Published in final edited form as:
Arthritis Rheum. 2011 July ; 63(7): 2049–2057. doi:10.1002/art.30356.

Genetic Analyses of Interferon Pathway-Related Genes Reveals Multiple New Loci Associated with Systemic Lupus Erythematosus (SLE)

Paula S. Ramos, PhD[*], **Adrienne H. Williams, MS**, **Julie T. Ziegler, MS**, **Mary E. Comeau, MS**, **Richard T. Guy, MS**, **Christopher J. Lessard, PhD**, **He Li, BSc**, **Jeffrey C. Edberg, PhD**, **Raphael Zidovetzki, PhD**, **Lindsey A. Criswell, MD, MPH**, **Patrick M. Gaffney, MD**, **Deborah Cunninghame Graham, PhD**, **Robert R. Graham, PhD**, **Jennifer A. Kelly, MS**, **Kenneth M. Kaufman, PhD**, **Elizabeth E. Brown, PhD, MPH**, **Graciela S. Alarcón, MD, MPH**, **Michelle A. Petri, MD, MPH**, **John D. Reveille, MD**, **Gerald McGwin, PhD**, **Luis M. Vilá, MD**, **Rosalind Ramsey-Goldman, MD, DrPH**, **Chaim O. Jacob, MD, PhD**, **Timothy J. Vyse, MRCP, PhD**, **Betty P. Tsao, PhD**, **John B. Harley, MD, PhD**, **Robert P. Kimberly, MD**, **Marta E. Alarcón-Riquelme, MD, PhD**, **Carl D. Langefeld, PhD**, and **Kathy L. Moser, PhD**

Abstract

Objective—The overexpression of interferon (IFN)-inducible genes is a prominent feature of SLE, serves as a marker for active and more severe disease, and is also observed in other autoimmune and inflammatory conditions. The genetic variations responsible for sustained activation of IFN responsive genes are unknown.

Methods—We systematically evaluated association of SLE with a total of 1,754 IFN-pathway related genes, including IFN-inducible genes known to be differentially expressed in SLE patients and their direct regulators. We performed a three-stage design where two cohorts (total n=939 SLE cases, 3,398 controls) were analyzed independently and jointly for association with SLE, and the results were adjusted for the number of comparisons.

Results—A total of 16,137 SNPs passed all quality control filters of which 316 demonstrated replicated association with SLE in both cohorts. Nine variants were further genotyped for confirmation in an average of 1,316 independent SLE cases and 3,215 independent controls. Association with SLE was confirmed for several genes, including the transmembrane receptor CD44 (rs507230, $P = 3.98\times10^{-12}$), cytokine pleiotrophin (PTN) (rs919581, $P = 5.38\times10^{-04}$), the heat-shock DNAJA1 (rs10971259, $P = 6.31\times10^{-03}$), and the nuclear import protein karyopherin alpha 1 (KPNA1) (rs6810306, $P = 4.91\times10^{-02}$).

Conclusion—This study expands the number of candidate genes associated with SLE and highlights the potential of pathway-based approaches for gene discovery. Identification of the causal alleles will help elucidate the molecular mechanisms responsible for activation of the IFN system in SLE.

Introduction

Systemic lupus erythematosus (SLE [MIM152700]) is a chronic and severe systemic autoimmune disease characterized by the production of high titers of autoantibodies directed against native DNA and a wide variety of other cellular constituents. The prevalence of SLE

Corresponding author: Paula S. Ramos, Wake Forest University Health Sciences, PHS, Department of Biostatistics, Medical Center Blvd., WC 23, Winston Salem, NC 27157, USA; Phone: 336-713-1127; Fax: 336-716-6427; pramos@wfubmc.edu.

Disclosures: R.R.G. is an employee of Genentech.

in the U.S. is estimated between 0.05% and 0.1% of the population, disproportionately affecting women and African Americans (1). SLE susceptibility is strongly influenced by genetic factors (2–7). To date, association of SLE with about 38 loci have been convincingly established. Clustering of some genetic associations identified to date appears to fall into at least three major pathways including immune complex processing, lymphocyte signaling, and interferon (IFN) pathways (8).

Numerous studies have clearly demonstrated that dysregulation of the IFN system occurs in SLE and closely related autoimmune phenotypes including Sjögren's syndrome, psoriasis, and others (9). Genome-wide transcriptional profiling in SLE has shown that many patients overexpress IFN-inducible genes (10–19). This observed overexpression of IFN-inducible genes, known as the "IFN signature", is a marker for patients with active and severe disease. Dysregulation of IFN responses also correlates with several clinical and laboratory criteria, and is present in virtually all pediatric cases (9). Furthermore, some individuals treated with IFN-α later develop anti-nuclear antibodies or even SLE (20). High serum IFN-α activity, consistent with overexpression of IFN-inducible genes, is a heritable trait in families with SLE (21). Sustained overproduction of IFNs activates dendritic cells, autoreactive T cells, autoreactive B cells and cytotoxic effector cells. Thus, many of the immunological disturbances observed in SLE, such as peripheral tolerance breakdown, nuclear autoantibody production, immune complex formation and systemic tissue damage, may be explained at least in part by an impaired IFN system (22).

The role of IFNs in the homeostasis of the immune system and their observed dysregulation in patients with SLE makes any gene in this system a potential candidate for SLE susceptibility. To date, association analyses have established the interferon regulatory factor 5 gene (IRF5) and a few others related to IFN pathways (e.g. STAT4, SPP1, and TREX1) as risk factors for SLE (8;23). Given that the likelihood that additional IFN-related genes important in SLE have yet to be identified, it is crucial to investigate the genetic contributions of such genes to SLE.

In this study, we utilized a pathway centric approach to perform the first comprehensive genetic association analysis of genes known to constitute the IFN signature, their direct regulators and all other known IFN-pathway genes based on literature and database searches. Independent discovery (Stage 1) and replication (Stage 2) datasets consisted of both the observed and imputed IFN-related single nucleotide polymorphisms (SNPs) from the genome-wide association studies (GWAS) by Harley et al. and Graham et al., respectively. We performed single locus tests of association, admixture adjustments, and adjusted our results for the number of comparisons. We then confirmed the top findings in a third confirmation cohort (Stage 3). In addition, we also performed two-way interaction tests of association and applied a more novel approach, alternating decision trees (ADTrees), to test the predictive ability of these polymorphisms and their potential higher order architecture. We report novel SLE risk loci with confirmed evidence for association in all the cohorts.

Patients and Methods

Criteria for definition of interferon (IFN) pathway-related genes

We compiled two lists with different sets of IFN-related genes:

- Set 1: all genes reported as differentially expressed IFN-inducible genes in SLE plus all known IFN genes. This list was compiled from gene expression profiling studies that observed an IFN signature in SLE patients (10–19), or by searching

NCBI and Ingenuity Pathway Analysis (IPA) (www.ingenuity.com) for genes (and pseudogenes) with "IFN" in the gene or protein name or alias.

- Set 2: full set of direct regulators of the differentially expressed IFN-inducible genes compiled from the literature and included in Set 1. We used IPA to identify all the regulators (cytokines, transporters, kinases, peptidases, phosphatases, growth factors, ion channels, nuclear receptors, transmembrane receptors, G-protein coupled receptors, transcription and translation regulators) that are upstream and have direct interactions with the above literature genes.

Association analysis in the Discovery Cohort (Stage 1)—As described in Harley et al. (3), a genome-wide association study (GWAS) using 317,501 SNPs was performed by the International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEN; www.slegen.org) in 706 Caucasian females with SLE and 2,317 controls genotyped on the Illumina Infinium HumanHap300 BeadChip. Genotypes from these subjects were imputed using the program IMPUTE (24) version 0.5 for SNPs not genotyped or poorly genotyped. Imputation was performed using high quality genotype data from the SLEGEN GWAS (3) and phased HapMap Phase II (NCBI B35 assembly) genotype data from 60 CEU HapMap founders. We considered SNPs that mapped within 50 kb upstream and 10 kb downstream of each IFN-related gene. We used SNPs that met the following quality criteria: 1) No statistically significant differences in the proportions of missing genotype data between cases and controls (i.e., $P > 0.05$); 2) overall < 10% missing genotype data; 3) Hardy-Weinberg Expectations (HWE) in controls $P > 0.01$, HWE in cases $P > 0.0001$; and 4) minor allele frequencies (MAFs) of controls within a 95% or 99.99% confidence interval for ethnicity matched HapMap MAFs, for genotyped and imputed SNPs, respectively. Retained SNPs had an estimated MAF > 0.01 in the control samples, an information score > 0.50 and a confidence score > 0.90. Imputed SNPs were analyzed using SNPTEST (24). We report the lowest *P-value* among the additive, dominant and recessive models; however, since these tests can be affected by low genotype counts, we require at least 30 homozygotes for the minor allele to consider the recessive or additive models. All genetic models were defined relative to the minor allele. To account for potential population stratification, we computed Principal Component analysis (PCs) using all SNPs, as described (3). After adjustment for four PCs, the genome-wide inflation factor was λ=1.05, indicating minimal inflation of the test statistics.

Association analysis in the Replication Cohort (Stage 2)—Replication studies were carried out in an independent set of subjects from the University of Minnesota (MN) SLE Cohort using GWAS data as described by Graham et al. (2). Genotypes from 412 cases and 1,081 controls were imputed using the program IMPUTE (24) version 0.5 for SNPs not genotyped or poorly genotyped in the MN GWAS. Imputation was performed using high quality MN GWAS genotype data (Affymetrix Genome-Wide Human SNP Array 5.0 platform) and phased HapMap Phase II (NCBI B35 assembly) genotype data from 60 CEU HapMap founders. Retained SNPs had an estimated MAF > 0.01 in the control samples, an information score > 0.50 and a confidence score > 0.90. Imputed SNPs were analyzed using SNPTEST (24). We used SNPs that met the same quality criteria and present the *P-value* chosen as described above. After adjustment for four PCs, the genome-wide inflation factor in this analysis was λ=1.05. We only report those replications that show consistency of the risk allele in both the discovery and replication studies.

Joint-analysis of the Discovery and Replication Cohorts (Stages 1 and 2)—We combined the genotypic and imputed data from both the Discovery (Stage 1)(3) and Replication (Stage 2)(2) GWAS datasets and performed a joint-analysis. A total of 179 SLE cases were duplicates or first degree relatives between both studies and were removed from

the Replication set, bringing the number of MN cases in the joint analysis to 233. We used SNPs that met the same quality criteria as described above, and report the *P-value* chosen as described above. These analyses were adjusted for four PCs to account for admixture. The genome-wide inflation factor in the joint analysis was λ=1.08. The reported *P-values* for the tests of association are not adjusted for the number of comparisons. However, we corrected for multiple comparisons by applying a False Discovery Rate (FDR) (25) to all SNPs in the joint-analysis that passed QC within Set 1 (2,169) and Set 2 (12,997), and only declare statistical significance for those that met a FDR-adjusted threshold of significance ($P < 0.05$) within each Set.

Meta-analysis with the Confirmation Cohort (Stage 3)—We genotyped most of the SNPs that replicated in an additional, independent Caucasian Confirmation Cohort. This Confirmation Cohort consisted of 474 cases and 539 controls from the Lupus Family Registry and Repository (LFRR)(26), 739 cases from the PROFILE study (27), 902 cases and 214 from the UK, and 2,729 out-of-study controls. The UK sample was collected by Timothy Vyse at King's College, London and includes 200 controls taken from the British 1958 Birth Control Cohort. These probands conformed to the American College of Rheumatology criteria for SLE (28). Written consent was obtained from all participants. In the UK, ethical approval was obtained from Multi-Centre Research Ethics Committee. The out-of-study controls were available from the following studies at the dbGaP database (http://www.ncbi.nlm.nih.gov/gap): the Finland-United States Investigation of NIDDM Genetics (FUSION) Study, the NIDDK IBDGC Crohn's Disease GWAS, the CGEMS Prostate Cancer GWAS - Stage 1 – PLCO, the CGEMS Breast Cancer GWAS - Stage 1 – NHS, and the NINDS Parkinson's Disease study. We excluded duplicates and relatives between these studies, and used all GWAS SNPs to compute PCs and exclude all outliers. The following SNPs were genotyped in the LFRR samples: rs10124051, rs1880791, rs2285210, rs2613310, rs507230, rs6810306, rs749701, rs755690. All these SNPs plus the following were genotyped in the PROFILE collection: rs10971259, rs366078, rs4659444. Only the following three SNPs were genotyped in the UK samples: rs10124051, rs10971259 and rs919581. In the LFRR and PROFILE samples SNPs were genotyped using TaqMan Pre-Designed SNP Genotyping Assays following the manufacturer's protocol (Applied Biosystems, Foster City, CA, USA). The UK samples were genotyped on a custom Illumina array. All SNPs had overall < 5% missing genotype data and HWE $P > 0.05$ in their genotyping cohorts. The out-of-study controls had < 14% missing genotypes and HWE $P > 0.01$. We performed a weighted Z-score meta-analysis as implemented in METAL (www.sph.umich.edu/csg/abecasis/metal) between Stages 1, 2 and 3 SNPs. METAL combines study-specific *P-values* and direction of effect independently of β-estimates, and then converts them into signed Z-statistics that are summed with weights proportional to the square root of the sample size for each dataset.

Two-locus interaction analysis—Using the Discovery Cohort (Stage 1), we computed all two-locus interaction analyses among all SNPs that had an individual locus *P-value* < 0.2. Specifically, we computed a logistic regression model with each SNP modeled under an additive genetic model and the interaction as the centered crossproduct of the SNPs under the additive model. This analysis included 6,324 SNPs with $P < 0.2$ identified in the Discovery Cohort analysis. These SNPs met the same quality criteria defined above. In order to reduce false positive interactions due to low MAFs, we rejected all the pairs for which the expected number of individuals in our dataset was less than 10 for minor allele homozygotes. In addition, we excluded all SNP pairs with a linkage disequilibrium (LD) measure of $r^2 > 0.2$. The interactions were adjusted for four principal components to account for admixture. We then performed these analyses in the Replication Cohort. Although no

pairs remained significant after correcting for the number of comparisons, several interactions passed all quality control criteria and were replicated.

Alternating Decision Tree (ADTrees)—We have also used an ADTree approach to identify the variants that best distinguish case vs. control status and the multilocus relationships among the variants. This analysis included the 6,324 SNPs used for the two-locus interaction analysis described above. In order to minimize biased results, we only used autosomal non-HLA region SNPs that met the quality criteria defined previously. We used the Weka 3 software (www.cs.waikato.ac.nz/ml/weka/index.html) to create one hundred bootstrap samples (random sampling with replacement) from the Discovery Cohort dataset and an ADtree (29;30) was constructed for each sample. Common structural elements across the 100 trees were recorded and the SNPs involved in each common element that appeared in at least 5% of the trees were removed from the data set. Structural elements are defined as paths through a tree that include the root and a leaf node. In the event that no structural element appeared in at least 5% of the trees, the criteria was reduced to paths through the tree that include only the root node and only paths appearing in at least 10% of the trees. Analysis was repeated until no common structural elements met the threshold of inclusion. The resulting paths were ranked by order in which they were found, then number of trees in which they appeared.

SNP functionality evaluation—Linkage disequilibrium between SNPs was assessed using Haploview 4.1 (www.broadinstitute.org/mpg/haploview) with HapMap CEU Data (Release 24/Phase II) (hapmap.ncbi.nlm.nih.gov). SNP function was evaluated with the UCSC genome browser (genome.ucsc.edu) March 2006 assembly. SNP effects on expression levels were examined using the SCAN database (scan.bsd.uchicago.edu).

Results

Association Analysis

We compiled a list of 323 IFN-inducible genes identified through gene expression profiling experiments in SLE and 136 additional IFN pathway genes for Set 1. An additional 1,761 genes identified as direct regulators of the IFN-inducible genes in Set 1 were included in Set 2 (Table 1). Thus, a total of 2,220 genes were selected for this targeted IFN pathway study.

We next searched for SNPs that mapped within 50 kb upstream and 10 kb downstream of each IFN-related gene in Set 1 that were present in the Discovery Cohort GWAS dataset (3) (Stage 1). Set 1 had 4,759 SNPs and Set 2 had 26,265 SNPs, both genotyped and imputed, that were then tested for association with SLE. We repeated this process using data from our Replication Cohort dataset (2) (Stage 2) and selected 4,680 SNPs in Set 1 and 25,196 SNPs in Set 2, also genotyped and imputed, to test for association with SLE. A total of 2,169 SNPs in Set 1 and 12,997 SNPs in Set 2 passed all quality control filters in both datasets. Of these, 56 SNPs within 26 loci in Set 1, and 249 SNPs on 119 loci in Set 2, resulted in $P < 0.05$ in both the Discovery and Replication Cohorts (Stages 1 and 2). It is noteworthy that we observed an enrichment of results above the null distribution: from the 2,169 SNPs tested in Set 1 and 12,997 tested in Set 2, we observed 325 SNPs in Set 1 and 1608 SNPs in Set 2 with $P<0.05$, well above the 109 SNPs expected in Set 1 and 650 SNPs expected in Set 2 by chance. We herein report the unadjusted *P-values* in the joint-analysis (P_{joint}) and, unless explicitly noted, only report variants that survived a False Discovery Rate (FDR) (25) correction for the number of comparisons in the joint-analysis (P_{FDR}) of Stages 1 and 2. Finally, we attempted to confirm 11 of the most significant novel SNPs observed in this joint-analysis in an independent Confirmation Cohort (Stage 3). A meta-analysis between

the joint-analysis of Stages 1 and 2, and the Confirmation Cohort (Stage 3) was then computed.

Examination of the 561 loci with evidence for association in both Stages 1 and 2 revealed both known and novel SLE risk genes. Not surprisingly, the most significant SNPs reside in genes in the *HLA* region (complement factor B (CFB), rs1270942, $P_{joint} = 5.29\times10^{-25}$, $P_{FDR} = 1.11\times10^{-21}$). Other genes with well-established SLE associations included *IRF5-TNPO3* (rs10488631, $P_{joint} = 1.45\times10^{-19}$, $P_{FDR} = 1.71\times10^{-16}$), *ITGAM-ITGAX* (rs11150610, $P_{joint} = 1.17\times10^{-7}$, $P_{FDR} = 3.18\times10^{-05}$), *STAT4* (rs3024896, $P_{joint} = 1.10\times10^{-04}$, $P_{FDR} = 1.91\times10^{-02}$) and *TNFAIP3* (rs3757173, $P_{joint} = 4.37\times10^{-7}$, $P_{FDR} = 7.89\times10^{-05}$) *loci* (2–4;31;32).

As shown in Table 2, the most significant association of SLE with a novel locus in Set 1 is that of the heat-shock protein *DnaJ (Hsp40) homolog, subfamily A, member 1* (DNAJA1) (rs10124051, $P_{joint} = 1.12\times10^{-04}$, $P_{FDR} = 1.43\times10^{-02}$). Another SNP in this locus in LD with the former (r^2=0.8) remained associated in the meta-analysis between Stage 1, 2 and 3 cohorts (rs10971259, $P = 6.31\times10^{-03}$). Differential expression in peripheral blood of SLE patients versus controls for this gene was reported by Baechler et al (10).

Two other *loci* in this set, those of *RPS6KA1* (rs4659444, $P_{joint} = 1.14\times10^{-03}$, $P_{FDR} = 8.85\times10^{-02}$) and *IRF8* (rs366078, $P_{joint} = 1.53\times10^{-05}$, $P_{FDR} = 3.11\times10^{-03}$), showed associations that were further confirmed in a joint-analysis between Stages 1, 2, and additional cases from Stage 3 ($P = 2.14\times10^{-03}$ and $P= 9.63\times10^{-03}$, respectively).

The loci in Set 2 (Table 3) show stronger associations with SLE than those in Set 1. The most significant variant in Set 2 lies in the *spleen tyrosine kinase* (SYK) gene (rs2613310, $P_{joint} = 1.53\times10^{-05}$, $P_{FDR} = 3.11\times10^{-03}$), but it did not remain associated in the meta-analysis between all cohorts (Table 4).

The next strongest effect resides in the *pleiotrophin* (PTN) gene, where one out of three replicated variants passed a FDR adjustment (rs919581, $P_{joint}= 5.86\times10^{-05}$, $P_{FDR} = 1.12\times10^{-02}$). This SNP was genotyped on the UK samples of the Confirmation (Stage 3) cohort, and continues to show evidence for association in the meta-analysis between these samples and the original cohorts ($P = 5.38\times10^{-04}$). Nevertheless, this variant is not in LD (r^2<0.2) with any known functional SNP in the gene. PTN has direct interactions with several reported IFN-inducible molecules: PSMB10 (15), APOBEC3G (10), CSTA (10), AGRN (9;11), PLXNB2 (10), and VEGFA (10).

CD44 is important in lymphocyte activation and homing. One replicated variant 31.2 kb upstream the *CD44* antigen gene, located in a large CNV, survived an FDR adjustment (rs507230, $P_{joint} = 1.42\times10^{-04}$, $P_{FDR} = 2.03910^{-02}$) (Table 3). Furthermore, this SNP shows a significant association in the meta-analysis of all cohorts ($P = 3.98\times10^{-12}$) (Table 4). This variant does not show LD ($r^2 > 0.2$) with any known functional SNP in *CD44*. This receptor has direct interactions with four reported (10) IFN-inducible molecules: CD9, MMP9, VCAN, and EPB41L3. Furthermore, a ligand for CD44, SPP1, has previously been established as an SLE risk locus involved in IFN pathways (8;33).

A variant in the karyopherin alpha 1 (KPNA1) region showed association in Stage 1 and Stage 2 analyses (rs6810306, $P_{joint} = 1.67E^{-04}$, $P_{FDR} = 2.68\times10^{-02}$) (Table 3) and was subsequently confirmed as associated in the meta-analysis ($P = 4.91\times10^{-02}$) (Table 4). This variant locates 15 kb upstream of KPNA1, in PARP9, in a 120 kb region of very high LD ($r^2 > 0.8$) comprising both genes. KPNA1 is a direct regulator of the interferon signature genes STAT1 (10;11;18) and STAT2 (10;18).

The *ubiquitin-conjugating enzyme E2L 3* (UBE2L3) is another associated locus, but we chose not to attempt to confirm it since our most significant variant has already been established as associated with SLE (3;34).

Interaction and ADTree analyses

The two-locus interaction analysis revealed replicated interactions not due to linkage disequilibrium between pairs of SNPs (Table 5). The most interesting pairs include serpin peptidase inhibitor, clade D, member 1 (SERPIND1, also known as heparin cofactor II, HCF2) with beta-parvin (PARVB), and Fyn-related kinase (FRK) with receptor-type protein-tyrosine phosphatase D (PTPRD). An intronic variant in the SERPIND1 gene showed interaction with a variant 37 kb upstream of actin-binding protein PARVB gene, a focal adhesion protein. Both molecules are involved in cellular growth and proliferation. The interacting SNPs in FRK and PTPRD are also intronic. Interestingly, the latter lies in a CpG island and a region of CNV. None of these pairs are known to date to have any direct biological interaction, and none have been previously associated with SLE. Although these interactions did not remain significant after adjusting for the number of combinations tested, they passed all quality control criteria and were replicated.

After exclusion of the HLA and chromosome X, the ADTree approach identified SNPs in the *IRF5-TNPO3* and *ITGAM* loci as the best discriminators between case-control status, thus confirming the association results. rs10488631 and rs4728142, in the IRF5 region, were found in 16% and 8% of the bootstrap samples, respectively, and rs9888739 and rs9937837, in ITGAM, were present in 10% and 8% of the bootstrap samples. Three novel loci were chosen in 6% of all bootstrap data sets: rs11605818 at ATG16L2-FCHSD2, rs11655550 at MED1-CRKRS, and rs2651843 at TSPAN32. Also, rs9888739 in ITGAM and rs2850724 in NFATC1 were present in the same structural feature in 6% of the bootstrap samples, suggesting a potential interaction in the Discovery Cohort data.

Discussion

The discovery that a signature of coordinately overexpressed IFN-inducible genes is prominent in a substantial fraction of lupus patients has fueled interest in the IFN pathway as a potential target for therapeutic intervention. This molecular signature is correlated with increased disease activity and specific clinical manifestations such as low complement levels, high levels of anti-dsDNA autoantibodies, higher sedimentation rates, and increased renal complications (35). The goal of this study was to identify the genetic variation that leads to the dysregulation of the IFN-related pathways and genes, including IFN-inducible genes and their direct regulators. This essentially Bayesian approach of selecting candidate genes based on prior knowledge serves to increase the reliability and likelihood of finding genes truly associated with disease (36–38). To our knowledge, this study represents the most comprehensive IFN pathway-based genetic analysis to date. Using independent cohorts for discovery and replication, we have evaluated a total of 1,754 genes. Eight genes were confirmed as associated with SLE.

The overall most significant associations of SLE with IFN-related genes were observed with SNPs located in regions that have been previously reported and firmly established as risk factors (*HLA*, *IRF5-TNPO3*, *ITGAM-ITGAX*, *STAT4*, *and TNFAIP3*). Two additional genes with previous evidence for association include the ubiquitin enzyme *UBE2L3* gene (3) and *IRF8* (34). The majority of additional loci we report represent novel genetic associations with SLE, underscoring the power of this candidate pathway approach.

The strongest novel genetic effect locates in the CD44 gene. CD44 is an integral cell membrane glycoprotein important for cell-cell interactions. As a key regulator of many

molecules, including IFNG and LCK, CD44 has important roles in lymphocyte activation, recirculation and homing, hematopoiesis, tumor metastasis and inflammation. It has also been shown that T cells from SLE patients display an increased and abnormal distribution of CD44 (39). Kaufman et al. (40) genotyped 4 SNPs within CD44 in 13 African-American families and found no evidence of association. However, given the number of SNPs and families tested, this study was likely underpowered to detect the effect that we report.

Another significant association signal was identified in the heparin-binding pleiotrophin (PTN) gene. PTN is a developmentally regulated cytokine with fibrinolytic, anti-apoptotic, mitogenic, transforming, angiogenic, and chemotactic activities (41). PTN has recently been shown to induce the expression of inflammatory cytokines including TNF-α, IL-1β and IL-6 in quiescent human peripheral blood mononuclear cells (PBMCs) (42). Other than being a regulator of several of the IFN-signature genes, no link between PTN and SLE has been established to date. Nevertheless, its expression is upregulated in experimental autoimmune encephalomyelitis (EAE) (43).

We have also confirmed an association with *DNAJ* (also known as *heat shock protein* (*Hsp) 40) homolog, subfamily A, member 1* gene. DnaJ is a heat shock protein that assists the chaperone Hsp70 in protein translation, folding, unfolding, translocation, and degradation. As a stress response protein, DnaJ is involved in repair and removal of damaged proteins, and is therefore important for maintaining cell viability. Hsps are potential targets of an autoimmune response and have been implicated in the induction and propagation of autoimmunity in several diseases, including rheumatoid arthritis and type 1 diabetes (44). Experimental evidence suggests that improper protein folding may promote autoimmunity (45). Thus, DnaJ and related proteins have potentially important but poorly understood roles in autoimmune diseases.

Karyopherin alpha 1 (KPNA1) binds RAG1, a lymphoid-specific recombinase essential for V(D)J recombination, and influences its sub-nuclear localization, hence controlling the generation of immunoglobulins and T cell receptors (46). It has also been shown to bind activated STAT1 and IRF5 proteins and transport them to the nucleus (47;48).

Even though our objective was to perform a comprehensive analysis of all IFN related genes, we cannot exclude the possibility that strong associations were missed due to the genomic coverage of the genotyping arrays or the *a priori* selection of specific genes found in the literature to be IFN-related or interactors. We, therefore, could have missed some unknown interactions. In addition, the dysregulation of IFN pathway genes is not a uniform feature across all lupus patients, and as such we would expect to detect moderate genetic effects that affect probably half of our cases. Nonetheless, replication of the novel effects we have identified in a second cohort, correction for multiple comparisons, and confirmation in a third cohort increase our confidence in the robustness of these associations.

In addition to the conventional statistical approaches, we used a data mining approach, Alternating Decision Trees (ADTrees), to try to corroborate the association results and identify novel variants that best discriminate cases vs. controls, as well as confirm and unveil potential interactions between genetic variants. The ADTrees validate the association results, and replication analyses are underway to confirm the uncovered two-locus interactions.

In summary, we have identified multiple IFN pathway-related genes that show confirmatory evidence for association with SLE. For the majority of loci, this is the first report of a genetic association with SLE with confirmation. Taken together, these new data expand the growing list of genes that show association with SLE, and emphasize the genetic contribution of dysregulated IFN pathways. Understanding how these genetic factors might

contribute to pathogenesis should ultimately lead to important opportunities for developing therapeutic targets to control the active IFN signature seen in SLE patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funded by the Alliance for Lupus Research [grant numbers 156395 and 52104] and as part of the International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEN, www.slegen.org), the National Institutes of Health [grant numbers AR043274, RR020143, R01 AI063274, AR043814, R01 AR44804, K24 AR02175, P60 AR0533008, 5 M01 RR-00079, P01 AR49084, R01 AR445650, AR042460, AI024717, AR049084, AR062277, K24 AR 002138, P60 2 AR30692, P01 AR49084, UL1RR025741], the US Department of Veterans Affairs, the Kirkland Scholar Program, the Lupus Research Institute, the Lupus Foundation of Minnesota, the Swedish Research Council, the Swedish Association Against Rheumatism, the Gustaf-Vth-80th-Year Jubilee, the Instituto de Salud Carlos III [grant number PS09/00129], the Arthritis Research Campaign [grant number 17761], the Wellcome Trust [grant number 085492], and the Center for Public Health Genomics at Wake Forest University Health Sciences.

We are extremely grateful to Miranda Marion, Joshua Grab and Lingyi Lu for assistance with analyses of these data, and to Debbie McDuffie for assistance with the genotyping assays.

References

1. Helmick CG, Felson DT, Lawrence RC, Gabriel S, Hirsch R, Kwoh CK, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. Arthritis Rheum. 2008; 58(1):15–25. [PubMed: 18163481]
2. Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, et al. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. Nat Genet. 2008; 40(9):1059–61. [PubMed: 19165918]
3. 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(2):204–10. [PubMed: 18204446]
4. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, et al. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. N Engl J Med. 2008; 358(9):900–9. [PubMed: 18204098]
5. Kozyrev SV, Abelson AK, Wojcik J, Zaghlool A, Linga Reddy MV, Sanchez E, et al. Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nat Genet. 2008; 40(2):211–6. [PubMed: 18204447]
6. Nath SK, Han S, Kim-Howard X, Kelly JA, Viswanathan P, Gilkeson GS, et al. A nonsynonymous functional variant in integrin-alpha(M) (encoded by ITGAM) is associated with systemic lupus erythematosus. Nat Genet. 2008; 40(2):152–4. [PubMed: 18204448]
7. Wakeland EK, Liu K, Graham RR, Behrens TW. Delineating the genetic basis of systemic lupus erythematosus. Immunity. 2001; 15(3):397–408. [PubMed: 11567630]
8. Moser KL, Kelly JA, Lessard CJ, Harley JB. Recent insights into the genetic basis of systemic lupus erythematosus. Genes Immun. 2009; 10(5):373–9. [PubMed: 19440199]
9. Baechler EC, Batliwalla FM, Reed AM, Peterson EJ, Gaffney PM, Moser KL, et al. Gene expression profiling in human autoimmunity. Immunol Rev. 2006; 210:120–37. [PubMed: 16623768]
10. Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proc Natl Acad Sci U S A. 2003; 100(5):2610–5. [PubMed: 12604793]
11. Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, Banchereau J, et al. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. J Exp Med. 2003; 197(6):711–23. [PubMed: 12642603]

12. Feng X, Wu H, Grossman JM, Hanvivadhanakul P, FitzGerald JD, Park GS, et al. Association of increased interferon-inducible gene expression with disease activity and lupus nephritis in patients with systemic lupus erythematosus. Arthritis Rheum. 2006; 54(9):2951–62. [PubMed: 16947629]
13. Han GM, Chen SL, Shen N, Ye S, Bao CD, Gu YY. Analysis of gene expression profiles in human systemic lupus erythematosus using oligonucleotide microarray. Genes Immun. 2003; 4(3):177–86. [PubMed: 12700592]
14. Ishii T, Onda H, Tanigawa A, Ohshima S, Fujiwara H, Mima T, et al. Isolation and expression profiling of genes upregulated in the peripheral blood cells of systemic lupus erythematosus patients. DNA Res. 2005; 12(6):429–39. [PubMed: 16769699]
15. Kirou KA, Lee C, George S, Louca K, Papagiannis IG, Peterson MG, et al. Coordinate overexpression of interferon-alpha-induced genes in systemic lupus erythematosus. Arthritis Rheum. 2004; 50(12):3958–67. [PubMed: 15593221]
16. Mandel M, Gurevich M, Pauzner R, Kaminski N, Achiron A. Autoimmunity gene expression portrait: specific signature that intersects or differentiates between multiple sclerosis and systemic lupus erythematosus. Clin Exp Immunol. 2004; 138(1):164–70. [PubMed: 15373920]
17. Nikpour M, Dempsey AA, Urowitz MB, Gladman DD, Barnes DA. Association of a gene expression profile from whole blood with disease activity in systemic lupus erythaematosus. Ann Rheum Dis. 2008; 67(8):1069–75. [PubMed: 18063674]
18. Nzeusseu TA, Galant C, Theate I, Maudoux AL, Lories RJ, Houssiau FA, et al. Identification of distinct gene expression profiles in the synovium of patients with systemic lupus erythematosus. Arthritis Rheum. 2007; 56(5):1579–88. [PubMed: 17469140]
19. Peterson KS, Huang JF, Zhu J, D'Agati V, Liu X, Miller N, et al. Characterization of heterogeneity in the molecular pathogenesis of lupus nephritis from transcriptional profiles of laser-captured glomeruli. J Clin Invest. 2004; 113(12):1722–33. [PubMed: 15199407]
20. Ronnblom L, Alm GV. A pivotal role for the natural interferon alpha-producing cells (plasmacytoid dendritic cells) in the pathogenesis of lupus. J Exp Med. 2001; 194(12):F59–F63. [PubMed: 11748288]
21. Niewold TB, Hua J, Lehman TJ, Harley JB, Crow MK. High serum IFN-alpha activity is a heritable risk factor for systemic lupus erythematosus. Genes Immun. 2007; 8(6):492–502. [PubMed: 17581626]
22. Kyogoku C, Tsuchiya N. A compass that points to lupus: genetic studies on type I interferon pathway. Genes Immun. 2007; 8(6):445–55. [PubMed: 17581625]
23. Kozyrev SV, Alarcon-Riquelme ME. The genetics and biology of Irf5-mediated signaling in lupus. Autoimmunity. 2007; 40(8):591–601. [PubMed: 18075793]
24. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet. 2007; 39(7):906–13. [PubMed: 17572673]
25. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B. 1995; 57:289–300.
26. Rasmussen A, Sevier S, Kelly JA, Glenn SB, Aberle T, Cooney CM, et al. The Lupus Family Registry and Repository. Rheumatology (Oxford). 2010
27. Alarcon GS, McGwin G Jr, Petri M, Reveille JD, Ramsey-Goldman R, Kimberly RP. Baseline characteristics of a multiethnic lupus cohort: PROFILE. Lupus. 2002; 11(2):95–101. [PubMed: 11958584]
28. 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(11): 1271–7. [PubMed: 7138600]
29. Freund, Y.; Mason, L. The alternating decision tree learning algorithm. Proceedings of the Sixteenth International Conference in Machine Learning; 1999. p. 124-33.
30. Pfahringer B, Holmes G, Kirkby R. Optimizing the induction of alternating decision trees. Proceedings of the Fifth Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD2001). 2001:477–87.

31. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med. 2007; 357(10):977–86. [PubMed: 17804842]
32. Namjou B, Sestak AL, Armstrong DL, Zidovetzki R, Kelly JA, Jacob N, et al. High-density genotyping of STAT4 reveals multiple haplotypic associations with systemic lupus erythematosus in different racial groups. Arthritis Rheum. 2009; 60(4):1085–95. [PubMed: 19333953]
33. Kariuki SN, Moore JG, Kirou KA, Crow MK, Utset TO, Niewold TB. Age- and gender-specific modulation of serum osteopontin and interferon-alpha by osteopontin genotype in systemic lupus erythematosus. Genes Immun. 2009; 10(5):487–94. [PubMed: 19339987]
34. Gateva V, Sandling JK, Hom G, Taylor KE, Chung SA, Sun X, et al. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. Nat Genet. 2009; 41(11):1228–33. [PubMed: 19838195]
35. Petri M, Singh S, Tesfasyone H, Dedrick R, Fry K, Lal P, et al. Longitudinal expression of type I interferon responsive genes in systemic lupus erythematosus. Lupus. 2009; 18(11):980–9. [PubMed: 19762399]
36. Colhoun HM, McKeigue PM, Davey SG. Problems of reporting genetic associations with complex outcomes. Lancet. 2003; 361(9360):865–72. [PubMed: 12642066]
37. Freimer N, Sabatti C. The use of pedigree, sib-pair and association studies of common diseases for genetic mapping and epidemiology. Nat Genet. 2004; 36(10):1045–51. [PubMed: 15454942]
38. Armstrong DL, Reiff A, Myones BL, Quismorio FP Jr, Klein-Gitelman M, McCurdy D, et al. Identification of new SLE-associated genes with a two-step Bayesian study design. Genes Immun. 2009; 10(5):446–56. [PubMed: 19440200]
39. Crispin JC, Tsokos GC. Novel molecular targets in the treatment of systemic lupus erythematosus. Autoimmun Rev. 2008; 7(3):256–61. [PubMed: 18190888]
40. Kaufman KM, Rankin J, Harley IT, Kelly JA, Harley JB, Scofield RH. A genetic marker within the CD44 gene confirms linkage at 11p13 in African-American families with lupus stratified by thrombocytopenia, but genetic association with CD44 is not present. Genes Immun. 2002; 3 (Suppl 1):S86–S88. [PubMed: 12215908]
41. Kadomatsu K, Muramatsu T. Midkine and pleiotrophin in neural development and cancer. Cancer Lett. 2004; 204(2):127–43. [PubMed: 15013213]
42. Achour A, M'bika JP, Baudouin F, Caruelle D, Courty J. Pleiotrophin induces expression of inflammatory cytokines in peripheral blood mononuclear cells. Biochimie. 2008; 90(11–12):1791–5. [PubMed: 18486628]
43. Liu X, Mashour GA, Webster HF, Kurtz A. Basic FGF and FGF receptor 1 are expressed in microglia during experimental autoimmune encephalomyelitis: temporally distinct expression of midkine and pleiotrophin. Glia. 1998; 24(4):390–7. [PubMed: 9814819]
44. Rajaiah R, Moudgil KD. Heat-shock proteins can promote as well as regulate autoimmunity. Autoimmun Rev. 2009; 8(5):388–93. [PubMed: 19121415]
45. Todd DJ, Lee AH, Glimcher LH. The endoplasmic reticulum stress response in immunity and autoimmunity. Nat Rev Immunol. 2008; 8(9):663–74. [PubMed: 18670423]
46. Jones JM, Simkus C. The roles of the RAG1 and RAG2 "non-core" regions in V(D)J recombination and lymphocyte development. Arch Immunol Ther Exp (Warsz). 2009; 57(2):105–16. [PubMed: 19333736]
47. Nardozzi J, Wenta N, Yasuhara N, Vinkemeier U, Cingolani G. Molecular basis for the recognition of phosphorylated STAT1 by importin alpha5. J Mol Biol. 2010; 402(1):83–100. [PubMed: 20643137]
48. Yeon SI, Youn JH, Lim MH, Lee HJ, Kim YM, Choi JE, et al. Development of monoclonal antibodies against human IRF-5 and their use in identifying the binding of IRF-5 to nuclear import proteins karyopherin-alpha1 and -beta1. Yonsei Med J. 2008; 49(6):1023–31. [PubMed: 19108028]

Table 1

Number of IFN Pathway-related genes and SNPs attempted, captured and replicated in Stages 1 and 2.

	Set 1	Set 2
# genes compiled	459	1,761
# genes attempted	415	1,739
Attempted loci*	304	1,450
SLEGEN SNPs	4,759	26,265
MN SNPs	4,680	25,196
QC[1] joint SNPs	2,169	12,997
Replicated[2] SNPs	56	249
Replicated Loci	26	119

*As defined by the interval between 50 kb upstream and 10 kb downstream each gene; 1 – SNPs that met quality criteria (QC) described in the Methods section for the joint-analysis; 2 - SNPs that met QC and *P*<0.05 in both cohorts and joint-analysis. Used SNPs with P<0.05 under best model.

Table 2

Most significant replicated SNPs in Set 1.

Marker[1]	Ch	Position (Mb)	Locus	MA	Stage 1 P-value[2]	Stage 2 P-value[2]	Joint-analysis			
							Case MAF	Control MAF	P-value[2]	OR [95% CI]
rs4659444*	1	26.698	RPS6KA1	T	1.48E-02	5.97E-03^{D}	0.24	0.21	1.14E-03^{D}	1.29 [1.11–1.50]
rs4970514^{S2}*	1	26.719	RPS6KA1	T	2.36E-02	6.43E-03^{D}	0.24	0.20	3.61E-03^{D}	1.26 [1.08–1.47]
rs2285210^{S2}	1	173.218	RABGAP1L	C	3.11E-03^{R}	3.35E-02^{R}	0.32	0.29	2.71E-04^{R}	1.64 [1.26–2.14]
rs10971259^{S2}*	9	32.969	DNAJA1	T	8.66E-03	3.14E-02	0.12	0.15	4.89E-03	0.81 [0.69–0.94]
rs10124051^{S2}	9	32.997	DNAJA1	C	2.87E-04	4.53E-02	0.08	0.11	1.12E-04	0.72 [0.60–0.85]
rs366078^{S2}	16	84.522	IRF8	C	1.15E-02	4.20E-03^{D}	0.10	0.13	2.41E-04	0.74 [0.63–0.87]

Table excludes MHC, IRF5-TNPO3, ITGAM-ITGAX, TNFAIP3 and STAT4 *loci*. The reported *P-values* are not adjusted for the number of comparisons, but except for the markers denoted with *, all met a FDR-adjusted threshold of significance in the joint-analysis, as described in the Methods. The smallest *P-value* is presented and, unless noted otherwise, it is under the additive model. OR and CI calculated under the model presented. Ch – chromosome; Mb – Megabases; MA –minor allele; MAF – Minor allele frequency; OR – odds ratio; CI – 95% confidence interval;

[1]The superscript after the marker denotes if it was imputed in any cohort: ^{S1}Stage 1, ^{S2}Stage 2;

[2]The superscript after the P-value denotes its genetic model, when other than the additive: ^{D}dominant, ^{R}recessive.

Table 3

Most significant replicated SNPs in Set 2.

Marker[1]	Ch	Position (Mb)	Locus	MA	Stage 1 P-value[2]	Stage 2 P-value[2]	Joint-analysis			
							Case MAF	Control MAF	P-value[2]	OR [95% CI]
rs6810306	3	123.731	KPNA1	G	1.25E-03^{D}	7.85E-03^{D}	0.44	0.41	1.67E-04^{D}	1.35 [1.15–1.58]
rs919581^{S2}	7	136.631	PTN	G	1.94E-03^{R}	8.66E-03^{R}	0.20	0.18	5.86E-05^{R}	2.39 [1.56–3.66]
rs2613310	9	92.670	SYK	C	1.34E-03	5.49E-04^{D}	0.25	0.20	1.53E-05	1.33 [1.17–1.52]
rs507230	11	35.086	CD44	G	1.25E-05^{D}	1.85E-02^{R}	0.44	0.49	1.42E-04^{D}	0.72 [0.61–0.86]
rs755690^{S2}	19	43.859	ACTN4	A	1.60E-03^{R}	3.87E-03^{R}	0.46	0.42	1.34E-04^{R}	1.46 [1.20–1.77]
rs749701^{S2}	19	43.882	ACTN4	T	2.40E-03^{R}	6.51E-03^{R}	0.46	0.42	2.05E-04^{R}	1.44 [1.19–1.75]
rs5754217^{S2}	22	20.270	UBE2L3	T	1.22E-04	3.72E-02^{R}	0.23	0.19	8.01E-05	1.31 [1.14–1.49]
rs2298428^{S2}	22	20.313	UBE2L3	T	3.65E-04	3.72E-02^{R}	0.22	0.18	3.07E-04^{R}	2.08 [1.40–3.10]

Table excludes the MHC region. The reported *P-values* are not adjusted for the number of comparisons, but all met a FDR-adjusted threshold of significance in the joint-analysis, as described in the Methods. The smallest *P-value* is presented and, unless noted otherwise, it is under the additive model. OR and CI calculated under the model presented. Ch – chromosome; Mb – Megabases; MA – minor allele; MAF – Minor allele frequency; OR – odds ratio; CI – 95% confidence interval;

[1] The superscript after the marker denotes if it was imputed in any cohort: ^{S1}Stage 1, ^{S2}Stage 2;

[2] The superscript after the P-value denotes its genetic model, when other than the additive: ^{D}dominant, ^{R}recessive.

Table 4

Results of the confirmation analysis.

				Stage 3						Meta-analysis	
Marker	**Ch**	**Position (Mb)**	**Locus**	**Case N**	**Control N**	**Case MAF**	**Control MAF**	**P-value**	**OR [95%CI]**	**P-value**	**OR [95%CI]**
Set 1											
rs4659444*	1	26.698	RPS6KA1	739	0	0.22	na	na	na	2.14E-03	1.21 [1.07–1.36]
rs2285210	1	173.218	RABGAP1L	1194	3267	0.31	0.30	3.63E-01	1.11 [0.89–1.39]	1.98E-01	1.35 [1.13–1.60]
rs10971259	9	32.969	DNAJA1	1636	2943	0.15	0.14	5.35E-02	1.13 [1.00–1.27]	6.31E-03	0.96 [0.87–1.05]
rs10124051	9	32.997	DNAJA1	2111	3481	0.11	0.11	2.26E-01	1.08 [0.95–1.22]	8.68E-02	0.89 [0.81–0.99]
rs366078*	16	84.522	IRF8	727	0	0.13	na	na	na	9.63E-03	1.18 [1.04–1.34]
Set 2											
rs6810306	3	123.731	KPNA1	1211	3260	0.43	0.41	1.64E-01	1.10 [0.96–1.27]	4.91E-02	1.22 [1.10–1.36]
rs919581	7	136.631	PTN	726	2935	0.19	0.16	1.44E-02	1.69 [1.11–2.57]	5.38E-04	2.02 [1.50–2.73]
rs2613310	9	92.67	SYK	1210	3267	0.22	0.22	6.21E-01	1.03 [0.92–1.15]	4.84E-01	1.17 [1.07–1.27]
rs507230	11	35.086	CD44	1190	3259	0.44	0.50	9.31E-07	0.70 [0.60–0.80]	3.98E-12	0.71 [0.63–0.79]
rs755690	19	43.859	ACTN4	1199	3260	0.45	0.44	5.23E-01	1.06 [0.89–1.25]	3.66E-01	1.24 [1.09–1.41]
rs749701	19	43.882	ACTN4	1207	3262	0.45	0.44	5.23E-01	1.06 [0.89–1.25]	3.67E-01	1.23 [1.08–1.40]

Ch – chromosome; Mb – Megabases; N – sample size; MAF – Minor allele frequency; OR – odds ratio; CI – 95% confidence interval; na – not available; P-values, OR and CI reported were calculated under the same genetic model reported for the joint-analysis of Stages 1 and 2 (Tables 2 and 3); *Given the absence of Stage 3 results, the final results reported under “meta-analysis” reflect a joint-analysis between Stages 1, 2, and the cases from Stage 3.

Table 5

Replicated results of the two-locus logistic regression interaction analysis for most significant non-MHC pairs.

SNP1				SNP2				Stage 1		Stage 2	
ID	Ch	Pos (kb)	locus	ID	Ch	Pos (kb)	locus	P	OR [CI]	P	OR [CI]
rs165862*	22	19460	SERPIND1	rs2073082*	22	42685	PARVB	3.00E-04	1.50 [1.21–1.86]	3.40E-03	1.60 [1.17–2.18]
rs1544155	6	116453	FRK	rs1021720	9	8391	PTPRD	2.00E-04	1.46 [1.19–1.79]	9.50E-03	1.47 [1.10–1.97]

Ch – chromosome; Pos – position; OR – odds ratio; CI – 95% confidence interval.

*Stage 1 r^2=0.00038.