8 is required for promoting type I interferon responses in DCs upon viral stimulation.[50] Interestingly, the overexpression of genes induced by type I interferons has been widely reported in SLE and other autoimmune conditions.[12,51,52] In the cytosol, IRF8 is involved in the TLR9-MyD88-dependent signaling by binding to TRAF6 in both DCs and macrophages.[50] After TLR9 stimulation, DCs from mice that are $Irf8^{-/-}$ cannot activate NF-κB or MAPKs.[50] Of note, rs17445836, which was not included

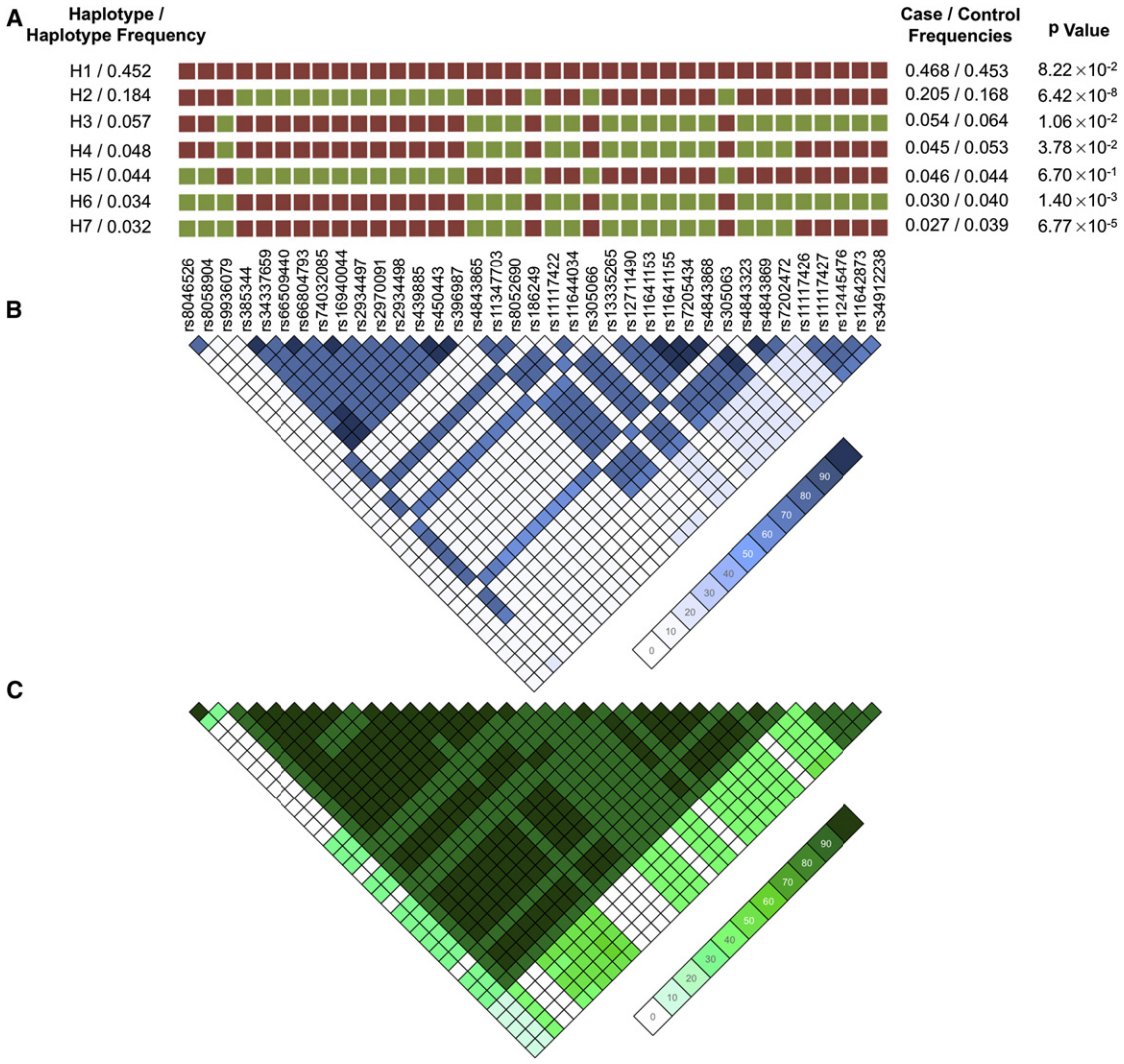

Figure 3. *IRF8* Haplotype and LD in the European Ancestral Population
(A) Haplotype structure in Europeans present at a frequency >3%. Major alleles are represented by red squares, whereas the green squares are minor alleles.
(B and C) A LD plot of r^2 (B) and D′ (C) in Europeans illustrates that the variants tagged by rs450443 and those tagged by rs4843869 are in weak r^2 but strong D′, providing evidence that these variants are inherited together.

in our study but has been associated with multiple sclerosis (MIM 126200), lies approximately 61 kb telomeric to *IRF8* and is far removed from the regions identified in SLE.[53]

Fine mapping, resequencing, imputation, and haplotype analysis of the *IRF8* locus in Europeans identified a single haplotype requiring the presence of three independent effects to confer risk. Additionally, several variants within the *IRF8* risk haplotype might influence binding to the many regulatory elements present within the region. Thus, we hypothesize that the likely functional effect would result in altered IRF8 mRNA and protein expression. Although we believe that most of the common variation within the *IRF8* region has been evaluated in this study, it is possible that some variants with minor allele frequencies <1% also play a role in SLE risk but were not detected in our study because of the number of samples resequenced.

The *TMEM39A*-associated coding SNP (rs1132200) results in an amino acid change from alanine to threonine at position 487 of the protein. Although almost no biological data have been published suggesting its relevance to SLE, it has been found to be associated with multiple sclerosis.[54] Better understanding whether the coding SNP in *TMEM39A* is functionally relevant or is merely correlated with other unexamined causal polymorphism(s) will require mechanistic and fine-mapping experiments.

Although the region surrounding the *IKZF3-ZPBP2* locus at 17q21 has been associated with multiple phenotypes, the extensive LD in the region has prohibited investigators from clearly determining the relevant gene. Crohn disease (MIM 266600), ulcerative colitis (MIM 266600), primary biliary cirrhosis (MIM 109720), and rheumatoid arthritis (MIM 180300) all have reported associations

Table 3. Replicated Loci that Demonstrate Suggestive Evidence of SLE Risk[a]

			European					Test of Heterogeneity	
SNP	Locus	Alleles[b]	$p_{GWA\ scan}$[c]	$OR_{GWA\ scan}$ (95% CI)	p_{REP}	OR_{REP} (95% CI)	$p_{META\text{-}Euro}$	p_Q[d]	I^2
rs7542235	*CFHR1*	A/G	3.94×10^{-3}	1.30 (1.11–1.54)	1.10×10^{-3}	1.15 (1.06–1.25)	1.85×10^{-5}	0.180	44.4%
rs485499	*LOC730109/ IL12A*	A/G	2.14×10^{-3}	0.75 (0.65–0.87)	1.47×10^{-4}	0.87 (0.81–0.94)	1.31×10^{-6}	0.076	68.2%
rs669003	*LOC730109/ IL12A*	A/G	2.16×10^{-3}	0.75 (0.65–0.87)	1.15×10^{-4}	0.87 (0.81–0.93)	1.02×10^{-6}	0.081	67.2%
rs7631930	*LPP*	A/G	3.60×10^{-3}	1.25 (1.06–1.49)	1.66×10^{-3}	1.15 (1.05–1.25)	2.71×10^{-5}	0.375	0.00%
rs9310002	*CADM2*	G/A	2.09×10^{-3}	2.06 (1.38–3.07)	6.12×10^{-3}	1.39 (1.10–1.76)	8.30×10^{-5}	0.099	63.3%
rs1075059	*LOC63920*	A/C	7.30×10^{-4}	0.82 (0.71–0.95)	7.26×10^{-3}	0.91 (0.85–0.97)	5.27×10^{-5}	0.221	33.2%
rs1895321	*SLU7*	A/C	4.21×10^{-3}	1.22 (1.06–1.41)	3.11×10^{-3}	1.11 (1.04–1.19)	6.09×10^{-5}	0.260	21.2%
rs7039790	*ADAMTSL1*	C/A	5.36×10^{-3}	1.62 (1.24–2.12)	1.14×10^{-3}	1.27 (1.10–1.47)	2.38×10^{-5}	0.124	57.7%
rs2940712	*C10orf64*	G/A	4.72×10^{-3}	0.79 (0.67–0.91)	8.73×10^{-4}	0.88 (0.82–0.95)	1.62×10^{-5}	0.178	45.0%
rs10790605	*OR8D4*	G/A	2.02×10^{-3}	0.80 (0.67–0.95)	4.35×10^{-3}	0.88 (0.81–0.96)	5.39×10^{-5}	0.321	0.00%
rs7960162	*FAM19A2*	A/G	4.95×10^{-3}	0.76 (0.61–0.94)	4.31×10^{-3}	0.87 (0.79–0.96)	9.79×10^{-5}	0.253	23.6%
rs749373	*STXBP6*	A/G	5.44×10^{-3}	1.34 (1.11–1.62)	2.32×10^{-3}	1.16 (1.05–1.27)	5.26×10^{-5}	0.171	46.7%

The following abbreviation is used: GWA, genome-wide association; OR, odds ratio; CI, confidence interval; and REP, replication.
[a]Table S7 contains results for all populations evaluated within this study.
[b]Major/minor alleles.
[c]GWA scan previously reported in Graham et al.[1]
[d]Cochran's Q test statistic.

with genes between 34.62–35.51 Mb of chromosome 17.[55–60] Fine mapping and resequencing of this region in Europeans and African Americans are needed if researchers are to more precisely refine this association and determine the loci associated with risk. *IKZF3* is a member of the IKAROS family of transcription factors involved in lymphocyte development; *IKZF1* in this family has already been reported as a risk locus for SLE.[22] Mice with a mutant form of IKZF3 produce anti-dsDNA autoantibodies, making it an interesting candidate gene for human SLE.[61] Moreover, mice that are null for *IKZF3* and *OBF-1* (POU class 2 associating factor 1) do not mount an autoimmune response.[61] The peak signal in our study was in a region containing multiple regulatory elements, so it is likely that the associated SNP could affect expression of *IKZF3* or *ZPBP2,* which both share the promoter region. However, no known function of *ZPBP2* has been reported.

Eleven additional regions were replicated in the European subjects but did not surpass genome-wide significance. Of these regions, *LOC730108/IL12A* was previously reported as a risk locus for primary biliary cirrhosis and multiple sclerosis.[60,62] IL-12A induces interferon-gamma and helps differentiate Th1 and Th2 cells.[63] The response of lymphocytes to IL-12A is mediated by STAT4, which is also implicated in SLE pathogenesis.[64] The LIM domain containing preferred translocation partner in lipoma (LPP) is involved in focal adhesions, cell-cell adhesion, and cell motility. Variants within the *LPP* region have been associated with vitiligo and celiac disease.[65,66] Confirmation of these associations will require replication in a larger independent and equally diverse population.

In conclusion, we have robustly established three additional susceptibility loci for SLE: *IRF8, TMEM39A,* and *IKZF3-ZPBP2*. Eleven other regions were replicated but did not exceed the genome-wide threshold of significance. Collectively, these data, along with other previously reported loci, demonstrate the growing complexity of the heritable contribution to SLE pathogenesis. A complete understanding of how genetics influence the pathophysiology of SLE will only be possible once we have identified all contributing loci and functional and/or causal variants for each association and have extensively evaluated the role of rare variants. More work will be required if we are to increase our understanding of how the loci identified in this study influence SLE etiology.

Supplemental Data

Supplemental Data include supplemental acknowledgments, six figures, and seven tables and can be found with this article online at http://www.cell.com/AJHG.

Acknowledgments

We are grateful to all the individuals with SLE and those serving as healthy controls who participated in this study. We thank the following individuals for contributing samples: Sandra Marc Bijl, D'Alfonso, Emoke Endreffy, Iñigo Rúa-Figueroa, Cintia Garcilazo, Carmen Gutierrez, Peter Junker, Helle Laustrup, Rafaella Scorza,

Berta Martins da Silva, Ana Suarez, and Carlos Vasconcelos. For the GENLES collaboration, we thank Eduardo Acevedo, Mario Cardiel, Ignacio García de la Torre, Mabel Busajm, Cecilia Castel, Marco Maradiaga, José F. Moctezuma, and Jorge Musuruana. For the Asociación Andaluza de Enfermedades Autoimmunes collaboration, we thank Juan Jimenez-Alonso, Norberto Ortego-Centeno, Enrique de Ramon, and Julio Sanchez-Roman. We would like to thank Summer Frank and Mei Li Zhu for their assistance in genotyping, quality-control analyses, and clinical data management. We would also like to thank Emily Cole for her assistance in preparing figures. Grant support information is provided in the Supplemental Acknowledgments available online.

Received: October 26, 2011
Revised: February 22, 2012
Accepted: February 22, 2012
Published online: March 29, 2012

Web Resources

The URLs for data presented herein are as follows:

dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/index.html
Encyclopedia of DNA Elements (ENCODE), http://genome.ucsc.edu/ENCODE/
HUGO Gene Nomenclature Committee, http://www.genenames.org/
LocusZoom, http://csg.sph.umich.edu/locuszoom/
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org
UCSC Genome Browser, http://genome.ucsc.edu/
VCFtools, http://www.vcftools.sourceforge.net

References

1. Petri, M. (2002). Epidemiology of systemic lupus erythematosus. Best Pract. Res. Clin. Rheumatol. *16*, 847–858.
2. Moser, K.L., Kelly, J.A., Lessard, C.J., and Harley, J.B. (2009). Recent insights into the genetic basis of systemic lupus erythematosus. Genes Immun. *10*, 373–379.
3. Alarcón-Segovia, D., Alarcón-Riquelme, M.E., Cardiel, M.H., Caeiro, F., Massardo, L., Villa, A.R., and Pons-Estel, B.A.; Grupo Latinoamericano de Estudio del Lupus Eritematoso (GLADEL). (2005). Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort. Arthritis Rheum. *52*, 1138–1147.
4. Anaya, J.M., Tobon, G.J., Vega, P., and Castiblanco, J. (2006). Autoimmune disease aggregation in families with primary Sjögren's syndrome. J. Rheumatol. *33*, 2227–2234.
5. Arora-Singh, R.K., Assassi, S., del Junco, D.J., Arnett, F.C., Perry, M., Irfan, U., Sharif, R., Mattar, T., and Mayes, M.D. (2010). Autoimmune diseases and autoantibodies in the first degree relatives of patients with systemic sclerosis. J. Autoimmun. *35*, 52–57.
6. Sestak, A.L., Shaver, T.S., Moser, K.L., Neas, B.R., and Harley, J.B. (1999). Familial aggregation of lupus and autoimmunity in an unusual multiplex pedigree. J. Rheumatol. *26*, 1495–1499.
7. Deng, Y., and Tsao, B.P. (2010). Genetic susceptibility to systemic lupus erythematosus in the genomic era. Nat. Rev. Rheumatol. *6*, 683–692.
8. Reinertsen, J.L., Klippel, J.H., Johnson, A.H., Steinberg, A.D., Decker, J.L., and Mann, D.L. (1978). B-lymphocyte alloantigens associated with systemic lupus erythematosus. N. Engl. J. Med. *299*, 515–518.
9. Nies, K.M., Brown, J.C., Dubois, E.L., Quismorio, F.P., Friou, G.J., and Terasaki, P.I. (1974). Histocompatibility (HL-A) antigens and lymphocytotoxic antibodies in systemic lupus erythematosus (SLE). Arthritis Rheum. *17*, 397–402.
10. Graham, R.R., Ortmann, W.A., Langefeld, C.D., Jawaheer, D., Selby, S.A., Rodine, P.R., Baechler, E.C., Rohlf, K.E., Shark, K.B., Espe, K.J., et al. (2002). Visualizing human leukocyte antigen class II risk haplotypes in human systemic lupus erythematosus. Am. J. Hum. Genet. *71*, 543–553.
11. McCulloch, D.K., Klaff, L.J., Kahn, S.E., Schoenfeld, S.L., Greenbaum, C.J., Mauseth, R.S., Benson, E.A., Nepom, G.T., Shewey, L., and Palmer, J.P. (1990). Nonprogression of subclinical beta-cell dysfunction among first-degree relatives of IDDM patients. 5-yr follow-up of the Seattle Family Study. Diabetes *39*, 549–556.
12. Baechler, E.C., Batliwalla, F.M., Karypis, G., Gaffney, P.M., Ortmann, W.A., Espe, K.J., Shark, K.B., Grande, W.J., Hughes, K.M., Kapur, V., et al. (2003). Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proc. Natl. Acad. Sci. USA *100*, 2610–2615.
13. Bennett, L., Palucka, A.K., Arce, E., Cantrell, V., Borvak, J., Banchereau, J., and Pascual, V. (2003). Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. J. Exp. Med. *197*, 711–723.
14. Kirou, K.A., Lee, C., George, S., Louca, K., Papagiannis, I.G., Peterson, M.G., Ly, N., Woodward, R.N., Fry, K.E., Lau, A.Y., et al. (2004). Coordinate overexpression of interferon-alpha-induced genes in systemic lupus erythematosus. Arthritis Rheum. *50*, 3958–3967.
15. Sigurdsson, S., Nordmark, G., Göring, H.H., Lindroos, K., Wiman, A.C., Sturfelt, G., Jönsen, A., Rantapää-Dahlqvist, S., Möller, B., Kere, J., et al. (2005). Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. Am. J. Hum. Genet. *76*, 528–537.
16. Hom, G., Graham, R.R., Modrek, B., Taylor, K.E., Ortmann, W., Garnier, S., Lee, A.T., Chung, S.A., Ferreira, R.C., Pant, P.V., et al. (2008). Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. N. Engl. J. Med. *358*, 900–909.
17. Harley, J.B., Alarcón-Riquelme, M.E., Criswell, L.A., Jacob, C.O., Kimberly, R.P., Moser, K.L., Tsao, B.P., Vyse, T.J., Langefeld, C.D., Nath, S.K., et al; International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEN). (2008). Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat. Genet. *40*, 204–210.
18. Yang, W., Shen, N., Ye, D.Q., Liu, Q., Zhang, Y., Qian, X.X., Hirankarn, N., Ying, D., Pan, H.F., Mok, C.C., et al; Asian Lupus Genetics Consortium. (2010). Genome-wide association study in Asian populations identifies variants in ETS1 and WDFY4 associated with systemic lupus erythematosus. PLoS Genet. *6*, e1000841.
19. Graham, R.R., Cotsapas, C., Davies, L., Hackett, R., Lessard, C.J., Leon, J.M., Burtt, N.P., Guiducci, C., Parkin, M., Gates, C., et al. (2008). Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. Nat. Genet. *40*, 1059–1061.

20. Kozyrev, S.V., Abelson, A.K., Wojcik, J., Zaghlool, A., Linga Reddy, M.V., Sanchez, E., Gunnarsson, I., Svenungsson, E., Sturfelt, G., Jönsen, A., et al. (2008). Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nat. Genet. *40*, 211–216.
21. Han, J.W., Zheng, H.F., Cui, Y., Sun, L.D., Ye, D.Q., Hu, Z., Xu, J.H., Cai, Z.M., Huang, W., Zhao, G.P., et al. (2009). Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. Nat. Genet. *41*, 1234–1237.
22. Gateva, V., Sandling, J.K., Hom, G., Taylor, K.E., Chung, S.A., Sun, X., Ortmann, W., Kosoy, R., Ferreira, R.C., Nordmark, G., et al. (2009). A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. Nat. Genet. *41*, 1228–1233.
23. Lessard, C.J., Adrianto, I., Kelly, J.A., Kaufman, K.M., Grundahl, K.M., Adler, A., Williams, A.H., Gallant, C.J., Anaya, J.M., Bae, S.C., et al; Marta E. Alarcón-Riquelme on behalf of the BIOLUPUS and GENLES Networks. (2011). Identification of a systemic lupus erythematosus susceptibility locus at 11p13 between PDHX and CD44 in a multiethnic study. Am. J. Hum. Genet. *88*, 83–91.
24. Namjou, B., Kothari, P.H., Kelly, J.A., Glenn, S.B., Ojwang, J.O., Adler, A., Alarcón-Riquelme, M.E., Gallant, C.J., Boackle, S.A., Criswell, L.A., et al. (2011). Evaluation of the TREX1 gene in a large multi-ancestral lupus cohort. Genes Immun. *12*, 270–279.
25. Adrianto, I., Wen, F., Templeton, A., Wiley, G., King, J.B., Lessard, C.J., Bates, J.S., Hu, Y., Kelly, J.A., Kaufman, K.M., et al; BIOLUPUS and GENLES Networks. (2011). Association of a functional variant downstream of TNFAIP3 with systemic lupus erythematosus. Nat. Genet. *43*, 253–258.
26. Tan, W., Sunahori, K., Zhao, J., Deng, Y., Kaufman, K.M., Kelly, J.A., Langefeld, C.D., Williams, A.H., Comeau, M.E., Ziegler, J.T., et al; BIOLUPUS Network; GENLES Network. (2011). Association of PPP2CA polymorphisms with systemic lupus erythematosus susceptibility in multiple ethnic groups. Arthritis Rheum. *63*, 2755–2763.
27. Zhao, J., Wu, H., Khosravi, M., Cui, H., Qian, X., Kelly, J.A., Kaufman, K.M., Langefeld, C.D., Williams, A.H., Comeau, M.E., et al; BIOLUPUS Network; GENLES Network. (2011). Association of genetic variants in complement factor H and factor H-related genes with systemic lupus erythematosus susceptibility. PLoS Genet. *7*, e1002079.
28. Sanchez, E., Nadig, A., Richardson, B.C., Freedman, B.I., Kaufman, K.M., Kelly, J.A., Niewold, T.B., Kamen, D.L., Gilkeson, G.S., Ziegler, J.T., et al; BIOLUPUS and GENLES. (2011). Phenotypic associations of genetic susceptibility loci in systemic lupus erythematosus. Ann. Rheum. Dis. *70*, 1752–1757.
29. Hochberg, M.C. (1997). Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum. *40*, 1725.
30. Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2005). Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics *21*, 263–265.
31. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909.
32. McKeigue, P.M., Carpenter, J.R., Parra, E.J., and Shriver, M.D. (2000). Estimation of admixture and detection of linkage in admixed populations by a Bayesian approach: Application to African-American populations. Ann. Hum. Genet. *64*, 171–186.
33. Hoggart, C.J., Parra, E.J., Shriver, M.D., Bonilla, C., Kittles, R.A., Clayton, D.G., and McKeigue, P.M. (2003). Control of confounding of genetic associations in stratified populations. Am. J. Hum. Genet. *72*, 1492–1504.
34. Hoggart, C.J., Shriver, M.D., Kittles, R.A., Clayton, D.G., and McKeigue, P.M. (2004). Design and analysis of admixture mapping studies. Am. J. Hum. Genet. *74*, 965–978.
35. Smith, M.W., Patterson, N., Lautenberger, J.A., Truelove, A.L., McDonald, G.J., Waliszewska, A., Kessing, B.D., Malasky, M.J., Scafe, C., Le, E., et al. (2004). A high-density admixture map for disease gene discovery in african americans. Am. J. Hum. Genet. *74*, 1001–1013.
36. Halder, I., Shriver, M., Thomas, M., Fernandez, J.R., and Frudakis, T. (2008). A panel of ancestry informative markers for estimating individual biogeographical ancestry and admixture from four continents: Utility and applications. Hum. Mutat. *29*, 648–658.
37. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.
38. Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: Fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190–2191.
39. Cochran, W.G. (1954). The Combination of Estimates from Different Experiments. Biometrics *10*, 101–129.
40. Higgins, J.P., Thompson, S.G., Deeks, J.J., and Altman, D.G. (2003). Measuring inconsistency in meta-analyses. BMJ *327*, 557–560.
41. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.
42. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.
43. Browning, S.R., and Browning, B.L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am. J. Hum. Genet. *81*, 1084–1097.
44. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.
45. Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M., et al; International HapMap Consortium. (2007). A second generation human haplotype map of over 3.1 million SNPs. Nature *449*, 851–861.
46. Howie, B.N., Donnelly, P., and Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. *5*, e1000529.
47. Via, M., Gignoux, C., and Burchard, E.G. (2010). The 1000 Genomes Project: New opportunities for research and social challenges. Genome Med. *2*, 3.
48. 1000 Genomes Project Consortium. (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061–1073.

49. Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., et al; ENCODE Project Consortium; NISC Comparative Sequencing Program; Baylor College of Medicine Human Genome Sequencing Center; Washington University Genome Sequencing Center; Broad Institute; Children's Hospital Oakland Research Institute. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799–816.
50. Tamura, T., Yanai, H., Savitsky, D., and Taniguchi, T. (2008). The IRF family transcription factors in immunity and oncogenesis. Annu. Rev. Immunol. *26*, 535–584.
51. Baechler, E.C., Gregersen, P.K., and Behrens, T.W. (2004). The emerging role of interferon in human systemic lupus erythematosus. Curr. Opin. Immunol. *16*, 801–807.
52. Baechler, E.C., Batliwalla, F.M., Reed, A.M., Peterson, E.J., Gaffney, P.M., Moser, K.L., Gregersen, P.K., and Behrens, T.W. (2006). Gene expression profiling in human autoimmunity. Immunol. Rev. *210*, 120–137.
53. De Jager, P.L., Jia, X., Wang, J., de Bakker, P.I., Ottoboni, L., Aggarwal, N.T., Piccio, L., Raychaudhuri, S., Tran, D., Aubin, C., et al; International MS Genetics Consortium. (2009). Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. Nat. Genet. *41*, 776–782.
54. International Multiple Sclerosis Genetics Consortium (IMSGC). (2010). Comprehensive follow-up of the first genome-wide association study of multiple sclerosis identifies KIF21B and TMEM39A as susceptibility loci. Hum. Mol. Genet. *19*, 953–962.
55. Franke, A., McGovern, D.P., Barrett, J.C., Wang, K., Radford-Smith, G.L., Ahmad, T., Lees, C.W., Balschun, T., Lee, J., Roberts, R., et al. (2010). Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat. Genet. *42*, 1118–1125.
56. Liu, X., Invernizzi, P., Lu, Y., Kosoy, R., Lu, Y., Bianchi, I., Podda, M., Xu, C., Xie, G., Macciardi, F., et al. (2010). Genome-wide meta-analyses identify three loci associated with primary biliary cirrhosis. Nat. Genet. *42*, 658–660.
57. Hirschfield, G.M., Liu, X., Han, Y., Gorlov, I.P., Lu, Y., Xu, C., Lu, Y., Chen, W., Juran, B.D., Coltescu, C., et al. (2010). Variants at IRF5-TNPO3, 17q12-21 and MMEL1 are associated with primary biliary cirrhosis. Nat. Genet. *42*, 655–657.
58. Stahl, E.A., Raychaudhuri, S., Remmers, E.F., Xie, G., Eyre, S., Thomson, B.P., Li, Y., Kurreeman, F.A., Zhernakova, A., Hinks, A., et al; BIRAC Consortium; YEAR Consortium. (2010). Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat. Genet. *42*, 508–514.
59. Anderson, C.A., Boucher, G., Lees, C.W., Franke, A., D'Amato, M., Taylor, K.D., Lee, J.C., Goyette, P., Imielinski, M., Latiano, A., et al. (2011). Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nat. Genet. *43*, 246–252.
60. Hirschfield, G.M., Liu, X., Xu, C., Lu, Y., Xie, G., Lu, Y., Gu, X., Walker, E.J., Jing, K., Juran, B.D., et al. (2009). Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. N. Engl. J. Med. *360*, 2544–2555.
61. Sun, J., Matthias, G., Mihatsch, M.J., Georgopoulos, K., and Matthias, P. (2003). Lack of the transcriptional coactivator OBF-1 prevents the development of systemic lupus erythematosus-like phenotypes in Aiolos mutant mice. J. Immunol. *170*, 1699–1706.
62. International Multiple Sclerosis Genetics Conssortium (IMSGC). (2010). IL12A, MPHOSPH9/CDK2AP1 and RGS1 are novel multiple sclerosis susceptibility loci. Genes Immun. *11*, 397–405.
63. Peluso, I., Pallone, F., and Monteleone, G. (2006). Interleukin-12 and Th1 immune response in Crohn's disease: Pathogenetic relevance and therapeutic implication. World J. Gastroenterol. *12*, 5606–5610.
64. Remmers, E.F., Plenge, R.M., Lee, A.T., Graham, R.R., Hom, G., Behrens, T.W., de Bakker, P.I., Le, J.M., Lee, H.S., Batliwalla, F., et al. (2007). STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N. Engl. J. Med. *357*, 977–986.
65. Jin, Y., Birlea, S.A., Fain, P.R., Gowan, K., Riccardi, S.L., Holland, P.J., Mailloux, C.M., Sufit, A.J., Hutton, S.M., Amadi-Myers, A., et al. (2010). Variant of TYR and autoimmunity susceptibility loci in generalized vitiligo. N. Engl. J. Med. *362*, 1686–1697.
66. Hunt, K.A., Zhernakova, A., Turner, G., Heap, G.A., Franke, L., Bruinenberg, M., Romanos, J., Dinesen, L.C., Ryan, A.W., Panesar, D., et al. (2008). Newly identified genetic risk variants for celiac disease related to the immune response. Nat. Genet. *40*, 395–402.