

## Research Article

# Gene-Expression-Guided Selection of Candidate Loci and Molecular Phenotype Analyses Enhance Genetic Discovery in Systemic Lupus Erythematosus

Yelena Koldobskaya,<sup>1</sup> Kichul Ko,<sup>1</sup> Akaash A. Kumar,<sup>1</sup> Sandra Agik,<sup>1</sup> Jasmine Arrington,<sup>1</sup> Silvia N. Kariuki,<sup>1</sup> Beverly S. Franek,<sup>1</sup> Marissa Kumabe,<sup>1</sup> Tammy O. Utset,<sup>1</sup> Meenakshi Jolly,<sup>2</sup> Andrew D. Skol,<sup>3</sup> and Timothy B. Niewold<sup>1</sup>

<sup>1</sup> Section of Rheumatology and Gwen Knapp Center for Lupus and Immunology Research, University of Chicago, Chicago, IL 60637, USA

<sup>2</sup> Section of Rheumatology and Rush Lupus Clinic, Rush University, Chicago, IL 60612, USA

<sup>3</sup> Section of Genetic Medicine, University of Chicago, Chicago, IL 60637, USA

Correspondence should be addressed to Timothy B. Niewold, [tniewold@medicine.bsd.uchicago.edu](mailto:tniewold@medicine.bsd.uchicago.edu)

Received 15 June 2012; Accepted 17 July 2012

Academic Editor: George N. Goulielmos

Copyright © 2012 Yelena Koldobskaya et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Systemic lupus erythematosus (SLE) is a highly heterogeneous autoimmune disorder characterized by differences in autoantibody profiles, serum cytokines, and clinical manifestations. We have previously conducted a case-case genome-wide association study (GWAS) of SLE patients to detect associations with autoantibody profile and serum interferon alpha (IFN- $\alpha$ ). In this study, we used public gene expression data sets to rationally select additional single nucleotide polymorphisms (SNPs) for validation. The top 200 GWAS SNPs were searched in a database which compares genome-wide expression data to genome-wide SNP genotype data in HapMap cell lines. SNPs were chosen for validation if they were associated with differential expression of 15 or more genes at a significance of  $P < 9 \times 10^{-5}$ . This resulted in 11 SNPs which were genotyped in 453 SLE patients and 418 matched controls. Three SNPs were associated with SLE-associated autoantibodies, and one of these SNPs was also associated with serum IFN- $\alpha$  ( $P < 4.5 \times 10^{-3}$  for all). One additional SNP was associated exclusively with serum IFN- $\alpha$ . Case-control analysis was insensitive to these molecular subphenotype associations. This study illustrates the use of gene expression data to rationally select candidate loci in autoimmune disease, and the utility of stratification by molecular phenotypes in the discovery of additional genetic associations in SLE.

## 1. Introduction

Systemic lupus erythematosus (SLE) is a severe multisystem autoimmune disease of unknown etiology. Genetic factors clearly play a role in susceptibility, and a number of genetic loci have been implicated in the disease [1]. Despite the successes of recent genetic association studies, only a fraction of the genetic liability for SLE has been explained to date. SLE is a heterogeneous disease clinically, and there is strong evidence that the molecular pathogenesis of the condition varies considerably between individuals as well. For example, specific autoantibodies are formed in some patients and not others, and these autoantibody specificities have been

associated with clinical features of the disease [2, 3]. In addition, approximately half of adult patients with SLE demonstrate overactivity of the interferon alpha (IFN- $\alpha$ ) pathway in their peripheral blood [2, 4]. Interestingly, high IFN- $\alpha$  and SLE-associated autoantibodies are heritable as traits in SLE families and can be found in family members who are not affected by SLE [5, 6]. Autoantibodies can be found in sera for many years prior to the clinical diagnosis of SLE [7], and it is thought that some of the autoantibodies may be themselves directly pathogenic. IFN- $\alpha$  is a cytokine involved in viral defense, capable of bridging the innate and adaptive immune systems [8]. Interestingly, when recombinant human IFN- $\alpha$  has been given as a treatment for chronic

viral hepatitis, some treated individuals have developed de novo SLE, which frequently resolves upon discontinuation of the IFN- $\alpha$  [9, 10]. These data support the concept that both IFN- $\alpha$  and SLE-associated autoantibodies represent causal factors in human SLE. Additionally, both IFN- $\alpha$  and SLE-associated autoantibodies are heritable within SLE families supporting a genetic contribution, and thus the idea that these molecular measurements could be used as a phenotype in genetic studies.

In previous work, we have begun to map genetic variants which are associated with high IFN- $\alpha$  and with the presence of particular autoantibodies in SLE patients [11–13]. Some well-established genetic risk factors for SLE have been associated with one or both of these molecular phenotypes [14–18]. In addition, we have performed a genome-wide association study (GWAS) using these two molecular traits as phenotypes to enable discovery of novel genetic variants associated with IFN- $\alpha$  and SLE-associated autoantibodies [19]. A number of novel genes have been validated from this screen to date [19, 20], although much of the variance in both serum IFN- $\alpha$  and the presence or absence of particular autoantibodies remains to be explained.

In prioritizing genetic variants to be followed up in our GWAS scan, we used gene ontology and expert literature search to prioritize variants which were in or near genes related to immune responses. This was based upon the supposition that SLE is an autoimmune disease, and many of the well-validated loci which have emerged from unbiased studies to date encode genes with immune function. This approach has some limitations, as genetic variations which were not near known genes were not prioritized, nor were those which did not have known function within the immune system. It is clear that genetic variants can sometimes impact the expression of a gene which is not nearby, and these genetic variants may be assigned to irrelevant nearby genes in gene ontology analysis. Additionally, many genes which could be critical to human disease pathogenesis may still be unstudied and unknown, and thus unlikely to be prioritized in follow up candidate studies.

To address these possibilities in our GWAS validation, we searched our top 200 SNPs in a public database which links genome-wide SNP data from the HapMap project to genome-wide gene expression data from the HapMap lymphoblastoid B-cell lines (SCAN) database, [21]. Genes which are disease associated are more commonly associated with alternate gene expression than genes which are not disease associated [22], and thus genes from our top 200 which were strongly associated with differences in gene expression should be more likely to be true associations. In this study, we leverage gene expression data sets to prioritize additional candidates from our trait-stratified GWAS for validation in an independent cohort. We found eleven SNPs which were significantly associated with alternate gene expression of multiple transcripts in public databases, and had not been prioritized for followup in our initial GWAS screen. Four of these eleven SNPs were significantly associated with the important molecular subphenotypes IFN- $\alpha$  and SLE-associated autoantibodies in our independent validation cohort, validating this method of genetic discovery.

## 2. Methods

**2.1. Initial GWAS Study Description.** The initial cohort of SLE patients studied in the GWAS scan was obtained from the Hospital for Special Surgery Lupus Registries, and consisted of 104 SLE patients [19]. This study was designed as a case-case analysis to compare SNP frequencies in SLE patients with high versus low IFN- $\alpha$  and those with and without SLE-associated autoantibodies. Patients were selected in an extremes-of-phenotype design from the top 33% and bottom 33% of serum IFN- $\alpha$  activity and were additionally stratified for the GWAS study by ancestry and the presence or absence of anti-RBP or anti-dsDNA antibodies. A study design incorporating multiple ancestral backgrounds was chosen as both autoantibodies and serum IFN- $\alpha$  levels are heritable pathogenic factors which are shared between all ancestral backgrounds. The top 200 SNPs were examined in detail using expert review of public databases, and seven top SNPs chosen for replication using a gene-centric algorithm demonstrated strong associations with either serology or serum IFN- $\alpha$  in an independent cohort, as would have been expected based upon the initial GWAS study design [19].

**2.2. Validation Cohort.** The independent validation cohort of 453 SLE patients was obtained from the University of Chicago Translational Research in the Department of Medicine (TRIDOM) registry and Rush University Medical Center and consisted of 282 African-American and 171 European-American SLE patients. All patients met the revised 1982 ACR criteria for the diagnosis of SLE [23]. Samples from 418 controls were obtained from the TRIDOM registry, including 300 African-American and 118 European-American subjects who were individually screened for the absence of autoimmune disease by medical record review. The subjects in this study were not related to each other. Informed consent was obtained from all subjects at each site, and the study was approved by the IRB at each institution.

**2.3. SCAN Database Query.** We searched the top 200 SNPs from the GWAS described above as query terms in the SNP and CNV Annotation (SCAN) database (<http://www.scandb.org/>) [21]. This database is a searchable index of genome-wide gene expression data linked to genome-wide SNP genotype data from the HapMap project. Gene expression data is derived from studies in which gene expression arrays were run on Epstein-Barr virus-transformed lymphoblastoid cell lines from individuals genotyped in the HapMap project. The SCAN database contains expression data from both European (Centre d'Etude du Polymorphisme Humain or CEPH) and West African (Yoruba or YRI) HapMap reference populations. We used a threshold  $P$  value of  $P < 9 \times 10^{-5}$  and searched both CEPH and YRI population datasets for each SNP. Because SNPs associated with alternate gene expression are more likely to be disease or trait associated [22], we selected SNPs which were associated with alternate expression of 15 or more transcripts in the SCAN database. This resulted in 11 SNPs,

and for each SNP at least one of the 15 or more associated transcripts was involved in immune function.

**2.4. SNP Genotyping in the Validation Cohort.** Individuals in the validation cohort were genotyped at the rs9521996, rs11199974, rs7785392, rs9568401, rs4892122, rs4778708, rs1340981, rs1408806, rs4894215, rs1569428, and rs1159916 SNPs. Genotyping was performed using ABI TaqMan Assays-by-Design primers and probes on an ABI 7900HT PCR machine with >98% genotyping success. Scatter plots were all reviewed individually for quality, and genotype frequencies did not deviate significantly from the expected Hardy-Weinberg proportions ( $P > 0.01$  in controls across all ancestral backgrounds).

**2.5. Reporter Cell Assay for IFN- $\alpha$ .** The reporter cell assay for IFN- $\alpha$  has been described in detail elsewhere [5, 24]. Reporter cells were used to measure the ability of patient sera to cause IFN-induced gene expression. The reporter cells (WISH cells, ATCC #CCL-25) were cultured with 50% patient sera for 6 hours and then lysed. mRNA was purified from cell lysates, and cDNA was made from total cellular mRNA. cDNA was then quantified using real-time PCR using an Applied Biosystems 7900HT PCR machine with the SYBR Green fluorophore system. Forward and reverse primers for the genes *MX1*, *PKR*, and *IFIT1*, which are known to be highly and specifically induced by IFN- $\alpha$ , were used in the reaction [5]. *GAPDH* was amplified in the same samples to control for background gene expression. The amount of PCR product of the IFN- $\alpha$ -induced gene was normalized to the amount of product for the housekeeping gene *GAPDH* in the same sample. The relative expression of each of the three tested IFN-induced genes was calculated as a fold increase compared to its expression in WISH cells cultured with media alone. Results from the IFN- $\alpha$  assay were standardized to a healthy multiethnic reference population as previously described, and a serum IFN- $\alpha$  activity score was calculated based upon the mean and SD of the reference population [5]. This assay has been highly informative when applied to SLE as well as other autoimmune disease populations [5, 25–27].

**2.6. Measurement of Autoantibodies.** Antibodies to anti-Ro, anti-La, anti-Sm, and anti-RNP were measured in all samples by ELISA methods using kits from INOVA Diagnostics (San Diego, CA), and anti-dsDNA antibodies were measured using *Crithidia luciliae* immunofluorescence, with detectable fluorescence considered positive. All samples were assayed in University of Chicago clinical laboratory by the same personnel that test clinical samples. For the ELISA assays, the standard cutoff points for a positive test designated by the manufacturer were used to categorize samples as positive or negative.

**2.7. Statistical Analysis.** To control for population structure and effects related to admixture, we used a principal component analysis of SNPs which varied in frequency by ancestral background. All subjects in the study had genotype data

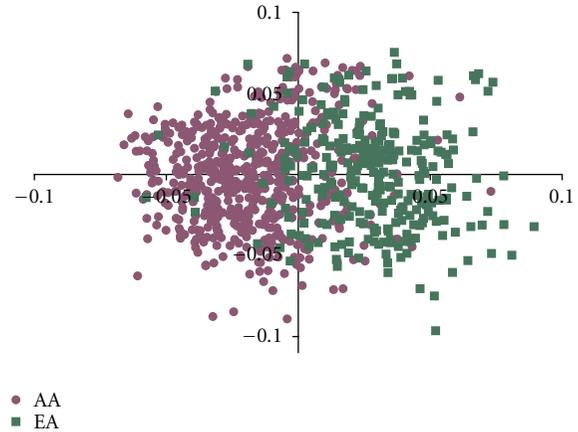


FIGURE 1: Principal component analysis of SNPs genotyped in all cases and controls. Component 1 is shown on the  $x$ -axis, and component 2 is shown on the  $y$ -axis. Each dot represents one subject, and the dots are color-coded by the self-reported ancestry of that subject.

available for 30 such SNPs, and principal component analysis was performed using the PCA option in the Cluster program by Eisen et al. [28]. The first two principal components are shown plotted on the  $x$  and  $y$  axes, respectively, in Figure 1, and the first component provides a strong separation of those subjects of self-reported African-American ancestry from those of self-reported European-American ancestry. We included the first and second principal components as covariates in all subsequent association analyses to provide control for differences in proportional ancestry in both cases and controls.

Logistic regression models were used to detect associations with SLE in case-control analysis or in case-case analyses examining the SLE-associated autoantibody traits and serum IFN- $\alpha$  activity. The SLE-associated autoantibodies anti-Ro, anti-La, anti-Sm, anti-RNP, and anti-dsDNA were all tested for association with each SNP in logistic regression models. Serum IFN- $\alpha$  was binned as high or low, using 2SD above the mean of healthy donor sera as the cutoff point, and then used as the outcome variable in logistic regression. Significant relationships observed in this regression were then explored by comparing quantitative IFN- $\alpha$  data between genotype categories. The IFN- $\alpha$  data was nonnormally distributed, and nonparametric Mann-Whitney  $U$  was used to compare quantitative IFN- $\alpha$  data between genotype subgroups.  $P$  values shown in the paper are uncorrected for multiple comparisons. To establish significance and account for multiple comparisons, we used a threshold  $P$  value of  $P < 4.5 \times 10^{-3}$  to allow for a type I error rate of 0.05 following a Bonferroni correction for the number of SNPs tested in this study.

### 3. Results

**3.1. Three of Eleven SNPs Demonstrate Association with Autoantibody Traits in SLE Patients.** We used logistic regression to detect associations between autoantibody traits and

TABLE 1: Summary of SNPs associated with autoantibody traits.

SNP	Chr.	Nearby Gene	Ancestry	Autoantibody	Odds ratio	$P$ value
rs9521996 C	13	ANKRD10	AA	Anti-RNP	2.01	$8.0 \times 10^{-4}$
rs1408806 G	9	TYRP1	EA	Anti-Sm	3.48	$1.5 \times 10^{-3}$
rs4894215 G	2	—	EA	Anti-Ro	2.16	$2.5 \times 10^{-3}$

SNP: single nucleotide polymorphism, chr.: chromosome, autoantibody: the antibody specificity associated with the particular SNP, odds ratio and  $P$ -value are calculated from the logistic regression model.

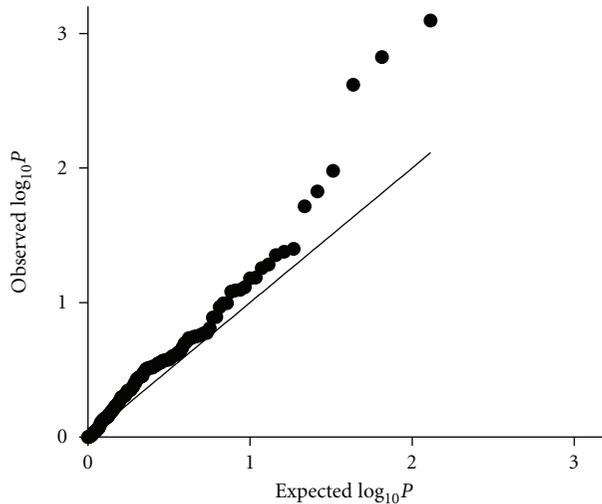


FIGURE 2: Q-Q plot showing the observed versus expected  $P$  values in the autoantibody analysis.  $P$  values that would be expected under the null hypothesis (no association between SNPs and autoantibody traits) are represented by the line, and the observed  $P$  values are represented by dots, one for each tested SNP-autoantibody association.

genotype at each of the 11 SNPs in each ancestral background separately. Three SNPs demonstrated associations which would withstand a Bonferroni correction for multiple comparisons correcting for the number of SNPs tested ( $P < 4.5 \times 10^{-3}$ , Table 1). Figure 2 shows a Q-Q plot of the distribution of probabilities observed in the antibody analyses versus the null distribution. In Figure 2, the top three SNP-antibody associations highlighted in Table 1 are represented by the three dots with the highest values on the  $y$ -axis which clearly deviate from the null distribution.

**3.2. Two SNPs Are Associated with Serum IFN- $\alpha$  in SLE Patients.** Regression models were also used to assess the association of each of the 11 SNPs with serum IFN- $\alpha$  activity in SLE patients. An association was observed between rs9568401 G and high serum IFN- $\alpha$  in both African and European Americans. In European-Americans, the rs1408806 G allele which was associated with anti-Sm antibodies was also associated with increased serum IFN- $\alpha$ . These associations are illustrated in Figure 3, which shows quantitative IFN- $\alpha$  by genotype category. The minor allele frequency of each SNP was relatively low, and thus minor allele homozygotes were rare and are combined with heterozygotes in this graph. Dominant or recessive models

could not be assessed due to the rarity of homozygous minor allele subjects, and the graph is not meant to represent a dominant relationship. While the rs1408806 SNP is also associated with an autoantibody trait, the rs9568401 SNP was not associated with any autoantibodies and was exclusively associated with serum IFN- $\alpha$  activity.

**3.3. Multiple Subphenotype Modeling Supports Complex Association Patterns between Genetic Variants, Autoantibodies, and Serum IFN- $\alpha$  Activity.** With regard to the rs1408806 G allele which was associated with both serum IFN- $\alpha$  and anti-Sm in European ancestry, the association between these two phenotypes appeared to be independent (Figure 4(a)). Given the strong relationship between serum IFN- $\alpha$  and autoantibodies in SLE [4], we also examined serum IFN- $\alpha$  in the context of the other SNP-autoantibody relationships we had identified rs9521996/anti-RNP in African Americans and rs4894215/anti-Ro in European-Americans, (Figures 4(b) and 4(c) resp.). Both of these SNPs demonstrated evidence for a secondary association with serum IFN- $\alpha$  which was dependent upon the associated autoantibody. Summarizing the four SNPs which demonstrate significant associations following multiple comparison corrections, one SNP is associated with serum IFN- $\alpha$  alone, two are associated with autoantibody profiles which are associated with higher IFN- $\alpha$ , and one SNP is associated with both serum IFN- $\alpha$  and autoantibody profile independently. These relationships are depicted in Figure 5.

**3.4. SCAN Database Search Results Predicted the Ancestral Background in Which the SLE Phenotype Association Was Observed.** The SCAN database search examined both European- and African-derived populations, and the SNPs which were associated with SLE subphenotypes were associated with alternate gene expression in the SCAN database in only one ancestral background. In each of the autoantibody associations, the ancestral background in which the autoantibody association was observed in SLE patients was the same ancestral background in which differential gene expression was observed in the SCAN database (Table 2). The association between rs9568401 and serum IFN- $\alpha$  was observed in both ancestral backgrounds, but the SNP was only associated with alternate gene expression in the SCAN database in African ancestry subjects. Overall this general concordance in ancestral backgrounds between the SLE phenotype associations further supports the idea that the SNPs which impact gene expression in human cell lines are more likely to be associated with molecular phenotypes in human

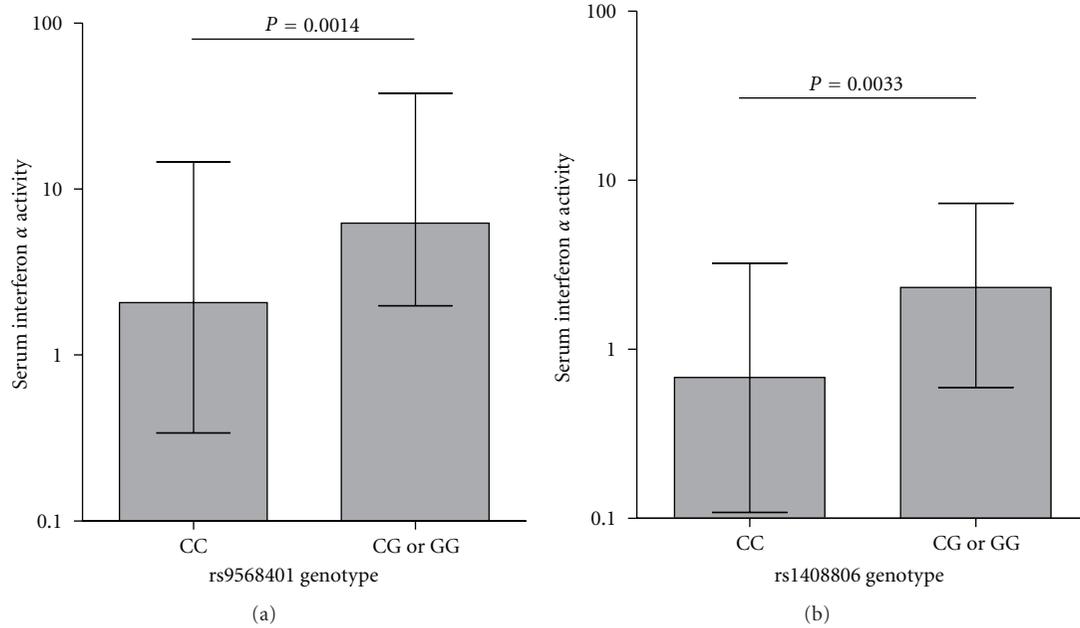


FIGURE 3: Serum IFN- $\alpha$  activity in SLE patients stratified by SNP genotype at rs9568401 (a) and rs1408806 (b). Minor allele homozygotes are combined with heterozygotes on the graph. Bars show the median error bars show the interquartile range.  $P$  value by Mann-Whitney  $U$  test.

disease. Representative transcripts that were differentially regulated by each associated SNP in the SCAN database are also shown in Table 2.

**3.5. Case-Control Analysis Does Not Show Large Differences in Allele Frequencies When Comparing All SLE Patients to Controls.** As shown in Table 3, we did not observe significant case-control associations for any of the 11 studied SNPs which would withstand statistical correction for multiple comparisons (all  $P > 4.5 \times 10^{-3}$ ). The initial GWAS was designed to detect associations with either autoantibodies or serum IFN- $\alpha$ , and the SNPs we followed up were most strongly associated with these traits. The lack of strong case-control associations at the same SNPs supports the idea that the genetic effects we observe are relevant to patient subsets, and that the power to discover these SNPs would be more limited in a standard case-control study design.

#### 4. Conclusions

In this study, we identify novel genetic variants associated with molecular phenotypes in SLE in two different ancestral backgrounds, using gene expression data as a guide for rational candidate gene selection from a previous GWAS study. In published overall case-control studies of SLE to date, there are examples of both shared associations across ancestral backgrounds [29], and associations which are particular to one or a few ancestral backgrounds [20, 30, 31]. In our study, it is striking that we did not find many associations which were shared between ancestral backgrounds and the majority were distinct to one ancestral background, despite

studying molecular phenotypes which are shared across ancestral backgrounds. This would support the hypothesis that while similar molecular pathways may be dysregulated in SLE patients of different ancestral backgrounds, the particular steps in that pathway which are dysregulated may differ by ancestry. These differences would be important to appreciate as we envision personalized therapy using agents which target these pathways, such as the category of anti-IFN- $\alpha$  drugs which are being developed for SLE currently. Presumably many autoimmune disease risk alleles which are common in the population have been maintained due to some benefit in increasing immune responses in response to pathogens. Infectious disease has been a major selective force in human history, and it seems likely that different world populations may have developed and selected for different immune system polymorphisms which could result in a similar end pathway output. A striking example of this type of human convergent evolution has been shown in the case of the human lactase gene [32], in which lactase persistence in adulthood was conferred by a number of different polymorphisms that had arisen separately in different world populations, converging upon a similar end pathway result.

Heterogeneity is not unexpected in SLE, as clinically the syndrome is very diverse. Overall case-control genetic studies are likely to be significantly limited due to heterogeneity, as different polymorphisms will be more or less relevant in different patient groups. In the case of physical phenotypes, a number of studies support the idea that different genetic variants will be associated with particular clinical disease manifestations, such as rash, renal disease, and others [33–36]. Diversity in autoantibody and cytokine phenotypes

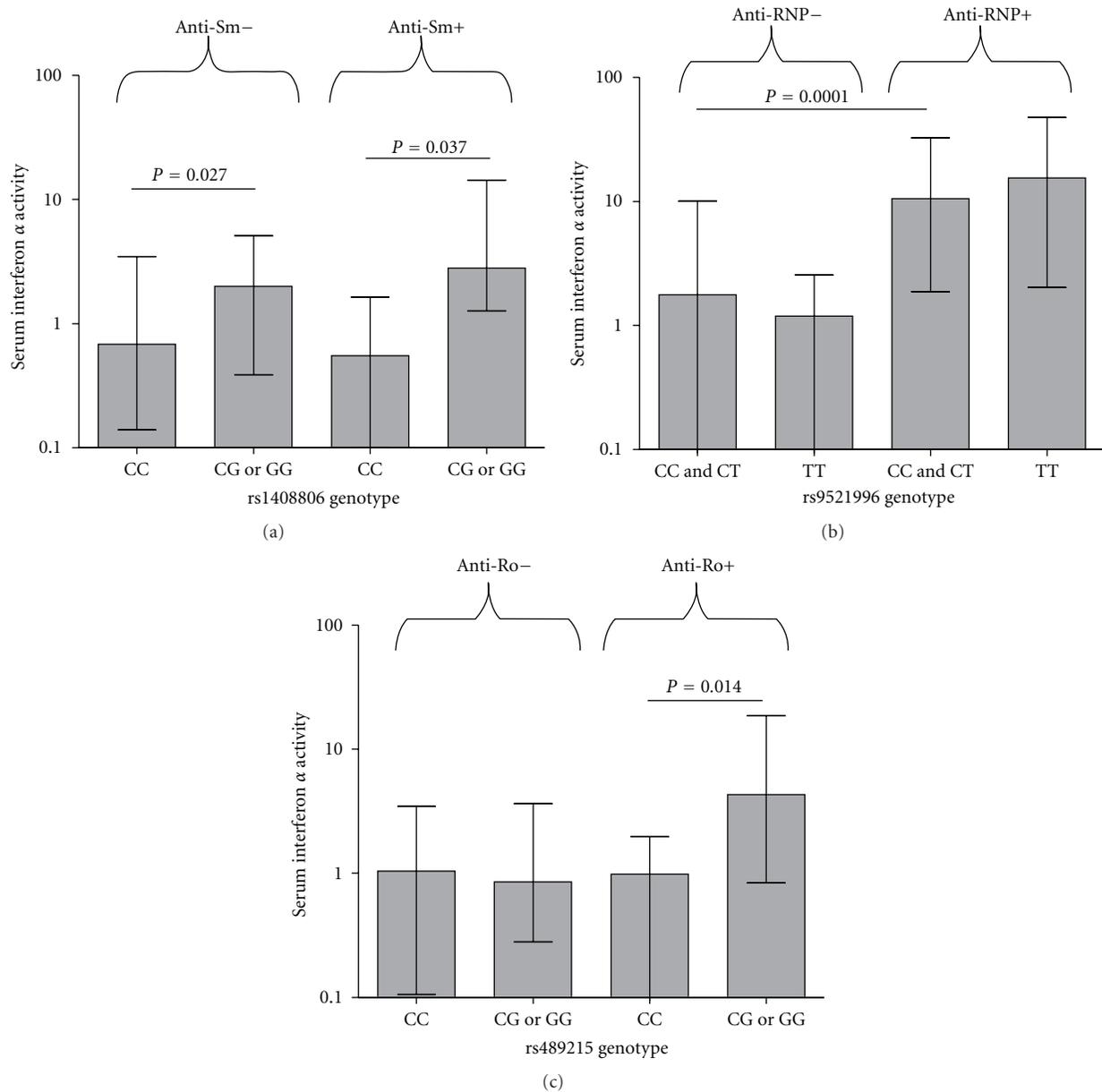


FIGURE 4: Serum IFN- $\alpha$  activity in SLE patients stratified by SNP genotype and the autoantibody associated with that particular SNP. Minor allele homozygotes are combined with heterozygotes on the graph. Bars show the median error bars show the interquartile range.  $P$  value by Mann-Whitney  $U$  test.

TABLE 2: Summary of the 4 SNPs associated with SLE phenotypes and the SCAN database results regarding ancestral background and representative associated transcripts.

SNP	Chr.	Nearby Gene	SLE association ancestry	Associated phenotype	SCAN ancestry	Representative SCAN transcripts
rs9521996 C	13	ANKRD10	AA	Anti-RNP	YRI	IRF3, MIF
rs1408806 G	9	TYRP1	EA	Anti-Sm	CEPH	CASP3, RIPK1
rs4894215 G	2	None within 200kb	EA	Anti-Ro	CEPH	HLADRB1, HLADQB1
rs9568401 G	13	DLEU2	EA, AA	IFN	YRI	IRAK2, NOD2

SNP: single nucleotide polymorphism, chr.: chromosome, SLE association ancestry: the ancestral background in which the SNP was associated with an SLE phenotype, SCAN ancestry: the ancestral background in which that SNP was associated with alternate gene expression, representative SCAN transcripts: genes which differentially expressed due to genotype at that SNP in the SCAN database; two transcripts of the >15 were chosen for inclusion in this table, with an emphasis on those transcripts with immune system relevance.

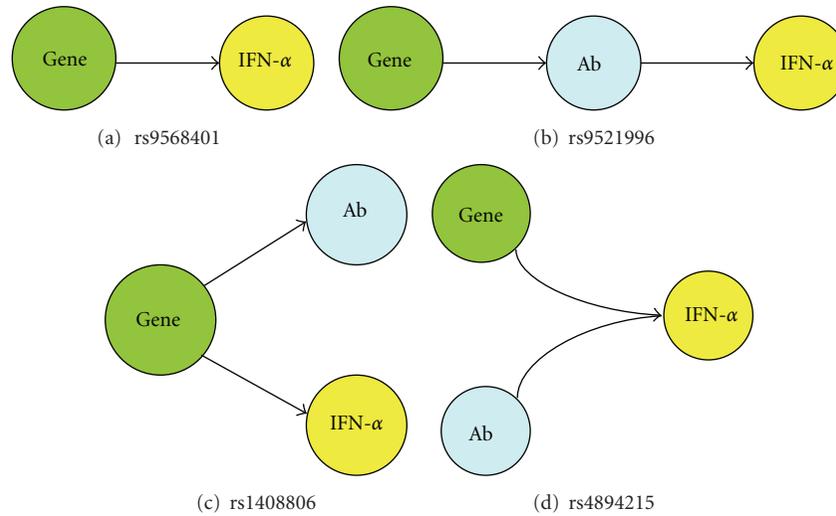


FIGURE 5: Diagrams depicting patterns of association observed between SNP genotype, autoantibodies, and serum IFN- $\alpha$ . Gene = SNP genotype at the indicated SNP, Ab = the particular autoantibody associated with that SNP, and arrows indicate the associations observed in the study.

TABLE 3: Case control analysis of 11 SNPs tested in this study in each ancestral background.

SNP	African Americans			European Americans		
	MAF	OR	<i>P</i> value	MAF	OR	<i>P</i> -value
rs9521996 C	0.285	1.02	0.86	0.136	1.46	0.12
rs11199974 G	0.258	0.89	0.44	0.482	1.11	0.56
rs7785392 T	0.473	0.80	0.084	0.612	0.78	0.16
rs9568401 G	0.122	0.74	0.12	0.085	1.04	0.90
rs4892122 G	0.279	1.14	0.32	0.295	1.19	0.39
rs4778708 T	0.407	0.95	0.68	0.268	1.10	0.64
rs1340981 A	0.161	0.92	0.61	0.397	0.88	0.47
rs1408806 G	0.174	0.80	0.19	0.246	0.85	0.44
rs4894215 G	0.358	0.94	0.64	0.430	1.08	0.67
rs1569428 G	0.341	0.70	0.0070	0.430	0.92	0.68
rs1159916 C	0.405	0.74	0.018	0.333	0.88	0.49

SNP: single nucleotide polymorphism, MAF: minor allele frequency in controls, OR: odds ratio, as calculated from the logistic regression model.

between SLE patients is also well recognized [4, 37, 38]. In this study we examine these two molecular phenotypes and find genetic associations which are relevant to subgroups of patients defined by these molecular characteristics. We have previously demonstrated strong subsetting of genetic risk related to molecular phenotypes in SLE in the case of the IRF5 gene. The majority of the genetic risk of SLE related to IRF5 was found within a particular serologic subgroup which constituted 40% of the overall SLE patient group studied [14]. This gene had been well validated as an SLE-risk gene in previous overall case-control studies [39, 40], but was later shown to have a very strong subgroup effect [14]. It seems likely that this phenomenon will be more widespread, and that many genetic loci could be very difficult to discover in an overall unstratified case-control study. Other autoimmune diseases such as rheumatoid arthritis have already set a strong precedent for the importance of serologic subsets in genetic analysis. The anti-CCP antibody positive versus

negative groups of rheumatoid arthritis patients demonstrate large differences in genetic association, including the HLA region [41]. The genes we report in this study have not been previously identified in case-control studies, and in our case-control analysis of these loci support we do not see large overall allele frequency differences. This does not mean that the loci are irrelevant, as they clearly impact important pathogenic subphenotypes in SLE. Instead, this supports the idea that “all cases versus all controls” study designs will have limits, and it is unlikely that we will be able to fully map the genetic architecture of complex diseases fully using only case-control designs, even if very large and well-powered cohorts are used. In summary, it seems likely that both physical or clinical phenotypes as well as molecular phenotypes will need to be incorporated in genetic study designs to address disease heterogeneity and enable continued genetic discovery in autoimmune disease.

Another benefit of including molecular subphenotypes and gene expression into genetic association studies is that the genetic loci discovered in this manner are immediately linked to some biological alteration. This is especially useful when genes which have not been previously studied are implicated, or if a particular associated genetic variant is not within or near a known gene. If these variants are found in an overall case-control analysis, it can be difficult to plan follow-up functional experiments if little is known about the function of that gene. In our study, we found SNPs which were not in obviously relevant genetic regions, but nonetheless impacted upon important molecular phenotypes and altered expression of immune system molecules. While we cannot know the mechanism by which the genetic variant impacts upon gene expression via our current study, these questions can be followed up and validated in functional experiments.

### Conflict of Interest

The authors report no conflict of interests.

### Acknowledgments

J. Arrington received NIH/NIDDK STEP-UP Scholar Award; S. N. Kariuki acknowledges receiving HHMI Gilliam Fellowship for Advanced Study; T. O. Utset acknowledges receiving Lupus Clinical Trials Consortium; T. B. Niewold acknowledges receiving NIH R01 AR060861, K08 AI083790, P30 DK42086, NIAID Clinical Research Loan Repayment AI071651, NIH CTSA Core Subsidy Grant, and CTSA Pilot Grants from UL1 RR024999, Lupus Research Institute Novel Research Grant, and an Alliance for Lupus Research Target Identification in Lupus Grant.

### References

- [1] K. L. Moser, J. A. Kelly, C. J. Lessard, and J. B. Harley, "Recent insights into the genetic basis of systemic lupus erythematosus," *Genes and Immunity*, vol. 10, no. 5, pp. 373–379, 2009.
- [2] K. Ko, B. S. Franek, M. Marion, K. M. Kaufman, C. D. Langefeld, J. B. Harley et al., "Genetic ancestry, serum interferon-alpha activity, and autoantibodies in systemic lupus erythematosus," *The Journal of Rheumatology*, vol. 39, no. 6, pp. 1238–1240, 2012.
- [3] D. Koffler, V. Agnello, R. I. Carr, and H. G. Kunkel, "Anti-DNA antibodies and renal lesions of patients with systemic lupus erythematosus," *Transplantation Proceedings*, vol. 1, no. 4, pp. 933–938, 1969.
- [4] C. E. Weckerle, B. S. Franek, J. A. Kelly et al., "Network analysis of associations between serum interferon- $\alpha$  activity, autoantibodies, and clinical features in systemic lupus erythematosus," *Arthritis & Rheumatism*, vol. 63, no. 4, pp. 1044–1053, 2011.
- [5] T. B. Niewold, J. Hua, T. J. A. Lehman, J. B. Harley, and M. K. Crow, "High serum IFN- $\alpha$  activity is a heritable risk factor for systemic lupus erythematosus," *Genes and Immunity*, vol. 8, no. 6, pp. 492–502, 2007.
- [6] P. S. Ramos, J. A. Kelly, C. Gray-McGuire et al., "Familial aggregation and linkage analysis of autoantibody traits in pedigrees multiplex for systemic lupus erythematosus," *Genes and Immunity*, vol. 7, no. 5, pp. 417–432, 2006.
- [7] M. R. Arbuckle, M. T. McClain, M. V. Rubertone et al., "Development of autoantibodies before the clinical onset of systemic lupus erythematosus," *The New England Journal of Medicine*, vol. 349, no. 16, pp. 1526–1533, 2003.
- [8] T. B. Niewold, D. N. Clark, R. Salloum, and B. D. Poole, "Interferon alpha in systemic lupus erythematosus," *Journal of Biomedicine and Biotechnology*, vol. 2010, Article ID 948364, 8 pages, 2010.
- [9] T. B. Niewold, "Interferon alpha-induced lupus: proof of principle," *Journal of Clinical Rheumatology*, vol. 14, no. 3, pp. 131–132, 2008.
- [10] T. B. Niewold and W. I. Swedler, "Systemic lupus erythematosus arising during interferon-alpha therapy for cryoglobulinemic vasculitis associated with hepatitis C," *Clinical Rheumatology*, vol. 24, no. 2, pp. 178–181, 2005.
- [11] S. N. Kariuki and T. B. Niewold, "Genetic regulation of serum cytokines in systemic lupus erythematosus," *Translational Research*, vol. 155, no. 3, pp. 109–117, 2010.
- [12] R. Salloum and T. B. Niewold, "Interferon regulatory factors in human lupus pathogenesis," *Translational Research*, vol. 157, no. 6, pp. 326–331, 2011.
- [13] T. B. Niewold, "Interferon alpha as a primary pathogenic factor in human lupus," *Journal of Interferon & Cytokine Research*, vol. 31, no. 12, pp. 887–892, 2011.
- [14] T. B. Niewold, J. A. Kelly, S. N. Kariuki, B. S. Franek, A. A. Kumar, K. M. Kaufman et al., "IRF5 haplotypes demonstrate diverse serological associations which predict serum interferon alpha activity and explain the majority of the genetic association with systemic lupus erythematosus," *Annals of the Rheumatic Diseases*, vol. 71, no. 3, pp. 463–469, 2012.
- [15] S. N. Kariuki, K. A. Kirou, E. J. MacDermott, L. Barillas-Arias, M. K. Crow, and T. B. Niewold, "Cutting edge: autoimmune disease risk variant of STAT4 confers increased sensitivity to IFN- $\alpha$  in lupus patients in vivo," *The Journal of Immunology*, vol. 182, no. 1, pp. 34–38, 2009.
- [16] R. Salloum, B. S. Franek, S. N. Kariuki et al., "Genetic variation at the IRF7/PHRF1 locus is associated with autoantibody profile and serum interferon- $\alpha$  activity in lupus patients," *Arthritis & Rheumatism*, vol. 62, no. 2, pp. 553–561, 2010.
- [17] S. Agik, B. S. Franek, A. A. Kumar, M. Kumabe, T. O. Utset, R. A. Mikolaitis et al., "The autoimmune disease risk allele of UBE2L3 in African American patients with systemic lupus erythematosus: a recessive effect upon subphenotypes," *The Journal of Rheumatology*, vol. 39, no. 1, pp. 73–78, 2012.
- [18] T. Robinson, S. N. Kariuki, B. S. Franek, M. Kumabe, A. A. Kumar, M. Badaracco et al., "Autoimmune disease risk variant of IFIH1 is associated with increased sensitivity to IFN-alpha and serologic autoimmunity in lupus patients," *The Journal of Immunology*, vol. 187, no. 3, pp. 1298–1303, 2011.
- [19] S. N. Kariuki, B. S. Franek, A. A. Kumar et al., "Trait-stratified genome-wide association study identifies novel and diverse genetic associations with serologic and cytokine phenotypes in systemic lupus erythematosus," *Arthritis Research and Therapy*, vol. 12, no. 4, article R151, 2010.
- [20] S. N. Kariuki, B. S. Franek, R. A. Mikolaitis et al., "Promoter variant of PIK3C3 is associated with autoimmunity against Ro and Sm epitopes in African-American lupus patients," *Journal of Biomedicine and Biotechnology*, vol. 2010, Article ID 826434, 7 pages, 2010.
- [21] E. R. Gamazon, W. Zhang, A. Konkashbaev et al., "SCAN: SNP and copy number annotation," *Bioinformatics*, vol. 26, no. 2, pp. 259–262, 2009.

- [22] D. L. Nicolae, E. Gamazon, W. Zhang, S. Duan, M. E. Dolan, and N. J. Cox, "Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS," *PLoS Genetics*, vol. 6, no. 4, Article ID e1000888, 2010.
- [23] E. M. Tan, A. S. Cohen, J. F. Fries et al., "The 1982 revised criteria for the classification of systemic lupus erythematosus," *Arthritis & Rheumatism*, vol. 25, no. 11, pp. 1271–1277, 1982.
- [24] J. Hua, K. Kirou, C. Lee, and M. K. Crow, "Functional assay of type I interferon in systemic lupus erythematosus plasma and association with anti-RNA binding protein autoantibodies," *Arthritis & Rheumatism*, vol. 54, no. 6, pp. 1906–1916, 2006.
- [25] T. B. Niewold, T. L. Rivera, J. P. Buyon, and M. K. Crow, "Serum type I interferon activity is dependent on maternal diagnosis in anti-SSA/Ro-positive mothers of children with neonatal lupus," *Arthritis & Rheumatism*, vol. 58, no. 2, pp. 541–546, 2008.
- [26] T. B. Niewold, S. N. Kariuki, G. A. Morgan, S. Shrestha, and L. M. Pachman, "Elevated serum interferon- $\alpha$  activity in juvenile dermatomyositis: associations with disease activity at diagnosis and after thirty-six months of therapy," *Arthritis & Rheumatism*, vol. 60, no. 6, pp. 1815–1824, 2009.
- [27] X. Feng, N. P. Reider, M. Yanamandala, A. Hill, B. S. Franek, T. B. Niewold et al., "Type I interferon signature is high in lupus and neuromyelitis optica but low in multiple sclerosis," *Journal of the Neurological Sciences*, vol. 313, no. 1–2, pp. 48–53, 2012.
- [28] M. B. Eisen, P. T. Spellman, P. O. Brown, and D. Botstein, "Cluster analysis and display of genome-wide expression patterns," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 25, pp. 14863–14868, 1998.
- [29] E. Sanchez, M. E. Comeau, B. I. Freedman, J. A. Kelly, K. M. Kaufman, C. D. Langefeld et al., "Identification of novel genetic susceptibility loci in African American lupus patients in a candidate gene association study," *Arthritis & Rheumatism*, vol. 63, no. 11, pp. 3493–3501, 2011.
- [30] J. P. Lodolce, L. E. Kolodziej, L. Rhee et al., "African-derived genetic polymorphisms in *TNFAIP3* mediate risk for autoimmunity," *The Journal of Immunology*, vol. 184, no. 12, pp. 7001–7009, 2010.
- [31] J. Pothlichet, T. B. Niewold, D. Vitour, B. Solhonne, M. K. Crow, and M. Si-Tahar, "A loss-of-function variant of the antiviral molecule MAVS is associated with a subset of systemic lupus patients," *EMBO Molecular Medicine*, vol. 3, no. 3, pp. 142–152, 2011.
- [32] S. A. Tishkoff, F. A. Reed, A. Ranciaro et al., "Convergent adaptation of human lactase persistence in Africa and Europe," *Nature Genetics*, vol. 39, no. 1, pp. 31–40, 2007.
- [33] I. T. W. Harley, T. B. Niewold, R. M. Stormont, K. M. Kaufman, S. B. Glenn, B. S. Franek et al., "The role of genetic variation near interferon-kappa in systemic lupus erythematosus," *Journal of Biomedicine and Biotechnology*, vol. 2010, Article ID 706825, 11 pages, 2010.
- [34] T. Trivedi, B. S. Franek, S. L. Green, S. N. Kariuki, M. Kumabe, R. A. Mikolaitis et al., "Osteopontin alleles are associated with clinical characteristics in systemic lupus erythematosus," *Journal of Biomedicine and Biotechnology*, vol. 2011, Article ID 802581, 6 pages, 2011.
- [35] C. E. Weckerle and T. B. Niewold, "The unexplained female predominance of systemic lupus erythematosus: clues from genetic and cytokine studies," *Clinical Reviews in Allergy and Immunology*, vol. 40, no. 1, pp. 42–49, 2011.
- [36] E. Sanchez, A. Nadig, B. C. Richardson, B. I. Freedman, K. M. Kaufman, J. A. Kelly et al., "Phenotypic associations of genetic susceptibility loci in systemic lupus erythematosus," *Annals of the Rheumatic Diseases*, vol. 70, no. 10, pp. 1752–1757, 2011.
- [37] C. E. Weckerle, D. Imbuka, B. S. Franek, J. A. Kelly, M. Kumabe, J. A. James et al., "Large scale analysis of tumor necrosis factor alpha levels in systemic lupus erythematosus," *Arthritis Rheum*, vol. 64, no. 9, pp. 2947–2952, 2012.
- [38] L. L. Ritterhouse, S. R. Crowe, T. B. Niewold, J. T. Merrill, V. C. Roberts, A. B. Dedeker et al., "B lymphocyte stimulator levels in systemic lupus erythematosus: higher circulating levels in African American patients and increased production after influenza vaccination in patients with low baseline levels," *Arthritis & Rheumatism*, vol. 63, no. 12, pp. 3931–3941, 2011.
- [39] R. R. Graham, S. V. Kozyrev, E. C. Baechler et al., "A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus," *Nature Genetics*, vol. 38, no. 5, pp. 550–555, 2006.
- [40] S. Sigurdsson, H. H. H. Göring, G. Kristjansdottir et al., "Comprehensive evaluation of the genetic variants of interferon regulatory factor 5 (IRF5) reveals a novel 5 bp length polymorphism as strong risk factor for systemic lupus erythematosus," *Human Molecular Genetics*, vol. 17, no. 6, pp. 872–881, 2008.
- [41] L. Padyukov, M. Seielstad, R. T. H. Ong et al., "A genome-wide association study suggests contrasting associations in ACPA-positive versus ACPA-negative rheumatoid arthritis," *Annals of the Rheumatic Diseases*, vol. 70, no. 2, pp. 259–265, 2011.



**Hindawi**  
Submit your manuscripts at  
<http://www.hindawi.com>

