ARTHRITIS & RHEUMATOLOGY
Vol. 68, No. 5, May 2016, pp 1197–1209
DOI 10.1002/art.39548

Identification of a Systemic Lupus Erythematosus Risk Locus Spanning *ATG16L2, FCHSD2*, and *P2RY2* in Koreans

Christopher J. Lessard,[1] Satria Sajuthi,[2] Jian Zhao,[3] Kwangwoo Kim,[4] John A. Ice,[1] He Li,[5] Hannah Ainsworth,[2] Astrid Rasmussen,[1] Jennifer A. Kelly,[1] Miranda Marion,[2] So-Young Bang,[4] Young Bin Joo,[4] Jeongim Choi,[4] Hye-Soon Lee,[4] Young Mo Kang,[6] Chang-Hee Suh,[7] Won Tae Chung,[8] Soo-Kon Lee,[9] Jung-Yoon Choe,[10] Seung Cheol Shim,[11] Ji Hee Oh,[12] Young Jin Kim,[12] Bok-Ghee Han,[12] Nan Shen,[13] Hwee Siew Howe,[14] Edward K. Wakeland,[15] Quan-Zhen Li,[15] Yeong Wook Song,[16] Patrick M. Gaffney,[1] Marta E. Alarcón-Riquelme,[17] Lindsey A. Criswell,[18] Chaim O. Jacob,[19] Robert P. Kimberly,[20] Timothy J. Vyse,[21] John B. Harley,[22] Kathy L. Sivils,[5] Sang-Cheol Bae,[4] Carl D. Langefeld,[2] and Betty P. Tsao[3]

Objective. **Systemic lupus erythematosus (SLE) is a chronic autoimmune disorder whose etiology is incompletely understood, but likely involves environmental triggers in genetically susceptible individuals. Using an unbiased genome-wide association (GWA) scan and replication analysis, we sought to identify the genetic loci associated with SLE in a Korean population.**

Methods. **A total of 1,174 SLE cases and 4,246 population controls from Korea were genotyped and analyzed with a GWA scan to identify single-nucleotide polymorphisms (SNPs) significantly associated with SLE, after strict quality control measures were applied. For select variants, replication of SLE risk loci was tested in an independent data set of 1,416 SLE cases and 1,145 population controls from Korea and China.**

Results. **Eleven regions outside the HLA exceeded the genome-wide significance level ($P = 5 \times 10^{-8}$). A**

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

Supported by grants from the NIH (5P01-AI-083194 to Drs. Lessard, Rasmussen, Gaffney, Alarcón-Riquelme, Criswell, Jacob, Kimberly, Vyse, Harley, Sivils, Langefeld, and Tsao; 5P01-AR-049084 to Drs. Kimberly and Langefeld; and R01-AR-043814 and R21-AR-065626 to Dr. Tsao), the Oklahoma Medical Research Foundation (to Drs. Lessard, Gaffney, and Sivils), the Alliance for Lupus Research (to Drs. Criswell and Tsao), and the Wake Forest School of Medicine Center for Public Health Genomics (to Dr. Langefeld). Dr. Criswell is recipient of a Kirkland Scholar Award. Sample collection and phenotyping of the subjects utilized in this study was funded by the Ministry for Health and Welfare, Republic of Korea (Korea Healthcare Technology R&D Project grant HI13C2124 to Dr. Bae, and HI13C1754 to Dr. Song). The out-of-study Korean control data were provided by the Korean Biobank Project, which is supported by the Korea Centers for Disease Control and Prevention at the Korea National Institute of Health.

[1]Christopher J. Lessard, PhD, John A. Ice, MD, Astrid Rasmussen, MD, PhD, Jennifer A. Kelly, MPH, Patrick M. Gaffney, MD: Oklahoma Medical Research Foundation, Oklahoma City; [2]Satria Sajuthi, MS, Hannah Ainsworth, BSc, Miranda Marion, MA, Carl D. Langefeld, PhD: Wake Forest University Health Sciences, Winston-Salem, North Carolina; [3]Jian Zhao, PhD, Betty P. Tsao, PhD: University of California, Los Angeles; [4]Kwangwoo Kim, PhD, So-Young Bang, MD, PhD, Young Bin Joo, MD, PhD, Jeongim Choi, MD, PhD, Hye-Soon Lee, MD, PhD, Sang-Cheol Bae, MD, PhD, MPH: Hanyang University Hospital for Rheumatic Diseases, Seoul, Republic of Korea; [5]He Li, PhD, Kathy L. Sivils, PhD: Oklahoma Medical Research Foundation and University of Oklahoma Health Sciences Center, Oklahoma City; [6]Young Mo Kang, MD, PhD: Kyungpook National University Hospital, Daegu, Republic of Korea; [7]Chang-Hee Suh, MD, PhD: Ajou University Hospital, Suwon, Republic of Korea; [8]Won Tae Chung, MD, PhD: Dong-A University Hospital, Busan, Republic of Korea; [9]Soo-Kon Lee, MD, PhD: Yonsei University College of Medicine, Seoul, Republic of Korea; [10]Jung-Yoon Choe, MD, PhD: Catholic University of Daegu School of Medicine, Daegu, Republic of Korea; [11]Seung Cheol Shim, MD, PhD: Chungnam National University Hospital, Daejeon, Republic of Korea; [12]Ji Hee Oh, PhD, Young Jin Kim, PhD, Bok-Ghee Han, PhD: Korea National Institute of Health, Osong, Republic of Korea; [13]Nan Shen, MD: Renji Hospital and Shanghai JiaoTong University School of Medicine, Shanghai, China; [14]Hwee Siew Howe, MBBS: Tan Tock Seng Hospital, Singapore; [15]Edward K. Wakeland, PhD, Quan-Zhen Li, MD, PhD: University of Texas Southwestern Medical Center, Dallas; [16]Yeong Wook Song, MD: Seoul National University Hospital, Seoul, Republic of Korea; [17]Marta E. Alarcón-Riquelme, MD, PhD: Oklahoma Medical Research Foundation, Oklahoma City, and Centro Pfizer–Universidad de Granada–Junta de Andalucía de Genómica e Investigaciones Oncológicas, Granada, Spain; [18]Lindsey A. Criswell, MD, MPH, DSc: University of California, San Francisco; [19]Chaim O. Jacob, MD, PhD: University of Southern California, Los Angeles; [20]Robert P. Kimberly, MD: University of Alabama at Birmingham; [21]Timothy J. Vyse, MBBS, FRCP, PhD, FMedSci: King's College London, London, UK; [22]John B. Harley, MD, PhD: Cincinnati Children's Hospital Medical Center and Cincinnati VA Medical Center, Cincinnati, Ohio.

novel SNP–SLE association was identified between *FCHSD2* and *P2RY2*, peaking at rs11235667 ($P = 1.03 \times 10^{-8}$, odds ratio [OR] 0.59) on a 33-kb haplotype upstream of *ATG16L2*. In the independent replication data set, the SNP rs11235667 continued to show a significant association with SLE (replication meta-analysis $P = 0.001$, overall meta-analysis $P = 6.67 \times 10^{-11}$; OR 0.63). Within the HLA region, the SNP–SLE association peaked in the class II region at rs116727542, with multiple independent effects observed in this region. Classic HLA allele imputation analysis identified *HLA–DRB1*1501* and *HLA–DQB1*0602*, each highly correlated with one another, as most strongly associated with SLE. Ten previously established SLE risk loci were replicated: *STAT1–STAT4*, *TNFSF4*, *TNFAIP3*, *IKZF1*, *HIP1*, *IRF5*, *BLK*, *WDFY4*, *ETS1*, and *IRAK1–MECP2*. Of these loci, previously unreported, independent second risk effects of SNPs in *TNFAIP3* and *TNFSF4*, as well as differences in the association with a putative causal variant in the *WDFY4* region, were identified.

Conclusion. **Further studies are needed to identify true SLE risk effects in other loci suggestive of a significant association, and to identify the causal variants in the regions of *ATG16L2*, *FCHSD2*, and *P2RY2*.**

Systemic lupus erythematosus (SLE; MIM ID 152700) is a chronic, heterogeneous autoimmune disease characterized by the loss of tolerance to self antigens, dysregulated type I interferon responses, and inflammation, often resulting in systemic end-organ damage (1). Immune dysfunction in SLE involves both B and T lymphocytes of the adaptive immune system, together with elements of the innate immune system, including dendritic cells and the complement system (1). The clinical manifestations of SLE can be quite variable and can involve virtually any organ system. Although the precise etiology of SLE is largely unknown, the pathogenic mechanism likely involves environmental triggers in a genetically susceptible host (2). Few effective treatment options exist, largely because understanding of the pathophysiologic basis of the disease is incomplete.

Drs. Lessard, Sajuthi, Zhao, and K. Kim contributed equally to this work. Drs. Sivils, Bae, Langefeld, and Tsao contributed equally to this work.

Address correspondence to Sang-Cheol Bae, MD, PhD, MPH, Department of Rheumatology, Hanyang University Hospital for Rheumatic Diseases, 222 Wangsimni-ro, Seondong-gu, Seoul 133-792, Republic of Korea (e-mail: scbae@hanyang.ac.kr); or to Betty P. Tsao, PhD, 936C Clinical Science Building, Medical University of South Carolina, 96 Jonathan Lucas Street, Charleston, SC 29425 (e-mail: tsaob@musc.edu).

Submitted for publication May 6, 2015; accepted in revised form December 8, 2015.

Genetic predisposition leading to an increased risk of SLE is supported by observations of high heritability (>66%), increased risk among siblings of affected patients (sibling risk ratio [λ_s] ~30), and an ~25% disease concordance rate in monozygotic twins (3). To date, associations of more than 50 loci with SLE susceptibility have been identified and confirmed (4). Many of these genes fall into known pathways that are key to innate and adaptive immune responses, lymphocyte activation and/or function, and immune complex clearance (4). However, a significant proportion of heritable SLE risk has yet to be explained (5). The identification of SLE-associated genes and their pathogenic mechanisms will greatly enhance our understanding of the pathophysiologic processes of lupus and facilitate the development of effective diagnostic, prognostic, and therapeutic tools.

To date, large-scale genome-wide genetic studies of SLE in Asian populations have focused on Han Chinese subjects (6–8) and Japanese subjects (9). Moreover, several studies have shown that transracial mapping of SLE loci can aid in the dissection of risk effects (4). In this study, we performed a genome-wide association (GWA) scan to identify genes associated with SLE in an East Asian population from Korea.

PATIENTS AND METHODS

Subjects. A total of 1,174 patients with SLE were recruited from the Hanyang University Hospital for Rheumatic Diseases (HUHRD) and 6 other university hospitals in Korea (10). In addition, 552 ethnically matched healthy control subjects were recruited from the HUHRD, and 3,700 ethnically matched out-of-study population control subjects were recruited from the Korean National Institutes of Health (10). For replication studies, an independent cohort of 1,416 SLE cases and 1,145 population controls was used (11,12). This independent sample set consisted of 739 Korean SLE cases and 436 Korean controls, as well as 677 Chinese SLE cases and 709 Chinese controls (see Supplementary Table 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract).

Written informed consent was obtained from each participant, in accordance with the protocols approved by the Institutional Review Boards of the participating institutions. All cases used in this study fulfilled at least 4 of the 11 American College of Rheumatology revised criteria for SLE (13) as updated in 1997 (14), while healthy, population-based controls were subjects without a family history of SLE or any other autoimmune disease.

GWA analysis. *Genotyping, sample quality control, and ascertainment of population stratification.* Samples were genotyped using Illumina HumanOmni1-Quad or HumanOmniExpress arrays with Infinium chemistry, in accordance with the manufacturer's protocols, at the Oklahoma Medical Research Foundation (OMRF). The out-of-study population controls in the GWA analysis were genotyped using Illumina

HumanOmni1-Quad arrays performed at the Korea National Institutes of Health.

Strict quality control standards were implemented for the single-nucleotide polymorphisms (SNPs) retained in the GWA data set and tested for association, including the requirement for well-defined clusters in the calling of the genotypes. Samples were excluded if they had a SNP call rate of <90%. SNPs were considered high-quality SNPs if they had call rates of >95%, no evidence of differential missingness between cases and controls (at a significance level of $P < 0.05$), and no evidence of a departure from expected Hardy-Weinberg proportions (at a significance level of $P < 0.01$ for controls and $P < 0.000001$ for cases). Inference was based primarily on those SNPs that had a minor allele frequency (MAF) >1%.

Based on the SNPs that passed the above quality control thresholds, samples were removed if there were inconsistencies between recorded and genotype-inferred sex or if there was excess heterozygosity on the autosomes. Duplicates and first- or second-degree relatives were removed based on identity-by-descent statistics, computed using the program KING (15).

With the remaining study samples, principal components (PCs) were computed, and the data were then merged with HapMap phase III genotypes from the CEU (Utah residents with ancestry from northern and western Europe from the CEPH [Centre d'Étude du Polymorphisme Humain]), YRI (Yoruba from Ibadan, Nigeria), and CHB (Han Chinese from Beijing, China) reference populations (16), using the program EigenSoft (17). Principal components analysis (PCA) was performed on a subset of autosomal SNPs that were selected by removing regions of known high linkage disequilibrium (LD), by removing variants with an MAF <0.05, and by pruning markers to reduce extended pairwise LD. PCs 1–3 were plotted in these analyses, to remove genetic outliers (see Supplementary Figure 1, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). Thus, the final data set that passed laboratory and statistical quality controls was composed of 1,174 SLE cases (1,096 female patients and 78 male patients) and 548 within-study controls (547 female subjects and 1 male subject). In addition, genotype data for 3,698 out-of-study controls (2,330 female subjects and 1,368 male subjects) were merged with the within-study genotype data.

Statistical analysis. Logistic regression analysis was used to test for associations between each SNP and SLE status while adjusting for PC3 as a covariate, since no other PC significantly changed the inflation factor λ. Primary inference was based on the additive genetic model, unless there was significant lack-of-fit (at a significance level of $P < 0.05$). If there was evidence of a departure from an additive model, then inference was based on the most significant value from the dominant, additive, or recessive genetic model. The additive model and recessive model were computed only if there were at least 10 individuals and 30 individuals, respectively, who were homozygous for the minor allele. The analyses were completed using the program SNPGWA version 4.0 (available at http://www.phs.wfubmc.edu). For the analysis of chromosome X SNPs, the samples were stratified by sex, and the data were then meta-analyzed across the sexes using the program METAL (18).

To determine the number of independent associations within each SLE risk locus exceeding the genome-wide significance threshold, a manual stepwise model, or conditional analysis, was computed. The stepwise modeling, or conditional analysis, was implemented using forward selection with backward elimination, applying both an entry criterion and an exit criterion of $P < 0.0001$, which accounted for ~500 independent variants within a given genomic region. Specifically, for each region of interest, the top SNP was included as a covariate and the association statistics were recalculated. SNPs were allowed to be entered into and exited from the models in this stepwise manner until no additional SNPs met the significance threshold of $P < 0.0001$.

Replication analysis. Genotypes from the replication cohort were obtained using TaqMan assays (Life Technologies) for 4 SNPs: rs2267828, rs10901656, rs11235667, and rs1048257. Analyses were conducted for these cohorts independently, to allow for PCA using previously collected data. Ancestry adjustments for the Korean replication cohort have been described previously (11). For the Chinese subjects, the PCA was done using the method described in a study by Kaiser et al (12) with slight modification.

In the present study, 7,918 randomly selected autosomal ImmunoChip SNPs with an MAF >1%, low pairwise LD ($r^2 < 0.1$), and no evidence of association with SLE ($P > 0.01$) were used to perform PCA, carried out using the program EigenSoft (17). PCA plots utilizing the genotype data from our subjects as well as those from the 1000 Genomes Project reference populations of CHB and Southern Han Chinese (CHS) subjects were used to select and remove genetic outliers. PC1 (Chinese cohort) and PCs 1, 2, and 3 (Korean data sets from the OMRF and University of Southern California, Los Angeles) were included as covariates in the logistic regression models, based on the variance explained in each data set. These data sets were than meta-analyzed using the program METAL (18). To test for heterogeneity among the individual association results in the meta-analysis, we utilized both the Cochran's Q test statistic (19) and the I^2 index (20).

Imputation analysis. To help localize the associations in the regions meeting the genome-wide significance level, ungenotyped variants were imputed based on the reference panel from the 1000 Genomes Project (21). Specifically, the program SHAPEIT was used to pre-phase the genotype data (22). After phasing the data, IMPUTE2 was used for the imputation based on the 1000 Genomes Project phase I integrated reference panel (23). The imputed data were filtered using standard postimputation quality control, based on IMPUTE2 information scores >0.5 and confidence scores >0.9 for subsequent association tests. Post-association analysis required inclusion of genotyped SNPs in LD with the imputed variants, to support the inferred alleles as true signals. The program SNPTEST (version 2) (24) was used to test for association of the imputed variants.

Imputation of the classic HLA alleles in the genes *HLA–A*, *HLA–B*, *HLA–C*, *HLA–DPB1*, *HLA–DQA1*, *HLA–DQB1*, and *HLA–DRB1* was done using the program HiBAG (25), with imputation analysis based on the Asian reference panel. In this sample, ~21% of the reference SNPs used in the HiBAG imputation were missing genotype data. To address this issue, HLA imputation was repeated after filling in the missing genotype data with the "best guess" imputed SNP data from the 1000 Genomes Project data set described above. By using the "best guess" genotype data with a posterior probabil-

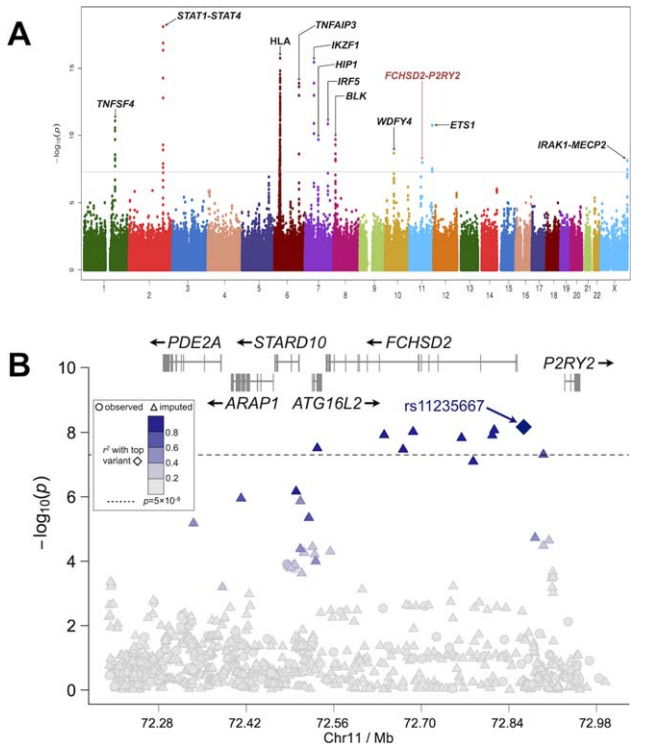

Figure 1. Summary of the results from the genome-wide association scan in 1,174 systemic lupus erythematosus (SLE) cases and 4,246 controls from Korea **(A)** and zoom plot of the region associated with SLE at 11q14 **(B)**. **A,** The Manhattan plot shows the $-\log_{10}$ (*P* values) for the association of each genotyped variant with SLE, according to chromosomal position. The shaded horizontal line indicates the genome-wide significance threshold of $P = 5 \times 10^{-8}$. **B,** The zoom plot shows the $-\log_{10}$ (*P* values) for the association of each genotyped variant (circles) and imputed variant (triangles), as well as for the variant showing peak association, rs11235667 (diamond), with SLE, according to basepair position on chromosome 11 (Chr11). Color key shows the values of linkage disequilibrium with rs11235667. The broken horizontal line indicates the genome-wide significance threshold of $P = 5 \times 10^{-8}$. Association exceeding this threshold was found in the region extending from *ATG16L2* through *FCHSD2* to the shared promoter region with *P2RY2*.

ity of >0.90, the percentage of missing variants in the reference set was reduced to 0.36%.

RESULTS

Summary findings of the GWA analysis. We observed modest inflation in the test statistic ($\lambda = 1.09$), with only slight deviation of observed *P* values from expected, once the HLA loci and other known SLE loci were removed (see Supplementary Figure 2, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). A total of 11 regions surpassed the genome-wide significance threshold of $P = 5 \times 10^{-8}$, with *STAT4* (MIM ID 600558) at rs11889341 yielding the most significant genotyped association with SLE ($P = 8.02 \times 10^{-19}$) (Figure 1A and Table 1). Of the non-HLA regions, 10 risk loci had been previously identified and confirmed as risk loci for SLE, including *STAT1* (MIM ID 600555)–*STAT4*, *IKZF1* (MIM ID 603023), *TNFAIP3* (MIM ID 191163), *TNFSF4* (MIM ID 603594), *HIP1* (MIM ID 601767), *IRF5* (MIM ID 607218), *ETS1* (MIM ID 164740), *BLK* (MIM ID 191305), *WDFY4* (MIM ID 613316), and *IRAK1* (MIM ID 300283)–*MECP2* (MIM ID 300005). In addition, a novel risk locus (i.e., not previously described as being associated with SLE) was observed at 11q14 (Figure 1A).

Association with SLE at 11q14. SNP–SLE association exceeding the genome-wide significance threshold was only observed for one genotyped variant, located between *FCHSD2* (MIM ID not available) and *P2RY2* (MIM ID 600041) ($P = 1.03 \times 10^{-8}$; odds ratio [OR] 0.59, 95% confidence interval [95% CI] 0.50–0.71) (Figure 1B and Table 2). Moreover, additional support from other genotyped variants in the region was observed but with $P > 5 \times 10^{-8}$ (a complete list of the variants identified by single-locus GWA analysis is available in Supplementary Table 2, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract).

After imputation of the 11q14 region showing association with SLE, rs11235667 remained the variant showing the most significant association (Figure 1B and Supplementary Table 2, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). However, a haplotype with 8 variants exceeding the genome-wide significance threshold was identified, spanning from *ATG16L2* (MIM ID not available) through *FCHSD2* to the shared promoter region with *P2RY2*. Stepwise logistic regression analysis adjusted for rs11235667 indicated that this variant accounted for all of the association in the region and likely the presence of only a single effect (see Supplementary Figure 3, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract).

Replication analysis of the primary signal in the region of *FCHSD2–P2RY2* was done using independent cohorts from Korea and China. The SNP rs11235667 between *FCHSD2* and *P2RY2* continued to show a significant association with SLE and had a similar effect size (replication meta-analysis *P* [$P_{\text{meta-rep}}$] = 0.001, OR 0.71, 95% CI 0.57–0.87). In the meta-analysis of the GWA and replication studies, the overall meta-analysis *P* was 6.67×10^{-11} (OR 0.63, 95% CI 0.55–0.72) (Table 2). No evidence of heterogeneity was observed in the meta-analysis (Table 2).

Bioinformatics database mining using HaploReg version 2 (26) identified that the region encompassing rs11235667 was hypersensitive to DNase I in B cells, as

Table 1. Single-locus analysis of previously identified systemic lupus erythematosus risk loci outside of the HLA region*

Marker	Chr.	Position	Upstream gene	Downstream gene	Within gene	MAF	P_{GWAS}	Model	Obs/Imp	OR (95% CI)	Major/minor allele
rs1234314	1	173177392	1 kb from *TNFSF4*	269 kb from *PRDX6*	–	0.37	9.25×10^{-11}	Add	Imp	1.37 (1.25–1.51)	C/G
rs2205960	1	173191475	15 kb from *TNFSF4*	255 kb from *PRDX6*	–	0.25	1.03×10^{-11}	Add	Imp	1.44 (1.30–1.60)	G/T
rs76413021	1	173206297	30 kb from *TNFSF4*	240 kb from *PRDX6*	–	0.23	3.26×10^{-13}	Add	Imp	1.52 (1.36–1.71)	G/A
rs844644	1	173209495	33 kb from *TNFSF4*	237 kb from *PRDX6*	–	0.4	4.47×10^{-11}	Add	Obs	1.37 (1.25–1.50)	A/C
rs10489265	1	173236065	60 kb from *TNFSF4*	211 kb from *PRDX6*	–	0.25	8.19×10^{-12}	Add	Obs	1.43 (1.29–1.58)	A/C
rs4916342	1	173347837	172 kb from *TNFSF4*	99 kb from *PRDX6*	–	0.3	8.22×10^{-9}	Add	Imp	0.75 (0.67–0.82)	A/G
rs16833239	2	191940260	–	–	*STAT4*	0.15	9.69×10^{-10}	Add	Imp	0.67 (0.59–0.76)	G/A
rs11889341	2	191943742	–	–	*STAT4*	0.32	8.02×10^{-19}	Add	Obs	1.53 (1.40–1.69)	C/T
rs12612769	2	191953998	–	–	*STAT4*	0.31	2.37×10^{-19}	Add	Imp	1.59 (1.43–1.75)	A/C
rs13192841	6	137967214	152 kb from *OLIG3*	221 kb from *TNFAIP3*	–	0.13	7.50×10^{-3}	Dom	Obs	1.23 (1.06–1.43)	G/A
rs5029937	6	138195151	–	–	*TNFAIP3*	0.07	3.98×10^{-14}	Dom	Imp	2.11 (1.74–2.55)	G/T
rs5029939	6	138195723	–	–	*TNFAIP3*	0.07	4.09×10^{-14}	Dom	Imp	2.10 (1.74–2.55)	C/G
rs2230926	6	138196066	–	–	*TNFAIP3*	0.07	2.34×10^{-14}	Dom	Obs	1.93 (1.63–2.28)	T/G
rs9373203	6	138289848	86 kb from *TNFAIP3*	120 kb from *PERP*	–	0.37	4.89×10^{-6}	Add	Imp	1.25 (1.14–1.37)	C/T
rs6922466	6	138444930	16 kb from *PERP*	38 kb from *KIAA1244*	–	0.18	0.297	Dom	Obs	1.08 (0.94–1.24)	A/G
rs11185602	7	50299077	100 kb from *C7orf72*	45 kb from *IKZF1*	–	0.33	1.53×10^{-16}	Add	Imp	0.66 (0.60–0.73)	A/G
rs17552904	7	50318308	120 kb from *C7orf72*	26 kb from IKZF1	–	0.33	3.51×10^{-16}	Add	Obs	0.65 (0.59–0.72)	G/T
rs6964720	7	75180344	–	–	*HIP1*	0.24	2.00×10^{-10}	Add	Obs	1.40 (1.26–1.56)	A/G
rs139110493	7	75209951	–	–	*HIP1*	0.06	1.21×10^{-12}	Dom	Imp	2.48 (1.93–3.19)	G/C
rs4728142	7	128573967	11 kb from *LOC392787*	4 kb from *IRF5*	–	0.14	1.38×10^{-11}	Add	Obs	1.53 (1.35–1.73)	G/A
rs113478424	7	128575797	13 kb from *LOC392787*	2 kb from *IRF5*	–	0.14	3.97×10^{-12}	Add	Imp	1.59 (1.39–1.81)	15-mer/T†
rs922483	8	11351912	–	–	*BLK*	0.25	2.00×10^{-10}	Add	Obs	0.71 (0.64–0.79)	T/C
rs2736345	8	11352485	–	–	*BLK*	0.24	7.88×10^{-11}	Add	Imp	0.70 (0.63–0.78)	G/A
rs10857631	10	49955821	–	–	*WDFY4*	0.13	1.67×10^{-4}	Add	Imp	1.30 (1.13–1.48)	A/G
rs7097397	10	50025396	–	–	*WDFY4*	0.37	2.10×10^{-9}	Add	Obs	1.33 (1.21–1.46)	A/G
rs877819	10	50042951	–	–	*WDFY4*	0.16	0.0558	Add	Imp	1.13 (1.00–1.28)	G/A
rs10776651	10	50084526	–	–	*WDFY4*	0.34	1.54×10^{-7}	Add	Imp	1.29 (1.18–1.43)	C/T
rs1913517	10	50119054	–	–	*WDFY4*	0.31	2.54×10^{-5}	Add	Obs	1.24 (1.12–1.37)	G/A
rs12576753	11	128304141	None within 500 kb	25 kb from *ETS1*	–	0.39	1.74×10^{-11}	Add	Obs	1.37 (1.25–1.56)	C/A
rs1128334	11	128328959	–	–	*ETS1*	0.38	7.17×10^{-12}	Add	Imp	1.39 (1.26–1.52)	G/A
rs5986948	X	153266172	17.3 kb from *TMEM187*	9.8 kb from *IRAK1*	–	0.24	4.36×10^{-10}	Rec	Imp	0.64 (0.56–0.74)	C/T
rs1059702	X	153284192	–	–	*IRAK1*	0.26	5.14×10^{-10}	Rec	Imp	0.65 (0.57–0.74)	A/G
rs2734647	X	153292180	–	–	*MECP2*	0.25	7.54×10^{-9}	Dom	Obs	0.62 (0.51–0.75)	T/C

* Chr. = chromosome; MAF = minor allele frequency; $P_{GWAS} = P$ value in the genome-wide association study; Obs/Imp = observed/imputed; OR = odds ratio; 95% CI = 95% confidence interval; Add = additive; Dom = dominant; Rec = recessive.
† 15-mer = CTTAGCTATTGCTC.

Table 2. Single-locus analysis of regions genotyped in the replication study*

Marker (region name), major/minor allele	MAF in cases/ controls	P_{GWAS}	Model	OR (95% CI)	Replication meta-analysis			Overall meta-analysis		
					P	Q/I^2	OR (95% CI)	P	Q/I^2	OR (95% CI)
rs2267828 *(GTF2IRD1)*, A/G	0.40/0.45	7.02×10^{-7}	Add	0.79 (0.72–0.87)	0.02	0.56/0	0.87 (0.77–0.98)	6.46×10^{-8}	0.41/0	0.81 (0.76–0.88)
rs10901656 *(DOCK1)*, C/T	0.27/0.23	6.91×10^{-7}	Dom	1.39 (1.22–1.58)	0.095	0.6/0	1.14 (0.98–1.31)	9.56×10^{-6}	0.23/28.72	1.21 (1.12–1.32)
rs11235667 *(FCHSD2–P2RY2)*, A/G	0.07/0.11	1.03×10^{-8}	Add	0.59 (0.50–0.71)	0.0014	0.29/18.43	0.71 (0.57–0.87)	6.67×10^{-11}	0.14/44.37	0.63 (0.55–0.72)
rs1048257 *(AHNAK2)*, T/C	0.34/0.39	1.67×10^{-6}	Add	0.79 (0.72–0.87)	0.086	0.29/17.06	0.90 (0.80–1.01)	8.66×10^{-7}	0.12/47.82	0.82 (0.76–0.89)

* MAF = minor allele frequency; $P_{GWAS} = P$ value in the genome-wide association study; OR = odds ratio; 95% CI = 95% confidence interval; Add = additive; Dom = dominant.

reported by the ENCODE project (27). This variant has been shown to be located within an enhancer element in multiple immunologic cell types, based on data from the Epigenetics Roadmap project (see Supplementary Table 3, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract) (28). Chromatin immunoprecipitation followed by sequencing (ChIP-seq) conducted by the ENCODE project identified 2 proteins, POL2 and YY1, that were cross-linked to this region. Moreover, studies using sequence prediction methods in HaploReg version 2 have indicated that rs11235667 can alter the binding motif for the FOXa family of transcription factors (26). These data suggest that the likely functional mechanism involves regulation of expression of *ATG16L2*, *FCHSD2*, and/or *P2RY2*. However, the data in current eQTL databases do not suggest that rs11235667 influences the expression of these loci (26,29). This could be due to the lack of data from the correct cell and/or tissue type, and/or could be attributed to the fact that some databases do not interrogate this SNP in their studies.

Of the 8 other variants that exceeded genome-wide significance on the haplotype, several findings pointed to rs11235604 as an intriguing potential causal variant. This variant is a missense allele (R220W) that resides in the coding region of *ATG16L2*. Although it has been predicted to be benign using the PolyPhen-2 predictive tool (30), it is possible that this variant may still have an impact on the risk of SLE. In HaploReg version 2 (26), rs11235604 has been reported to alter 8 predicted binding motifs, and this variant is thought to be located within an active enhancer element in several immunologically relevant cell types (for further details, see Supplementary Table 3, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). Further work is needed to conclusively identify the polymorphism or polymorphisms responsible for this association signal. These studies might include evaluating the potential impact of rs11235667 on the expression of *ATG16L2*, *FCHSD2*, and/or *P2RY2*. In addition, experiments are needed to assess the impact on *ATG16L2* of the missense allele arising from rs11235604 and/or any other variants within this haplotype.

Associations with the HLA region in Koreans with SLE. One of the most consistent associations with SLE has been with the HLA region. Although the HLA was not the most statistically significant genotyped region, the SNP rs116727542, located in an observed broad peak of association spanning from *HLA–DR* (MIM ID 142860) to *HLA–DQ* (MIM ID 146880), showed the strongest SNP–SLE association ($P = 6.15 \times 10^{-24}$) after imputation (Figure 2A and Table 3; see also Supplementary Table 4, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). The interval between *HLA–DR* and *HLA–DQ* has previously been implicated in studies of Korean patients with SLE (31). In an attempt to identify the number of independent effects in this complex region, we used a stepwise approach (as described in Patients and Methods) and identified 10 independent effects (for complete results of the stepwise regression analysis of the HLA region after imputation, see Supplementary Figure 4A, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). The first 4 variants identified in the stepwise regression analysis, rs116727542, rs9273371, rs114653103, and rs115253455, are all located in the HLA class II region (Figure 2A; see also Supplementary Figure 4A and Supplementary Table 4, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract).

To better understand the relationship between the variants reported in this study and the classic HLA alleles, we imputed alleles at *HLA–A*, *B*, *C*, *DPB1*, *DQA1*, *DQB1*, and *DRB1*. Peak statistical significance ($P = 5.55 \times 10^{-16}$) was observed for the 2 tightly linked alleles *HLA–*

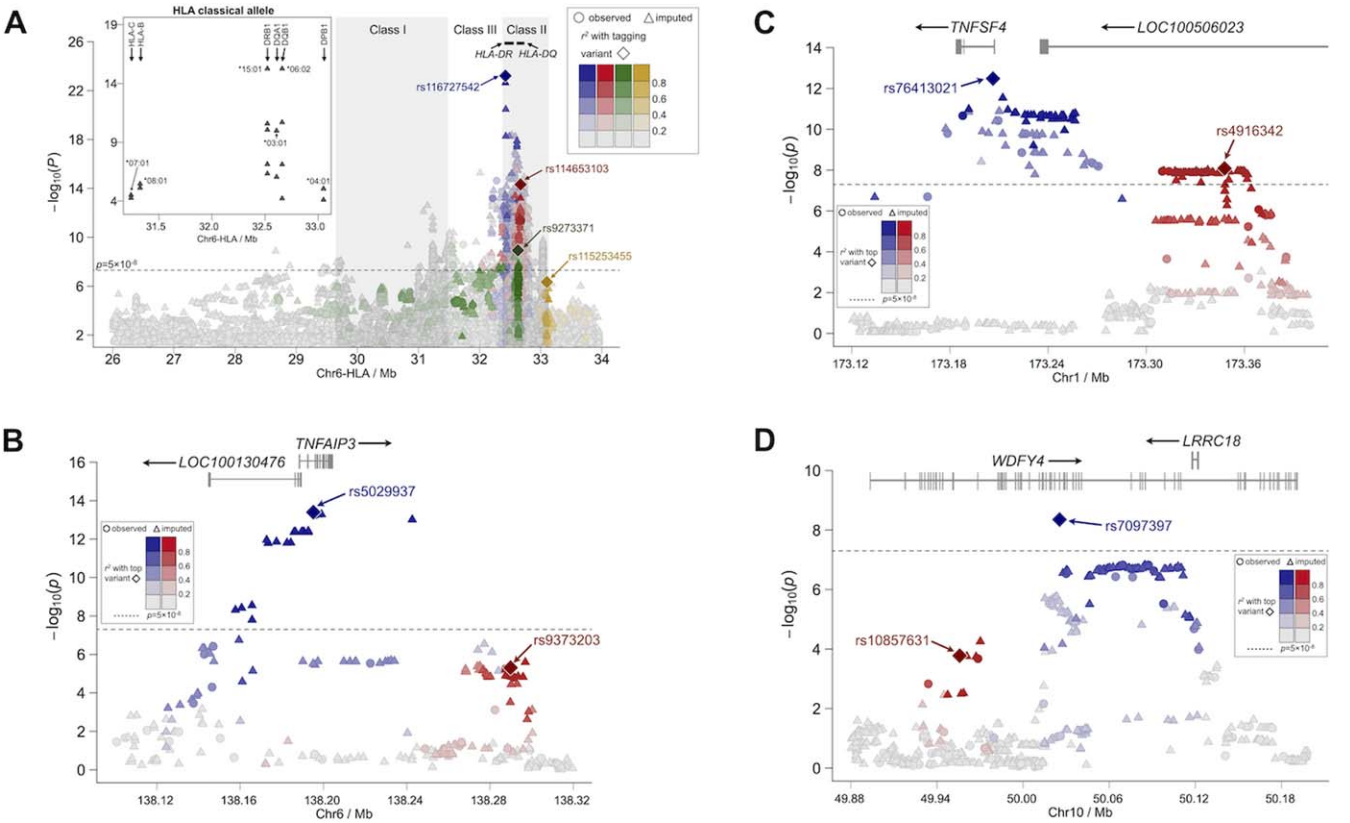

Figure 2. Expanded view of the association between systemic lupus erythematosus (SLE) and the HLA **(A)**, *TNFAIP3* **(B)**, *TNFSF4* **(C)**, and *WDFY4* **(D)** regions. **A,** The zoom plot shows the $-\log_{10}$ (*P* values) for the association with SLE of each observed variant (circles) and imputed variant (triangles) in the major histocompatibility complex region, according to basepair position from 26 Mb to 34 Mb on chromosome 6 (Chr6). Color key shows the values of linkage disequilibrium with the first 4 variants included in the stepwise logistic regression analysis (rs116727542, rs114653103, rs9273371, and rs115253455), all located within the HLA class II region. **Insert** on the left shows the $-\log_{10}$ (*P* values) for the association with SLE of the imputed classic alleles, according to basepair position from 31 Mb to 33.5 Mb. **B–D,** The zoom plot shows the $-\log_{10}$ (*P* values) for the association with SLE on chromosome 6 in the region of *TNFAIP3* **(B)**, chromosome 1 in *TNFSF4* **(C)**, and chromosome 10 in *WDFY4* **(D)**. For each independent effect, the variants showing peak association are represented by a diamond (in blue for the first effect, and in red for the second, if applicable). Color keys show the correlation of variants accounted for by each effect. The broken horizontal line in **A–D** indicates the genome-wide significance threshold of $P = 5 \times 10^{-8}$.

*DRB1*1501* (OR 1.85, 95% CI 1.59–2.14) and *HLA–DQB1*0602* (OR 1.90, 95% CI 1.62–2.21) (Figure 2A and Table 4; see also Supplementary Table 5, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). Stepwise logistic regression modeling of the HiBAG-imputed HLA alleles identified 13 independent effects (Table 4).

To better relate the classic alleles to the variants identified in this GWA scan, stepwise modeling with both the SNPs and classic HLA alleles was performed. After imputation analysis using the 1000 Genomes Project data, the peak effect observed was with the variant rs116727542, which was accounted for in *HLA–DQB1*0602* and *HLA–DRB1*0803* (for results of the additional 8 rounds of the stepwise regression analysis, see Supplementary Figure 4B, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract).

Confirmation of previously reported non-HLA associations with SLE and identification of novel independent effects. Several previously identified non-HLA loci associated with SLE were also replicated in this study, including *STAT1–STAT4*, *TNFSF4*, *TNFAIP3*, *IKZF1*, *HIP1*, *IRF5*, *BLK*, *WDFY4*, *ETS1*, and *IRAK1–MECP2* (Table 1 and Figures 2B–D; see also Supplementary Figures 5–11 and Supplementary Tables 6–15, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). Of these loci, the associations in the region of *TNFAIP3*, *TNFSF4*, and *WDFY4* had notable differences compared to those in previous studies.

After imputation of the *TNFAIP3* region, the primary independent effect in the stepwise model was observed at rs5029937, located within the second intron of *TNFAIP3* (Table 1 and Figure 2B). The second inde-

Table 3. Top 10 independent HLA associations identified in the single-locus and stepwise logistic regression analyses*

Marker†	Position	Upstream gene	Downstream gene	Within gene	MAF	*P*	OR (95% CI)	Major/minor allele	Stepwise *P*	Stepwise OR (95% CI)
rs116727542	32421227	8.4 kb from *HLA–DRA*	64 kb from *HLA–DRB5*	–	0.1700	6.15×10^{-24}	0.53 (0.47–0.60)	G/A	1.96×10^{-18}	1.74 (1.54–1.97)
rs9273371	32626565	14 kb from *HLA–DQA1*	675 bp from *HLA–DQB1*	–	0.1000	1.18×10^{-9}	1.61 (1.38–1.87)	C/T	8.43×10^{-5}	1.38 (1.18–1.63)
rs114653103	32668846	34 kb from *HLA–DQB1*	40 kb from *HLA–DQA2*	–	0.1200	7.31×10^{-15}	0.57 (0.49–0.66)	G/T	1.81×10^{-13}	0.50 (0.41–0.60)
rs115253455	33100021	43 kb from *HLA–DPB1*	30 kb from *COL11A2*	–	0.1300	4.68×10^{-7}	0.70 (0.61–0.80)	T/A	4.51×10^{-7}	0.66 (0.56–0.77)
Chr6: 31996524	31996524	–	–	*C4B*	0.1700	4.95×10^{-8}	0.71 (0.63–0.80)	C/A	4.87×10^{-10}	0.64 (0.55–0.73)
rs113833333	32594898	37 kb from *HLA–DRB1*	10 kb from *HLA–DQA1*	–	0.4100	2.30×10^{-5}	0.82 (0.74–0.90)	C/T	3.09×10^{-8}	0.74 (0.67–0.82)
rs116427960	31319226	79 kb from *HLA–C*	2.4 kb from *HLA–B*	–	0.0081	8.96×10^{-7}	4.57 (2.49–8.38)	C/T	1.69×10^{-5}	3.17 (1.87–5.35)
rs114904515	29362756	20 kb from *OR12D3*	1.7 kb from *OR12D2*	–	0.1000	1.55×10^{-6}	0.69 (0.60–0.80)	C/T	5.70×10^{-6}	0.67 (0.56–0.79)
rs118044183	30954150	–	–	*MUC21*	0.1200	3.61×10^{-6}	1.50 (1.27–1.79)	C/T	1.39×10^{-4}	1.39 (1.17–1.64)
rs2736191	31560910	~150 bp from *NCR3*	22 kb from *AIF1*	–	0.3800	7.53×10^{-7}	0.79 (0.71–0.86)	C/G	1.01×10^{-4}	0.81 (0.73–0.90)

* The results presented are the findings after adjustment for all other variants identified in the stepwise model (for complete results in the HLA region, see Supplementary Table 4), available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract. MAF = minor allele frequency; OR = odds ratio; 95% CI = 95% confidence interval; Chr6 = chromosome 6.

† All variants shown have been imputed.

Table 4. Multilocus model of the imputed HLA dosages in single-locus and stepwise analyses*

HLA allele	HLA dosage frequency		Best guess count		OR (95% CI)	Single locus *P*	Stepwise *P*	Stepwise OR (95% CI)
	Cases	Controls	Cases	Controls				
*DQB1*0602*†	0.25	0.15	301	636	1.90 (1.62–2.21)	5.55×10^{-16}	1.93×10^{-23}	2.35 (1.99–2.78)
*DRB1*0803*	0.20	0.13	251	613	1.59 (1.34–1.88)	7.37×10^{-8}	7.63×10^{-16}	2.14 (1.78–2.58)
*DQB1*0202*	0.18	0.12	212	502	1.60 (1.35–1.90)	7.57×10^{-8}	2.19×10^{-18}	2.50 (2.04–3.07)
*DQA1*0302*	0.19	0.16	286	847	1.41 (1.15–1.75)	1.27×10^{-3}	4.92×10^{-9}	1.98 (1.58–2.50)
*B*0801*	0.02	0.004	20	14	5.43 (2.66–11.08)	3.42×10^{-6}	5.15×10^{-6}	5.71 (2.70–12.07)
*DQA1*0401*	0.04	0.03	54	133	1.73 (1.17–2.57)	5.94×10^{-3}	3.11×10^{-5}	2.36 (1.58–3.54)
*C*0702*	0.22	0.17	261	711	1.37 (1.17–1.59)	5.45×10^{-5}	2.06×10^{-3}	1.30 (1.10–1.54)
*DRB1*0406*	0.03	0.07	41	354	0.15 (0.09–0.26)	2.63×10^{-11}	1.11×10^{-4}	0.32 (0.18–0.57)
*DPB1*0501*	0.79	0.72	930	3092	1.16 (1.05–1.28)	2.87×10^{-3}	1.06×10^{-4}	1.23 (1.11–1.36)
*DRB1*1602*	0.03	0.02	38	87	1.78 (1.16–2.72)	7.73×10^{-3}	1.53×10^{-3}	2.05 (1.32–3.20)
*DPB1*1701*	0.03	0.04	41	160	0.90 (0.62–1.30)	5.66×10^{-1}	1.07×10^{-3}	0.5 (0.33–0.76)
*C*0102*	0.29	0.32	359	1429	0.88 (0.77–1.01)	6.51×10^{-2}	5.82×10^{-3}	0.82 (0.71–0.94)
*DRB1*1202*	0.04	0.07	39	266	0.52 (0.37–0.74)	2.98×10^{-4}	7.07×10^{-3}	0.61 (0.42–0.87)

* Imputation analyses were done using the program HiBAG. The results presented are the findings after adjustment for all other variants identified in the stepwise model. OR = odds ratio; 95% CI = 95% confidence interval.
† The allele *HLA–DRB1*1501* had the same *P* value as that of *HLA–DQB1*0602*, but the latter was selected by the stepwise modeling procedure (for complete results, see Supplementary Table 5), available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract.

pendent effect was identified at rs9373203, which was located 3′ of the *TNFAIP3* coding region. A previous study of Han Chinese patients with SLE (6) demonstrated that rs2230926 was associated with disease, and another transracial mapping study of SLE in Koreans (a cohort that partially overlaps with subjects in the current study) and Europeans (32) identified an association of rs7749323 with risk of SLE. Both variants (rs2230926 and rs7749323) are highly correlated with rs5029937 ($D' = 1.0$, $r^2 > 0.98$), indicating that our results (see Supplementary Figure 12, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract) are consistent with the findings from these previous reports.

In addition, the second effect tagged by rs9373203 (Figure 2B) was not identified in studies by either Han et al (6) or Adrianto et al (32). Musone et al (33) identified effects of multiple variants, some of which spanned even further 3′ of *TNFAIP3* than rs9373203. After their stepwise analysis, they identified rs6922466 as the tagging variant accounting for this association; however, this variant was not found to be associated with SLE in Koreans in the present study. Moreover, the LD between these variants is very weak in Koreans ($D' = 0.43$, $r^2 = 0.00$; Supplementary Figure 12, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract), thus providing additional evidence that rs9373203 may be an independent effect warranting further study.

In the region of *TNFSF4*, after imputation, 2 independent effects were observed in the stepwise model. The first effect, peaking at rs76413021, was found to be located in the first intron of the *TNFSF4* coding region (Table 1 and Figure 2C). This variant is in LD with rs2205960 ($D' = 0.98$, $r^2 = 0.94$) and rs1234315 ($D' = 0.97$, $r^2 = 0.48$) (details in Supplementary Figure 13, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract) and was identified previously in a GWA scan of a Han Chinese population (6). Moreover, this effect was consistent with the findings reported in Europeans (34).

The second independent effect peaking at rs4916342 located in an intron of the gene *LOC100506023*, just 5′ of *TNFSF4*, was determined to be distinct from

signals found in previous studies (Figure 2C). Neither Han et al (6) nor Cunninghame Graham et al (34) reported association signals as far 5′ of *TNFSF4* as that observed in the present study for the second independent effect tagged by rs4916342. In a transracial mapping study of this region by Manku et al (35) that included subjects from East Asia, a second independent effect (tagged by rs1234314) was identified. In the present study of SLE in Koreans, we found that the first effect, tagged by rs76413021, accounted for the effect of rs1234314 in the stepwise model. Although rs1234314 and rs4916342 are located in the same general genomic location, the LD structure further supports the observation that they are not the same genetic effect; however, all of the risk variants are located on a single risk haplotype (see Supplementary Figure 13, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). This suggests that risk alleles for both rs76413021 and rs4916342 must be present to confer susceptibility to disease.

In the region of *WDFY4*, the present study did not replicate an effect of rs877819 (Table 1), which had been previously reported to result in a down-regulation of *WDFY4* through modification of a YY1 binding site (36). However, the results with regard to *WDFY4* in Koreans are consistent with those from 2 previous studies. First, the most statistically significant association within this region, the coding variant rs7097397 leading to the amino acid substitution R1816Q ($P = 2.10 \times 10^{-9}$), was previously reported by Yang et al (8). Second, we also demonstrated an association of rs1913517 ($P = 2.54 \times 10^{-5}$) (Table 1 and Figure 2D), which was identified previously by Han et al (6). Our haplotype and stepwise regression analysis indicated that there were 2 independent effects in the region, with rs7097397 accounting for the association observed at rs1913517, and rs10857631 tagging the second independent effect (Figure 2D and Supplementary Figures 14 and 15, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract).

Suggestive association with SLE identified in the GWA analysis. In total, 15 genotyped variants surpassed the suggestive significance threshold of $P = 2 \times 10^{-6}$ and were considered for further replication analysis (see Supplementary Table 16, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). Replication of the association was attempted for 3 additional variants, located within *GTF2IRD1*, *DOCK1,* and *AHNAK2* (Table 2), all of which have multiple genotyped variants that have been found to show suggestive significance in SLE and/or have been previously implicated in other related phenotypes. Only rs2267828 near *GTF2IRD1* yielded a $P_{\text{meta-rep}}$ value of <0.05, but this variant did not surpass genome-wide significance in the overall meta-analysis that included the data from the GWA scan (Table 2). The variant in the region of *AHNAK2,* rs1048257, was trending toward significance, while rs10901656 near *DOCK1* showed association in one replication cohort with the opposite allele (Table 2).

Outside of the 10 previously reported SLE loci described above, we observed 8 additional loci (on the list of ~50 that have been described previously) that showed a suggestive association at levels ranging from $P > 5 \times 10^{-8}$ to $P < 5 \times 10^{-5}$ (see Supplementary Table 17, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). This finding is likely due to the limited power of the present study and/or could be attributed to population-specific differences between this study and the studies in which these discoveries were originally identified.

DISCUSSION

The association in the 11q14 region observed in the present study peaked between 3 candidate genes, *ATG16L2*, *FCHSD2*, and *P2RY2*, all of which have the biologic potential to affect the pathophysiologic processes of SLE. Although this locus has not been identified as significant in other systemic autoimmune diseases, variants in this region have been found to be associated with Crohn's disease (MIM ID 266600) in Korean subjects (37). Moreover, the variant identified as showing peak significance in Crohn's disease, rs11235667, was also the variant revealed to be of peak significance in the present study of Korean patients with SLE. The missense variant, rs11235604, was also reported to be associated with Crohn's disease (37). GWA studies conducted in Europeans with SLE did not identify an association of the rs11235604 locus, since it is monomorphic in that population. Moreover, the GWA scans in the Han Chinese population (6–8) used the Illumina Human 610-Quad bead chip, which does not contain rs11235667.

ATG16L2 (autophagy-related 16–like 2) is a ubiquitously expressed homolog of *ATG16L1* (MIM ID 610767), a gene that has been implicated as a risk locus for Crohn's disease in patients of European descent (38,39). Both loci are involved in autophagy; however, little is known about the role that *ATG16L2* plays in the process. Interestingly, this pathway has been previously implicated in SLE. The gene *ATG5* (MIM ID 604261) has also been associated as a risk locus for lupus (6,40). Studies in mice have shown that *Apg16l* (the mouse equivalent of human *ATG16L*) interacts with *Apg5* (the mouse equivalent of human *ATG5*), suggesting that *ATG16L2* and *ATG5* may also interact in humans (41). More studies are needed to understand the functions of

ATG16L2 and whether it is involved in the association with SLE.

FCHSD2 (FCH and double SH3 domains 2) has been described as a regulator of F-actin assembly through interactions with WAS (also known as WASP) and WASL (also known as N-WASP) (42). *FCHSD2* is primarily expressed in CD19+ B cells, dendritic cells, myeloid cells, CD4+ T cells, and CD8+ T cells (39). Previous studies have shown that WAS plays an important role in the migration of T cells through reorganization of the actin cytoskeleton subsequent to interactions with dendritic or B cells (43).

P2RY2 (purinergic receptor P2Y, G protein–coupled, 2) is known to be involved in many cellular functions and is expressed in myeloid cells, including monocytes (39). *P2RY2* is a receptor for ATP and UTP that acts as a sensor for the release of nucleotides by apoptotic cells (44). Mice null for *P2RY2* showed a decreased ability to recruit monocytes and macrophages upon activation of nucleotides from apoptotic cells (44). *P2RY2* is also known to induce CCL2 secretion in macrophages, and coding variants in the receptor have been shown to influence secretion of this proinflammatory chemokine (45).

Although the HLA region has been associated with SLE susceptibility since the 1970s, the precise loci responsible for the risk conferred have not been fully characterized. A further cross-comparison of populations could be beneficial to take advantage of differences in LD and will likely help further refine the association signals seen in the GWA studies. For the classic alleles, previous studies have identified associations with alleles in the *HLA–DR* locus in Europeans, Chinese, Japanese, and Koreans (46–49), but *HLA–DQB1*0602* has not been implicated previously in Koreans. Two prominent classic HLA alleles identified in Europeans with SLE showed differences in association in Koreans. While *HLA–DRB1*1501* was among the alleles that were most significantly associated with SLE, *HLA–DRB1*0301* was found to be expressed at low frequency in this Korean population and was not associated with SLE (for details, see Supplementary Table 5, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). These results are consistent with those from a recent study in Koreans that evaluated the role of HLA in this population (50). Moreover, that study showed that amino acid changes to *HLA–DRB1* evident at positions 11, 13, and 26 accounted for the HLA association with SLE (50). However, it is possible that other amino acid–changing variants, noncoding RNAs, and/or transcriptional changes for other HLA loci that are coinherited with *HLA–DRB1* on these haplotypes may also be contributing to SLE risk.

The GWA scan in the present study replicated several loci that have been identified by prior studies as significantly associated with SLE, including *STAT1–STAT4*, *TNFSF4*, *TNFAIP3*, *IKZF1*, *HIP1*, *IRF5*, *BLK*, *WDFY4*, *ETS1*, and *IRAK1–MECP2*. It is important to note that a previous GWA scan of Korean women with SLE also demonstrated replication of the association of *STAT4* and *BLK* at a genome-wide significant level (31). Of these replicated regions, *TNFSF4*, *TNFAIP3*, *IKZF1*, *HIP1*, *IRF5*, *BLK*, and *ETS1* have functional effects that have been previously described (for details, see Supplementary Table 18, http://onlinelibrary.wiley.com/doi/10.1002/art.39548/abstract). Although most of the signals in these loci are identical, we did describe notable differences with independent effects in *TNFSF4* and *TNFAIP3*. Moreover, we did not observe association with rs877819, which had been proposed as a putative causal variant leading to expression differences of *WDFY4* (36).

In conclusion, in this study we performed a GWA scan of Korean SLE patients and population controls, resulting in identification of 12 regions that surpassed the genome-wide significance level. The region from *ATG16L2* through *FCHSD2* to the promoter region of the *P2RY2* locus was identified and confirmed as an SLE-associated region. Herein, we also observed strong associations in the HLA region and showed the relationship between the classic HLA alleles and the variants reported within this GWA study in Koreans. The 10 additional regions identified, *STAT1–STAT4*, *TNFSF4*, *TNFAIP3*, *IKZF1*, *HIP1*, *IRF5*, *BLK*, *WDFY4*, *ETS1*, and *IRAK1–MECP2*, have previously been implicated in SLE. Additional replication analyses of the loci suggestive of a significant association in this study are needed to determine their relationship with SLE. Although GWA approaches have been very successful in the identification of risk loci, continued efforts are needed to narrow down the association signals to the causal variants, and to determine the functional causal mechanisms contributing to the pathogenesis of SLE.

ACKNOWLEDGMENTS

We are grateful to all of the patients with SLE and individuals serving as healthy controls for their participation in this study.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Drs. Bae and Tsao had full access to

all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Lessard, Zhao, K. Kim, Bang, H. S. Lee, Shim, Alarcón-Riquelme, Criswell, Harley, Sivils, Bae, Langefeld, Tsao.

Acquisition of data. Lessard, Zhao, K. Kim, Ice, Rasmussen, Kelly, Bang, Joo, Choi, H. S. Lee, Kang, Suh, Chung, S. K. Lee, Choe, Shim, Han, Shen, Howe, Song, Gaffney, Jacob, Kimberly, Vyse, Harley, Sivils, Bae, Tsao.

Analysis and interpretation of data. Lessard, Sajuthi, Zhao, K. Kim, Ice, Li, Ainsworth, Marion, Bang, H. S. Lee, Oh, Y. J. Kim, Wakeland, Li, Alarcón-Riquelme, Sivils, Bae, Langefeld, Tsao.

REFERENCES

1. Tsokos GC. Systemic lupus erythematosus. N Engl J Med 2011; 365:2110–21.
2. Arbuckle MR, McClain MT, Rubertone MV, Scofield RH, Dennis GJ, James JA, et al. Development of autoantibodies before the clinical onset of systemic lupus erythematosus. N Engl J Med 2003;349:1526–33.
3. Moser KL, Kelly JA, Lessard CJ, Harley JB. Recent insights into the genetic basis of systemic lupus erythematosus. Genes Immun 2009;10:373–9.
4. Deng Y, Tsao BP. Advances in lupus genetics and epigenetics. Curr Opin Rheumatol 2014;26:482–92.
5. Dai C, Deng Y, Quinlan A, Gaskin F, Tsao BP, Fu SM. Genetics of systemic lupus erythematosus: immune responses and end organ resistance to damage. Curr Opin Immunol 2014;31:87–96.
6. 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.
7. Yang W, Tang H, Zhang Y, Tang X, Zhang J, Sun L, et al. Meta-analysis followed by replication identifies loci in or near CDKN1B, TET3, CD80, DRAM1, and ARID5B as associated with systemic lupus erythematosus in Asians. Am J Hum Genet 2013;92:41–51.
8. Yang W, Shen N, Ye DQ, Liu Q, Zhang Y, Qian XX, et al, Asian Lupus Genetics Consortium (ALGC). Genome-wide association study in Asian populations identifies variants in ETS1 and WDFY4 associated with systemic lupus erythematosus. PLoS Genet 2010;6:e1000841.
9. Okada Y, Shimane K, Kochi Y, Tahira T, Suzuki A, Higasa K, et al. A genome-wide association study identified AFF1 as a susceptibility locus for systemic lupus eyrthematosus in Japanese. PLoS Genet 2012;8:e1002455.
10. Kim K, Bang SY, Lee HS, Cho SK, Choi CB, Sung YK, et al. High-density genotyping of immune loci in Koreans and Europeans identifies eight new rheumatoid arthritis risk loci. Ann Rheum Dis 2015;74:e13.
11. Lessard CJ, Adrianto I, Ice JA, Wiley GB, Kelly JA, Glenn SB, et al. Identification of IRF8, TMEM39A, and IKZF3-ZPBP2 as susceptibility loci for systemic lupus erythematosus in a large-scale multiracial replication study. Am J Hum Genet 2012;90: 648–60.
12. Kaiser R, Taylor KE, Deng Y, Zhao J, Li Y, Nititham J, et al. Single-nucleotide polymorphisms in VKORC1 are risk factors for systemic lupus erythematosus in Asians. Arthritis Rheum 2013;65:211–5.
13. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25:1271–7.
14. Hochberg MC, for the Diagnostic and Therapeutic Criteria Committee of the American College of Rheumatology. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. Arthritis Rheum 1997;40:1725.
15. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics 2010;26:2867–73.
16. 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.
17. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38: 904–9.
18. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010; 26:2190–1.
19. Cochran WG. The combination of estimates from different experiments. Biometrics 1954;10:101–29.
20. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–60.
21. Durbin RM, Abecasis GR, Altshuler DL, Auton A, Brooks LD, Gibbs RA, et al. A map of human genome variation from population-scale sequencing. Nature 2010;467:1061–73.
22. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. Nat Methods 2012;9:179–81.
23. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009;5:e1000529.
24. Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet 2010;11:499–511.
25. Zheng X, Shen J, Cox C, Wakefield JC, Ehm MG, Nelson MR, et al. HIBAG—HLA genotype imputation with attribute bagging. Pharmacogenomics J 2014;14:192–200.
26. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res 2012;40: D930–4.
27. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57–74.
28. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 2009;462: 315–22.
29. Raj T, Rothamel K, Mostafavi S, Ye C, Lee MN, Replogle JM, et al. Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. Science 2014;344:519–23.
30. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248–9.
31. Lee HS, Kim T, Bang SY, Na YJ, Kim I, Kim K, et al. Ethnic specificity of lupus-associated loci identified in a genome-wide association study in Korean women. Ann Rheum Dis 2014;73: 1240–5.
32. Adrianto I, Wen F, Templeton A, Wiley G, King JB, Lessard CJ, et al. Association of a functional variant downstream of TNFAIP3 with systemic lupus erythematosus. Nat Genet 2011; 43:253–8.
33. Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W, et al. Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. Nat Genet 2008;40:1062–4.
34. Cunninghame Graham DS, Graham RR, Manku H, Wong AK, Whittaker JC, Gaffney PM, et al. Polymorphism at the TNF superfamily gene TNFSF4 confers susceptibility to systemic lupus erythematosus. Nat Genet 2008;40:83–9.
35. Manku H, Langefeld CD, Guerra SG, Malik TH, Alarcon-Riquelme M, Anaya JM, et al. Trans-ancestral studies fine map the SLE-susceptibility locus TNFSF4. PLoS Genet 2013;9: e1003554.
36. Zhao H, Yang W, Qiu R, Li J, Xin Q, Wang X, et al. An intronic variant associated with systemic lupus erythematosus

changes the binding affinity of Yinyang1 to downregulate WDFY4. Genes Immun 2012;13:536–42.
37. Yang SK, Hong M, Zhao W, Jung Y, Baek J, Tayebi N, et al. Genome-wide association study of Crohn's disease in Koreans revealed three new susceptibility loci and common attributes of genetic susceptibility across ethnic populations. Gut 2014;63: 80–7.
38. Hampe J, Franke A, Rosenstiel P, Till A, Teuber M, Huse K, et al. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. Nat Genet 2007;39:207–11.
39. Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S, et al. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biol 2009;10:R130.
40. 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, PXK, KIAA1542 and other loci. Nat Genet 2008;40: 204–10.
41. Mizushima N, Kuma A, Kobayashi Y, Yamamoto A, Matsubae M, Takao T, et al. Mouse Apg16L, a novel WD-repeat protein, targets to the autophagic isolation membrane with the Apg12-Apg5 conjugate. J Cell Sci 2003;116:1679–88.
42. Cao H, Yin X, Cao Y, Jin Y, Wang S, Kong Y, et al. FCHSD1 and FCHSD2 are expressed in hair cell stereocilia and cuticular plate and regulate actin polymerization in vitro. PLoS One 2013; 8:e56516.
43. Lafouresse F, Cotta-de-Almeida V, Malet-Engra G, Galy A, Valitutti S, Dupre L. Wiskott-Aldrich syndrome protein controls antigen-presenting cell-driven CD4+ T-cell motility by regulating adhesion to intercellular adhesion molecule-1. Immunology 2012;137:183–96.
44. Elliott MR, Chekeni FB, Trampont PC, Lazarowski ER, Kadl A, Walk SF, et al. Nucleotides released by apoptotic cells act as a find-me signal to promote phagocytic clearance. Nature 2009; 461:282–6.
45. Higgins KR, Kovacevic W, Stokes L. Nucleotides regulate secretion of the inflammatory chemokine CCL2 from human macrophages and monocytes. Mediators Inflamm 2014;2014:293925.
46. Lee HS, Chung YH, Kim TG, Kim TH, Jun JB, Jung S, et al. Independent association of HLA-DR and FCγ receptor polymorphisms in Korean patients with systemic lupus erythematosus. Rheumatology (Oxford) 2003;42:1501–7.
47. 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.
48. Furukawa H, Oka S, Shimada K, Sugii S, Hashimoto A, Komiya A, et al. Association of increased frequencies of HLA-DPB1* 05:01 with the presence of anti-Ro/SS-A and anti-La/SS-B antibodies in Japanese rheumatoid arthritis and systemic lupus erythematosus patients. PLoS One 2013;8:e53910.
49. Furukawa H, Kawasaki A, Oka S, Ito I, Shimada K, Sugii S, et al. Human leukocyte antigens and systemic lupus erythematosus: a protective role for the HLA-DR6 alleles DRB1*13:02 and *14:03. PLoS One 2014;9:e87792.
50. Kim K, Bang SY, Lee HS, Okada Y, Han B, Saw WY, et al. The HLA-DRβ1 amino acid positions 11-13-26 explain the majority of SLE-MHC associations. Nat Commun 2014;5:5902.