

# UCSF

## UC San Francisco Previously Published Works

### Title

Lupus Risk Variant Increases pSTAT1 Binding and Decreases ETS1 Expression

### Permalink

<https://escholarship.org/uc/item/3nh741p0>

### Journal

American Journal of Human Genetics, 96(5)

### ISSN

0002-9297

### Authors

Lu, Xiaoming  
Zoller, Erin E  
Weirauch, Matthew T  
et al.

### Publication Date

2015-05-01

### DOI

10.1016/j.ajhg.2015.03.002

Peer reviewed

# Lupus Risk Variant Increases pSTAT1 Binding and Decreases *ETS1* Expression

Xiaoming Lu,<sup>1,2</sup> Erin E. Zoller,<sup>2</sup> Matthew T. Weirauch,<sup>2,3</sup> Zhiguo Wu,<sup>4</sup> Bahram Namjou,<sup>2</sup> Adrienne H. Williams,<sup>5</sup> Julie T. Ziegler,<sup>5</sup> Mary E. Comeau,<sup>5</sup> Miranda C. Marion,<sup>5</sup> Stuart B. Glenn,<sup>6</sup> Adam Adler,<sup>6</sup> Nan Shen,<sup>2,7</sup> Swapan K. Nath,<sup>6</sup> Anne M. Stevens,<sup>8,9</sup> Barry I. Freedman,<sup>10</sup> Betty P. Tsao,<sup>11</sup> Chaim O. Jacob,<sup>12</sup> Diane L. Kamen,<sup>13</sup> Elizabeth E. Brown,<sup>14,15</sup> Gary S. Gilkeson,<sup>13</sup> Graciela S. Alarcón,<sup>15</sup> John D. Reveille,<sup>16</sup> Juan-Manuel Anaya,<sup>17</sup> Judith A. James,<sup>6,18</sup> Kathy L. Sivils,<sup>6</sup> Lindsey A. Criswell,<sup>19</sup> Luis M. Vilá,<sup>20</sup> Marta E. Alarcón-Riquelme,<sup>6,21</sup> Michelle Petri,<sup>22</sup> R. Hal Scofield,<sup>6,18,23</sup> Robert P. Kimberly,<sup>15</sup> Rosalind Ramsey-Goldman,<sup>24</sup> Young Bin Joo,<sup>25</sup> Jeongim Choi,<sup>25</sup> Sang-Cheol Bae,<sup>25</sup> Susan A. Boackle,<sup>26</sup> Deborah Cunninghame Graham,<sup>27</sup> Timothy J. Vyse,<sup>27</sup> Joel M. Guthridge,<sup>6</sup> Patrick M. Gaffney,<sup>6</sup> Carl D. Langefeld,<sup>5</sup> Jennifer A. Kelly,<sup>6</sup> Kenneth D. Greis,<sup>28</sup> Kenneth M. Kaufman,<sup>2,29</sup> John B. Harley,<sup>2,29,30</sup> and Leah C. Kottyan<sup>2,29,30,\*</sup>

Genetic variants at chromosomal region 11q23.3, near the gene *ETS1*, have been associated with systemic lupus erythematosus (SLE), or lupus, in independent cohorts of Asian ancestry. Several recent studies have implicated *ETS1* as a critical driver of immune cell function and differentiation, and mice deficient in *ETS1* develop an SLE-like autoimmunity. We performed a fine-mapping study of 14,551 subjects from multi-ancestral cohorts by starting with genotyped variants and imputing to all common variants spanning *ETS1*. By constructing genetic models via frequentist and Bayesian association methods, we identified 16 variants that are statistically likely to be causal. We functionally assessed each of these variants on the basis of their likelihood of affecting transcription factor binding, miRNA binding, or chromatin state. Of the four variants that we experimentally examined, only rs6590330 differentially binds lysate from B cells. Using mass spectrometry, we found more binding of the transcription factor signal transducer and activator of transcription 1 (STAT1) to DNA near the risk allele of rs6590330 than near the non-risk allele. Immunoblot analysis and chromatin immunoprecipitation of pSTAT1 in B cells heterozygous for rs6590330 confirmed that the risk allele increased binding to the active form of STAT1. Analysis with expression quantitative trait loci indicated that the risk allele of rs6590330 is associated with decreased *ETS1* expression in Han Chinese, but not other ancestral cohorts. We propose a model in which the risk allele of rs6590330 is associated with decreased *ETS1* expression and increases SLE risk by enhancing the binding of pSTAT1.

## Introduction

Systemic lupus erythematosus (SLE [MIM 152700]) is a heterogeneous autoimmune disease characterized by hyperactive T and B cells, autoantibody production, immune com-

plex deposition, and multi-organ tissue damage.<sup>1</sup> The prevalence of SLE is 20–150 cases per 100,000 individuals.<sup>2–4</sup> Although the precise etiology of SLE remains unclear, the contribution of genetic factors has been shown by association and family studies;<sup>5</sup> the concordance rate

<sup>1</sup>Immunology Graduate Program, College of Medicine, University of Cincinnati, Cincinnati, OH 45229, USA; <sup>2</sup>Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA; <sup>3</sup>Divisions of Biomedical Informatics and Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA; <sup>4</sup>Division of Immunobiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA; <sup>5</sup>Center for Public Health Genomics and Department of Biostatistical Sciences, Wake Forest School of Medicine, Wake Forest Baptist Medical Center, Winston-Salem, NC 27157, USA; <sup>6</sup>Arthritis and Clinical Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA; <sup>7</sup>Joint Molecular Rheumatology Laboratory, Institute of Health Sciences and Shanghai Renji Hospital, Shanghai Jiao Tong University School of Medicine in collaboration with Shanghai Institutes for Biological Sciences and Chinese Academy of Sciences, Shanghai 200031, PRC; <sup>8</sup>Center for Immunity and Immunotherapies, Seattle Children's Research Institute, Seattle, WA 98101, USA; <sup>9</sup>Division of Rheumatology, Department of Pediatrics, University of Washington, Seattle, WA 98195, USA; <sup>10</sup>Section on Nephrology, Department of Internal Medicine, Wake Forest School of Medicine, Wake Forest Baptist Medical Center, Winston-Salem, NC 27157, USA; <sup>11</sup>Division of Rheumatology, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, 90095, USA; <sup>12</sup>Department of Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089, USA; <sup>13</sup>Division of Rheumatology, Medical University of South Carolina, Charleston, SC 29425, USA; <sup>14</sup>Division of Molecular and Cellular Pathology, Department of Pathology, University of Alabama at Birmingham, Birmingham, AL 35294, USA; <sup>15</sup>Department of Medicine, University of Alabama at Birmingham, Birmingham, AL 35294, USA; <sup>16</sup>Rheumatology and Clinical Immunogenetics, University of Texas Health Science Center at Houston, Houston, TX 77030, USA; <sup>17</sup>Center for Autoimmune Diseases Research, Universidad del Rosario, 110010 Bogota, Colombia; <sup>18</sup>Department of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA; <sup>19</sup>Rosalind Russell / Ephraim P. Engleman Rheumatology Research Center, Department of Medicine, University of California, San Francisco, San Francisco, CA 94117, USA; <sup>20</sup>Division of Rheumatology, Department of Medicine, School of Medicine, University of Puerto Rico, Medical Sciences Campus, San Juan, PR 00936, USA; <sup>21</sup>Center for Genomics and Oncological Research, Pfizer-University of Granada-Junta de Andalucía, Granada 18016, Spain; <sup>22</sup>Department of Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD 21224, USA; <sup>23</sup>Oklahoma City Veterans Affairs Medical Center, Oklahoma City, OK 73104, USA; <sup>24</sup>Division of Rheumatology, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA; <sup>25</sup>Department of Rheumatology, Hanyang University Hospital for Rheumatic Diseases, Seoul 133-792, Korea; <sup>26</sup>Division of Rheumatology, School of Medicine, University of Colorado, Aurora, CO 80045, USA; <sup>27</sup>Divisions of Genetics and Molecular Medicine and Immunology, King's College London, Strand, London WC2R 2LS, UK; <sup>28</sup>Cancer Biology, University of Cincinnati Medical Center, Cincinnati, OH 45267, USA; <sup>29</sup>Cincinnati Veterans Affairs Medical Center, Cincinnati, OH 45220, USA

<sup>30</sup>These authors contributed equally to this work

\*Correspondence: [leah.kottyan@cchmc.org](mailto:leah.kottyan@cchmc.org)

<http://dx.doi.org/10.1016/j.ajhg.2015.03.002>. ©2015 by The American Society of Human Genetics. All rights reserved.

among monozygotic twins is between 20% and 59%, and siblings of affected subjects have an 8- to 30-fold higher risk of developing the disease than the general population.

Recently, several independent genome-wide association studies (GWASs) in Asian populations have confirmed that genetic variants in v-ets avian erythroblastosis virus E26 oncogene homolog 1 (*ETS1* [MIM 164720]) are associated with susceptibility to SLE.<sup>6–10</sup> These studies have established that the most strongly associated SNPs in *ETS1* are rs6590330 and rs1128334.

*ETS1* is known to play an important role in regulating immune cell proliferation and differentiation.<sup>11</sup> Moreover, *Ets1*-deficient mice develop a lupus-like disease characterized by high titers of immunoglobulin M (IgM) and IgG autoantibodies and immune-complex deposition in the kidneys.<sup>12</sup> Meanwhile, *ETS1* mRNA expression levels in peripheral-blood mononuclear cells (PBMCs) from SLE-affected individuals are considerably lower than those in healthy subjects.<sup>8</sup> Further, *ETS1* mRNA expression in PBMCs from chromosomes harboring lupus risk alleles is significantly lower than that in non-risk alleles of healthy subjects,<sup>8</sup> indicating that the risk variants at this locus are associated with reduced *ETS1* expression.

Previous studies have identified genetic association at *ETS1*, but no molecular mechanism has been presented to explain the association. Using a comprehensive dataset of genetic variants in this region in subjects from multiple ancestries, we employed frequentist and Bayesian fine-mapping strategies<sup>13</sup> to identify a set of variants that are likely to be causal.<sup>14</sup> This approach resulted in a total of 16 genetic variants comprising our credible set of most likely causal genetic variants. Importantly, we demonstrated that the minor allele of rs6590330, the most strongly associated genetic variant, leads to increased binding of the activated transcription factor signal transducer and activator of transcription 1 (pSTAT1), encoded by *STAT1* (MIM 600555), and is correlated with decreased *ETS1* expression. Altogether, our study provides insight into the mechanism driving the increased lupus risk at this locus in subjects of Asian ancestry.

## Material and Methods

### Subjects and Study Design

We used a large collection of samples from case and control subjects from multiple ethnic groups (Table S1). These samples were from the collaborative Large Lupus Association Study 2 (LLAS2)<sup>15</sup> and were contributed by participating institutions in the United States, Asia, and Europe. LLAS2, an SLE genetic-association study, used a candidate-gene approach to genotype 347 ancestral-informative markers and 31,851 candidate markers throughout the genome.<sup>16</sup> According to genetic ancestry, subjects were grouped into four ethnic groups, including European and European American (EU), African American (AA), Asian and Asian American (AS), and Hispanic American (HA). All SLE subjects met the American College of Rheumatology criteria for the classification of SLE<sup>17</sup> and were enrolled in this study through an

informed-consent process approved by the local institutional review boards.

### Genotyping of Genetic Variants and Sample Quality Control

We genotyped 69 SNPs covering the *ETS1* region (spanning 128.2–128.4 Mb on chromosome 11; GRCh37, UCSC Genome Browser hg19; Table S1) as part of a larger custom genotyping study. Specifically, the variants were chosen to span the association interval identified with the Infinium HumanHap330 array of the original GWAS that identified significant association at this locus. Genotyping of SNPs was completed with Infinium chemistry on an Illumina iSelect custom array according to the manufacturer's protocol. The following quality-control procedures were implemented for identifying SNPs for analysis: well-defined clusters for genotype calling, call rate > 90% across all samples genotyped, minor allele frequency (MAF) > 0.1%, and  $p < 0.05$  for differential missingness between case and control subjects. Markers with evidence of a departure from Hardy-Weinberg proportion expectation ( $p < 0.0001$  in control subjects) were removed from the initial analysis.

For LLAS2, we removed samples with a call rate < 90% or excess heterozygosity (the average call rate for *ETS1* was 99.3%). The remaining individuals were examined for excessive allele sharing as estimated by identity by descent (IBD). In sample pairs with excessive relatedness (IBD > 0.4), one individual was removed from the analysis on the basis of the following criteria: (1) remove the sample with the lower call rate, (2) remove the control sample and retain the case sample, (3) remove the male sample before the female sample, (4) remove the younger control sample before the older control sample, and (5) in a situation with two case samples, remove the sample whose available phenotype data are less complete.

### Ascertainment of Population Stratification

Genetic outliers from each ethnic and/or racial group were removed from further analysis as determined by principal-component (PC) analysis and admixture estimates, as previously described (Figure 1 in Lessard et al.<sup>16</sup> and McKeigue et al.<sup>18</sup> and Price et al.<sup>19</sup>). To distinguish the four continental ancestral populations, we used 347 ancestry-informative markers (AIMs) that were from the same custom genotyping study and that passed quality control in both EIGENSTRAT<sup>19</sup> and ADMIXMAP,<sup>20,21</sup> allowing identification of the substructure within the sample set.<sup>22,23</sup> The AIMs were selected to distinguish four continental ancestral populations: Africans, Europeans, American Indians, and East Asians. We utilized PCs from EIGENSTRAT outputs to identify outliers of each of the first three PCs for the individual population clusters with visual inspection.

### Statistical Analysis: Workflow

We initiated the analysis by assessing the association of genotyped variants in each of the four ancestral cohorts individually. Strategically, we analyzed the genotyped variants and then the imputed variants, performed full haplotype analysis, executed an analysis of linkage disequilibrium (LD), and finally built statistical models to account for the lupus-associated variability in each ancestry with genome-wide statistical association. In building the one-SNP models of association in the AS ancestry cohort, we comparatively evaluated every possible variant for its ability to account for the lupus-associated genetic variation.

### Statistical Analysis: Frequentist Approach

We tested each SNP for its association with SLE by using logistic regression models that included three estimates of admixture proportion as covariates as implemented in PLINK v.1.07<sup>24</sup> and SNPTEST v.2.5.<sup>25</sup> The additive genetic model was assessed as the initially tested model of inheritance for this locus. Other models were subsequently considered, but these were not found to be substantially superior.

We performed stepwise logistic regression to identify those single-nucleotide variants (SNVs) independently associated with the development of lupus in PLINK and SNPTEST. For these analyses, the allelic dosage(s) of specific variant(s), in addition to the admixture estimates, were added to the logistic model as covariates. LD and haplotypes were determined with Haploview v.4.2.<sup>26,27</sup> We calculated haplotype blocks for those haplotypes present at >3% frequency by using the four-gamete-rule algorithms with a minimum  $r^2$  value of 0.8. We performed haplotypic associations in PLINK by using a sliding-window approach and assessing the association of haplotypes defined by logistic regression, as described earlier.

### Statistical Analysis: Bayesian Approach

Using SNPTEST, we calculated the Bayesian factor (BF) for each genetic variant: we divided the probability of the genotype configuration at that genetic variant in case and control subjects under the alternative hypothesis that the genetic variant was associated with disease status by the probability of the genotype configuration at that genetic variant in case and control subjects under the null hypothesis that disease status was independent of genotype at that SNP (we used the methods developed and introduced in Maller et al.<sup>13</sup> and implemented in Kottyan et al.<sup>14</sup>). We used three admixture estimates as covariates, as we did for the frequentist approach. Large BF values correlate with robust evidence of association, given that small probabilities provide strong evidence in a frequentist approach. For well-powered studies, the BFs of relatively common variants are highly correlated with the p values (reviewed in Stephens and Balding<sup>28</sup>). We used the additive model. The linear predictor is  $\log(p_i / (1 - p_i)) = \mu + \beta G_i$ , and the prior is  $\mu \sim N(0, 1^2)$ ,  $\beta \sim N(0, 0.2^2)$  (variables are defined in SNPTEST and the supplemental note in Maller et al.<sup>13</sup>).

To identify the variants most likely to be driving the statistical association, we calculated a posterior probability under the assumption that any of the variants within a single genetic effect could be causal and that only one of these variants was causal for each genetic effect. Variants with a low posterior probability are highly unlikely to be causal regardless of the allele frequency or presence of the actual causal variant in the analysis, according to the procedure as presented.<sup>13</sup> Regardless of whether the causal variants have been genotyped in this experiment, variants with a low posterior probability are unlikely to be causal.<sup>13</sup>

### Imputation to Composite 1000 Genomes Reference Panel

To detect associated variants that were not directly genotyped, we imputed the *ETS1* region with IMPUTE2 by using a composite imputation reference panel based on 1000 Genomes Project sequence data from March 2012.<sup>25,29</sup> Imputed genotypes were included in the analysis if they had or exceeded a probability threshold of 0.5, an information measure of >0.4, and the same quality-control criteria described for the genotyped markers. We

used SNPTEST to incorporate the probability threshold from each imputed value into the statistical analysis.

### Electrophoretic Mobility Shift Assay

We annealed pairs of single-stranded 5' IRDye-700-infrared-dye-labeled 35-bp oligonucleotides (obtained from IDT) to generate double-stranded probes. We incubated 50 fmol of labeled probes with nuclear extract prepared from Epstein-Barr-virus-transformed B cell lines, poly dI-dC (a sodium salt complex of two strands, each with an alternating sequence of deoxyinosinic acid and deoxycytidylic acid), and buffers supplied by the Odyssey Infrared EMSA Kit (LI-COR Biosciences) according to LI-COR's recommended protocols. The binding reactions were analyzed by electrophoresis on 6% Tris-borate-EDTA polyacrylamide gels and detected by an infrared fluorescent procedure with the Odyssey Infrared Imaging System (LI-COR Biosciences). The oligonucleotide sequences are included in Table S2.

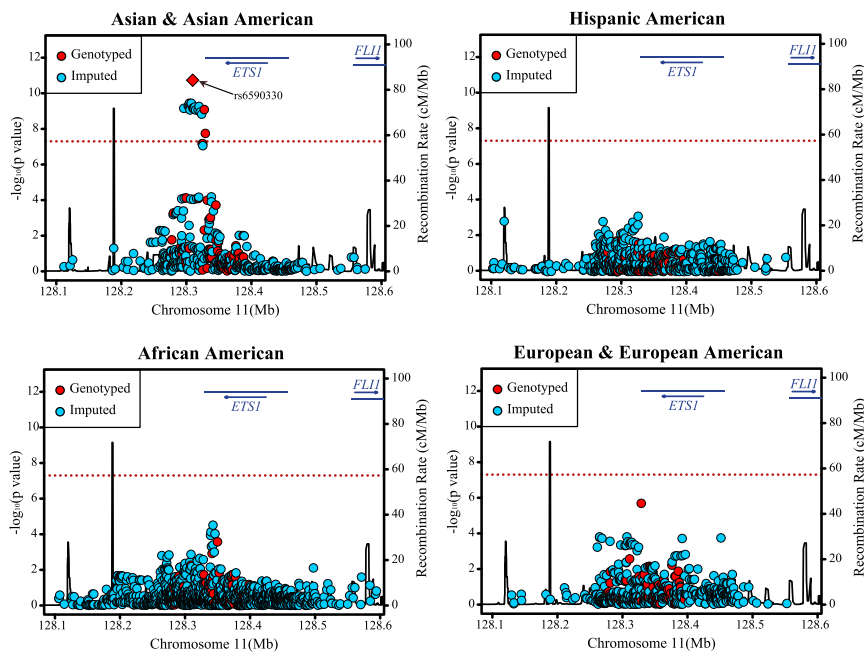
### DNA Affinity Precipitation Assay

We annealed pairs of single-stranded 5'-biotinylated 35-bp oligonucleotides (obtained from IDT) to generate double-stranded probes. Cell lysates were prepared from Epstein-Barr-virus-transformed B cell lines with cell-lysis buffer supplied by the  $\mu$ MACS Factor Finder Kit (Miltenyi Biotec) and the addition of protease inhibitor and phosphatase inhibitors (Pierce Biotechnology). Binding reactions were then performed with biotinylated probes, cell lysate, binding buffer, binding enhancer, protease inhibitor, phosphatase inhibitor, and 0.1  $\mu$ g poly(dI-dC) according to protocols supplied by the  $\mu$ MACS Factor Finder Kit. Eluted probe-bound proteins were identified by Nano liquid chromatography followed by tandem mass spectrometry analysis<sup>30</sup> and immunoblotting with anti-STAT1 or anti-pSTAT1 antibodies. The oligonucleotide sequences are included in Table S2.

### Chromatin Immunoprecipitation qPCR

Crosslinking of protein-chromatin complexes was achieved by incubation of Epstein-Barr-virus-transformed B cells in crosslinking solution (1% formaldehyde, 5 mM HEPES [pH 8.0], 10 mM sodium chloride [NaCl], 0.1 mM EDTA, and 0.05 mM EGTA) and shaking at room temperature for 10 min. Glycine was added to a final concentration of 0.125 M to quench the crosslinking. Cells were washed twice with ice-cold PBS, resuspended in lysis buffer L1 (50 mM HEPES [pH 8.0], 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.25% Triton X-100, and 0.5% NP-40), and incubated for 10 min on ice. Protease and phosphatase inhibitors were added to all buffers. Nuclei were harvested after centrifugation at 5,000 rpm for 10 min, resuspended in lysis buffer L2 (10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 200 mM NaCl, and 0.5 mM EGTA), and incubated at room temperature for 10 min. Nuclei were resuspended in sonication buffer (10 mM Tris [pH 8.0], 1 mM EDTA, and 0.1% SDS) after centrifuging. A S220 focused ultrasonicator (COVARIS) was used to shear genomic DNA (150- to 500-bp fragments) with 10% duty cycle, 175 peak power, and 200 bursts per cycle for 7 min. Sheared chromatin was precleared with 10  $\mu$ l Dynabeads Protein G (Life Technologies) at 4°C for 1 hr. Antibody (anti-STAT1, Santa Cruz Biotechnology) was incubated with 20  $\mu$ l Dynabeads Protein G at room temperature for 1 hr and then washed with PBS once. The antibody-coated beads were incubated with sheared chromatin at 4°C overnight. A volume of 1% sheared chromatin was used as input control. After immunoprecipitation, the beads were washed consecutively with low-salt wash buffer





**Figure 1. *ETS1*-Imputed Genetic Variants Demonstrate Genome-wide Lupus Association in a Cohort of Asian Ancestry**

Each variant is represented as a data point in the context of its genomic location and is colored on the basis of whether it was directly genotyped (red) or only imputed (blue). Genomic position is given with GRCh37 coordinates. rs6590330 was directly genotyped. The SLE association of genotyped and imputed variants in cohorts of Asian and Asian-American (AS) ancestry (12,57 case and 1,258 control subjects), Hispanic-American (HA) ancestry (952 case and 335 control subjects), African-American (AA) ancestry (1,524 case and 1,809 control subjects), and European and European-American (EU) ancestry (3,926 case and 3,490 control subjects) were assessed in a logistic regression with adjustment for admixture estimates. Genome-wide association was defined as  $p < 5 \times 10^{-8}$  and is indicated by a dashed red line in each figure panel.

(0.1% SDS, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM EDTA, 50 mM Tris-HCl [pH 8], and 150 mM NaCl) twice, high-salt wash buffer (as above with 500 mM NaCl) twice, LiCl wash buffer (0.5 M LiCl, 1% NP-40, 0.7% sodium deoxycholate, 1 mM EDTA, and 50 mM Tris-HCl [pH 8]) twice, and 1 mM EDTA and 10 mM Tris-HCl (pH 8) twice. Purified chromatin fragments were eluted from the beads with elution buffer (340 mM NaCl, 1 mM EDTA, and 10 mM Tris-HCl) and 1 mg/ml proteinase K and incubated at 37°C for 1 hr. DNA crosslinks were reversed by incubation of precipitates at 65°C for 5 hr. DNA was purified by the PureLink PCR Micro Kit (Life Technologies) and resuspended in H<sub>2</sub>O. DNA was then analyzed by qPCR with a single set of genotyping primers and differentially tagged fluorescent probes for the risk and non-risk alleles of rs6590330. This qPCR was performed by TaqMan assay on an ABI 7500 PCR system.

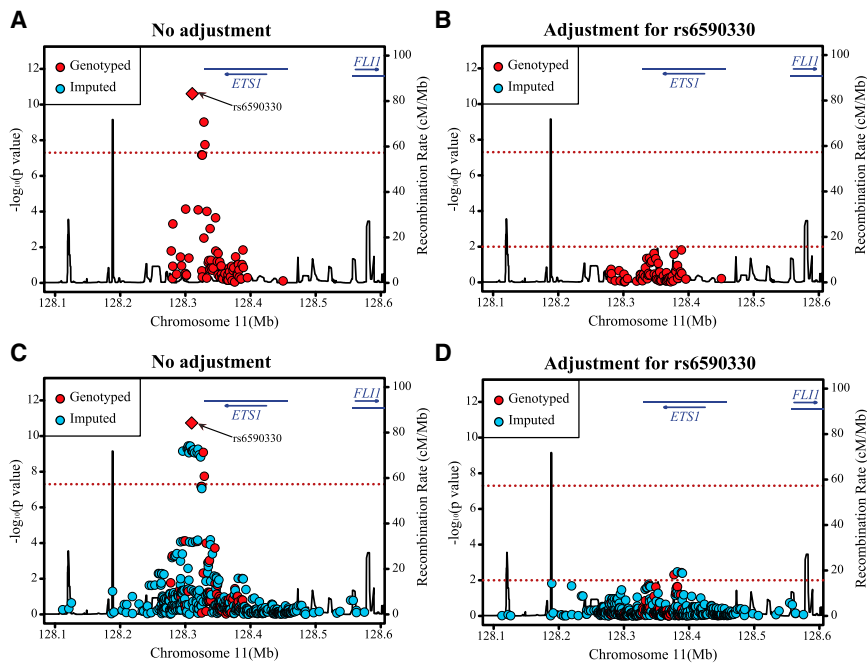
For calculating chromatin immunoprecipitation (ChIP) results as enrichment folds, the amount of immunoprecipitated DNA from the negative control site and the amount of immunoprecipitated DNA from the target site were first normalized against the amount of input DNA. The enrichment folds were then calculated as the amount of immunoprecipitated DNA from the target site divided by the amount of immunoprecipitated DNA from the negative control site. The sequences of primer pairs and probes are included in Table S2.

## Results

We genotyped a total of 69 genetic variants in *ETS1* in 7,659 SLE subjects and 6,892 control subjects (Table S1). Our trans-ancestral cohort included AS, HA, EU, and AA participants. We then imputed against 1000 Genomes data, acquiring 1,333 genetic variants with MAFs > 0.01. A total of 1,402 genetic variants were used for a fine-mapping and model-building study aimed at identifying the causal genetic variants driving the lupus association at this locus.

We tested each genetic variant individually for its association with SLE in each ancestry by using admixture estimates for subjects as covariates (Figure 1). As described in previous GWASs, the AS ancestry was the only one in which we identified variants with a probability (p) reaching genome-wide significance ( $p < 5 \times 10^{-8}$ ) (Figure 1). Apart from the AS ancestry, we also observed some suggestive association in other ancestries (Figure 1). Before performing this association study, we had enough power to detect associations with both the AA and EU ancestries (a prior power of 92% for AA and 99% for EU), whereas we only had 25% power to detect an association with HA ancestry. However, HA ancestry had variants with association at  $p < 10^{-7}$ , which was reduced to  $10^{-4}$  after admixture adjustment (Figure S1). This suggests that the association seen in HA ancestry might be due to a mixed Asian population structure. Therefore, we conclude that genetic variants in *ETS1* are associated with SLE only in the AS ancestry cohort.

After performing a haplotype association analysis in the AS ancestry cohort (Table S3), we found that no haplotypic model of association ( $p_{\text{lowest}} = 7.38 \times 10^{-9}$ ) was more significant than the SNP rs6590330 (the most significant SNV; p value =  $1.80 \times 10^{-11}$ , odds ratio = 1.407 [1.2585–1.573]), suggesting a single-variant model for lupus risk at this locus. We then performed a stepwise logistic regression analysis to identify independent genetic loci in the AS cohort; we started with adjustment for rs6590330. In the dataset containing only directly genotyped variants, no SNP retained a p value  $< 10^{-2}$  after adjustment for rs6590330 (Figure 2). Four SNPs retained p values in the range of 0.001–0.01 after adjustment for rs6590330 with the full dataset containing both genotyped and imputed variants. In summary, the great majority of the lupus



**Figure 2. A Single Genetic Effect Marked by Genotyped SNV rs6590330 Contributes to Lupus Risk in the AS Cohort**  
Genomic position is given with GRCh37 coordinates.

(A) The logistic regression association of genotyped variants in an AS cohort with an adjustment for admixture.  
(B) The logistic association of genotyped variants in an AS cohort with an adjustment for admixture and rs6590330.  
(C) The logistic association of genotyped and imputed variants in an AS cohort with an adjustment for admixture.  
(D) The logistic association of genotyped and imputed variants in an AS cohort with an adjustment for admixture and rs6590330.

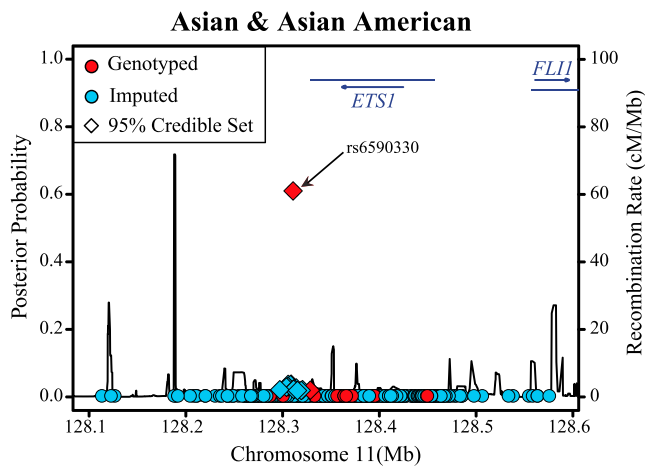
On the basis of our analysis, we hypothesized that the causal variant might reduce *ETS1* expression through differential miRNA binding, differential transcriptional factor binding, and/or changing chromatin

association at this locus was explained by rs6590330. Therefore, the frequentist approach is consistent with a model in which a single genetic effect marked by genotyped SNV rs6590330 contributes to lupus risk in our AS cohort.

To complement our frequentist approach, we used a Bayesian fine-mapping strategy to identify the set of genetic variants that account for 95% of the total posterior probability in the region. In total, this procedure was highly consistent with the frequentist approach and identified 16 genetic variants (3 genotyped and 13 imputed) as the 95% credible set that could be responsible for the *ETS1* association (Figure 3). These same 16 variants were also the most highly associated variants in the frequentist approach. Of the variants evaluated, rs6590330 made a 125-fold greater contribution than any other variant.

Most SNPs in the credible set are in a non-coding region of the genome, so the way they might act is by affecting *ETS1* expression levels. Yang et al. showed that the risk allele of rs1128334 is associated with reduced *ETS1* mRNA expression.<sup>8</sup> Because rs1128334 is an SNP of the credible set, and is in high LD with the other 15 SNPs, we hypothesized that each of the SNPs in the credible set was associated with reduced *ETS1* expression. We confirmed this hypothesis that members of the candidate credible set were associated with *ETS1* expression by using publicly available SNP-mRNA expression data. For example, we found that rs6590330 is a strong expression quantitative trait locus (eQTL) for *ETS1*, given that the risk allele is associated with reduced *ETS1* expression in HapMap subjects from the China-Beijing cohort (Figure S2). Meanwhile, this SNP is also found to be a significant eQTL for nearby gene *TP53AIP1* (MIM 605426), suggesting the potential for multiple genotype-dependent changes in gene expression at this locus (Figure S2).

interaction(s) or state(s). We used datasets and tools available from TargetScan,<sup>31</sup> Cis-BP,<sup>32</sup> Roadmap Epigenomics,<sup>33</sup> ENCODE,<sup>34</sup> and other sources to assess these possibilities for each of the SNVs in the credible set. The non-risk allele of rs1128334 was predicted to be bound by miR-300;<sup>31,35</sup> however, miR-300 is not expressed in cells of hematopoietic origin. The other 15 variants were not located within the *ETS1* transcript and thus were not predicted to disrupt miRNA binding. Altogether, we identified four promising functional variants with experimental evidence suggesting that they might affect transcription factor binding by locating to active chromatin regions. These four variants were identified according to the chromatin-state-segmentation hidden Markov model from the Roadmap Project,<sup>33</sup> chromatin looping to RNA polymerase II,<sup>36</sup> and DNase hypersensitivity clusters (Figure S3). We then experimentally analyzed these four most promising variants for differential transcriptional factor binding between risk and non-risk alleles with electrophoretic mobility shift assays using nuclear lysate from B cells (Figure S4). Of these variants, only rs6590330 exhibited differential binding of the risk and non-risk oligonucleotides to nuclear factors (Figure S4). For this variant, we identified the specific differentially bound protein by using a DNA affinity precipitation assay (DAPA) followed by mass spectrometry. The mass spectrometry results indicated that STAT1 binds to the risk allele but not the non-risk allele for rs6590330 (Table S4). We confirmed this finding by DAPA followed by immunoblotting for phosphorylated STAT1 and total STAT1 (Figure 4). Critically, STAT1 ChIP-qPCR analysis also confirmed that the risk allele of rs6590330 has more STAT1 binding enrichment than the non-risk allele in B cells heterozygous for rs6590330 (Figure 4).

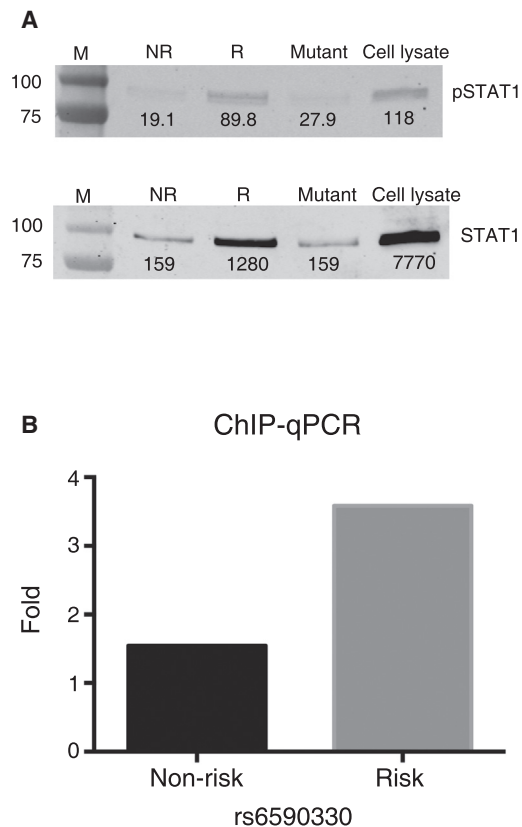


Variant name	Position	Bayes Factor	Posterior Probability
rs6590330	128,311,059	$2.73 \times 10^8$	0.61
rs72106982	128,308,991	$1.55 \times 10^7$	0.03
rs77989778	128,305,916	$1.52 \times 10^7$	0.03
rs12293665	128,308,361	$1.48 \times 10^7$	0.03
rs9736922	128,305,872	$1.23 \times 10^7$	0.03
rs9736939	128,305,871	$1.23 \times 10^7$	0.03
rs12576753	128,304,141	$1.15 \times 10^7$	0.03
rs11501246	128,314,424	$1.04 \times 10^7$	0.02
rs11221307	128,311,770	$9.06 \times 10^6$	0.02
rs61909063	128,312,806	$9.06 \times 10^6$	0.02
rs11602895	128,311,569	$8.65 \times 10^6$	0.02
rs114534998	128,320,877	$8.23 \times 10^6$	0.02
rs12576990	128,297,318	$8.12 \times 10^6$	0.02
rs1128334	128,328,959	$7.55 \times 10^6$	0.02
rs12574073	128,319,478	$7.20 \times 10^6$	0.02
rs140066073	128,315,613	$7.07 \times 10^6$	0.02

**Figure 3. Bayesian Association Plot Showing the Signal Strength in *ETS1* as the Posterior Probability of Each SNV**

Genomic position is given with GRCh37 coordinates; AS data are shown. SNVs are colored according to their origin: genotyped variants are in red, and imputed variants are in blue. Variants in the 95% credible set are marked by diamonds. Variants with larger posterior probabilities ( $>0.01$ ) represent those most likely to be causal.

rs6590330 is located 11 bases away from a putative STAT1 binding site predicted by a binding model identified in Factorbook<sup>37</sup> (Figure S5). When this STAT binding site was disrupted in the oligonucleotide used for the DAPA analysis containing the risk variant, the binding of STAT1 was also disrupted (Figure 4). STAT1 can be activated by type I interferons, including interferon alpha. In the context of data supporting chromatin looping of DNA surrounding rs650330 to the *ETS1* transcription start site, it is possible that this variant enhances activated STAT1 binding to the putative STAT1 binding site near rs6590330 and thus results in the repression of *ETS1* transcription initiated by RNA polymerase II. Taken together, these results strongly support a model in which the variant rs6590330 is highly associated with decreased expression of *ETS1* and increased lupus risk through pSTAT1 binding of a DNA sequence located next to the risk allele in B cells.



**Figure 4. The Lupus Risk Allele of rs6590330 Increases STAT1 Binding**

(A) STAT1 and pSTAT1 exhibit higher binding to oligonucleotides containing the rs6590330 risk allele than to oligonucleotides containing the non-risk allele. Biotin-labeled oligonucleotides were incubated with nuclear extract from Epstein-Barr-virus-transformed B cells. Proteins bound to the oligonucleotides were captured with the  $\mu$ MACs Factor Finder Kit. Proteins were then separated by SDS-PAGE and detected with anti-pSTAT1 (top) and anti-STAT1 (bottom). Abbreviations are as follows: M, marker; NR, oligonucleotide containing the non-risk allele of rs6590330; R, oligonucleotide containing the risk allele of rs6590330 (see Figure S5); mutant, oligonucleotide containing a disrupted putative STAT binding site downstream of rs6590330; and cell lysate, nuclear extract from B cells. The relative intensities of the bands are given above each band. Results are representative of four experiments; although all experiments demonstrated increased STAT1 binding to the probes with the risk allele, in two of four experiments, no STAT1 or pSTAT1 was detected in the immunoprecipitate from the non-risk oligonucleotide, whereas both were detected in the immunoprecipitate from the risk oligonucleotide. (B) rs6590330-heterozygous Epstein-Barr-virus-transformed B cells were used for ChIP-qPCR assessment of the differential binding of STAT1 to the risk and non-risk alleles. Crosslinked and sonicated chromatin was immunoprecipitated with an anti-STAT1 antibody. Site-specific primers and probes specific to the rs6590330 risk and non-risk alleles were used for determining STAT1 binding to immunoprecipitated DNA.

## Discussion

The genotyped and imputed data from 14,551 subjects facilitated fine mapping of *ETS1* and its association with SLE risk. A model consisting of a single genetic effect was identified by stepwise logistic regression analyses.

Importantly, a set of 16 variants that are likely to be causal were identified through frequentist and Bayesian analyses. Our data are consistent with a mechanistic model in which the risk allele of rs6590330 at *ETS1* contributes to increased SLE risk by facilitating binding of pSTAT1 to a nearby putative STAT binding site and subsequent repression of the expression of *ETS1*. Our fine-mapping results are corroborated by a recent study that used a different analytical method to estimate the probability that each variant within *ETS1* is a causal variant; this study concluded that rs6590330 is among the most likely causal variants proposed for the *ETS1* association with SLE,<sup>38</sup> further supporting the results of our genetic analysis.

The biological mechanism we propose herein is specific to the association of *ETS1* variants in the AS cohort. We performed the functional validation in B cells on the basis of the evidence that B cells play a critical role in the etiology of SLE. Previous studies have demonstrated that B cells are significantly enriched with the expression of genes near lupus risk loci.<sup>39</sup> Meanwhile, B cells from lupus-affected individuals produce autoantibodies, are hyperactivated, and have an exaggerated response to Toll-like receptor ligands and immune complexes.<sup>40</sup> *ETS1*-hypomorphic mice (producing *ETS1* lacking the Pointed domain [*Ets1<sup>P/P</sup>*]) develop a strikingly similar B cell phenotype.<sup>12,41</sup> This *ETS1* deficiency has been shown to drive intrinsic terminal differentiation of B cells into IgM-secreting plasma cells in a B-cell-intrinsic fashion.<sup>42</sup> Although we limited our functional analysis to B cells, it remains possible that the same genetic risk mechanism might also operate in other cell types, such as T cells.<sup>42,43</sup>

In all four ancestries, rs6590330 is polymorphic (MAF<sub>AS</sub> = 45%, MAF<sub>AA</sub> = 31%, MAF<sub>HA</sub> = 28%, and MAF<sub>EU</sub> = 26%). The *ETS1* association has only been observed in AS ancestry cohorts in our study and others before us. Perhaps *ETS1* variants do not increase SLE risk in AA and EU populations, a possibility consistent with epistasis or environmental factors (gene-environment interactions).<sup>44</sup> Asian-specific variants have been identified in previous studies,<sup>1,10</sup> and it is not surprising that ancestry-specific genetic factors affect the risk of SLE. Subjects of Asians ancestry have a higher SLE incidence than do those of European ancestry, in addition to an ancestry-specific distribution of clinical manifestations.<sup>45</sup>

We found that the risk allele of rs6590330 results in preferential binding of pSTAT1 to a nearby putative STAT binding site and is associated with reduced expression of *ETS1*. STAT1 plays a complex role in regulating gene expression and is capable of acting as either an activator or a repressor, depending on the cellular context.<sup>46</sup> This is an intriguing candidate causal mechanism in which disease risk is mediated through the disruption of transcription factor binding to a nearby site that does not contain the risk variant. Because the risk allele of rs6590330 is associated with reduced *ETS1* expression (Yang et al.<sup>8</sup> and Figure S2), it is possible that reduced *ETS1* expression associated with the risk allele of rs6590330 might skew B cell differentia-

tion to IgM plasma cells and thus subsequently contribute to SLE pathogenesis. Meanwhile, even though we present rs6590330 as a candidate causal variant responsible for disease risk, more experiments are necessary to establish causality and define the genotype-dependent immunological mechanisms driving lupus at this locus.

In conclusion, we conducted a large trans-ancestral fine-mapping study of *ETS1* to identify genetic variants that increase lupus risk. After determining a model of single genetic effect, we used frequentist and Bayesian association methods to identify a set of 16 variants that are most likely causal. Of these, we identified an allele-specific function for rs6590330. Altogether, we propose a model in which the risk allele of rs6590330 increases SLE risk through increased binding of pSTAT1 and depressed expression of *ETS1*.

### Supplemental Data

Supplemental Data include five figures and four tables and can be found with this article online at <http://dx.doi.org/10.1016/j.ajhg.2015.03.002>.

### Acknowledgments

We are grateful for support from the NIH (AI024717, AI063274, AI082714, AI083194, AR043274, AR043727, AR043814, AR051545, AR053483, AR056360, AR057172, AR058959, AR060366, AR063124, AR065626, AR62277, GM103456, GM104938, HG006828, K24AI078004, K24AR02318, MD007909, P01AR49084, P30AR053483, P30AR055385, P30GM103510, P60AR053308, P60AR062755, P60AR064464, R01AR44804, R21AI070304, S10RR027015, TR000077, U01AI101934, U01HG006828, U19AI082714, U54GM104938, UL1RR029882, UL1TR000004, and UL1TR000150). Support for this project was also provided by the United States Departments of Veteran Affairs and Defense (PR094002), a Kirkland Scholar Award, the National Basic Research Program of China (973 program, 2014CB541901), the National Natural Science Foundation of China (81230072 and 81421001), the State Key Laboratory of Oncogenes and Related Genes (grant 91-14-05), the Key Research Program of the Chinese Academy of Sciences (KJZD-EW-L01-3), the Program of the Shanghai Commission of Science and Technology (12JC1406000 and 12431900703), the Instituto de Salud Carlos III (partly financed by Fonds Européen de Développement Régional funds from the European Union [02558]), the Proyecto de Excelencia of the Junta de Andalucía (CTS2548), the Arthritis Foundation, the Alliance for Lupus Research, and the Korea Healthcare Technology R&D Project of the Ministry for Health and Welfare, Republic of Korea (HI13C2124).

Received: December 18, 2014

Accepted: March 5, 2015

Published: April 9, 2015

### Web Resources

The URLs for data presented herein are as follows:

ADMIXMAP, <http://homepages.ed.ac.uk/pmckeigu/admixmap/>  
CIS-BP, <http://cisbp.cabr.utoronto.ca/>



EIGENSTRAT, <http://www.hsph.harvard.edu/alkes-price/software/Factorbook>, <http://www.factorbook.org/mediawiki/index.php/STAT1>  
Haploview, <http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview>  
OMIM, <http://www.omim.org/>  
PLINK, <http://pngu.mgh.harvard.edu/~purcell/plink/>  
SNPTEST, [https://mathgen.stats.ox.ac.uk/genetics\\_software/snpctest/snpctest.html](https://mathgen.stats.ox.ac.uk/genetics_software/snpctest/snpctest.html)  
TargetScan, <http://www.targetscan.org/>  
UCSC Genome Browser, <http://genome.ucsc.edu/>

## References

- Cui, Y., Sheng, Y., and Zhang, X. (2013). Genetic susceptibility to SLE: recent progress from GWAS. *J. Autoimmun.* *41*, 25–33.
- Lawrence, R.C., Helmick, C.G., Arnett, F.C., Deyo, R.A., Felson, D.T., Giannini, E.H., Heyse, S.P., Hirsch, R., Hochberg, M.C., Hunder, G.G., et al. (1998). Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis Rheum.* *41*, 778–799.
- Pons-Estel, G.J., Alarcón, G.S., Scofield, L., Reinlib, L., and Cooper, G.S. (2010). Understanding the epidemiology and progression of systemic lupus erythematosus. *Semin. Arthritis Rheum.* *39*, 257–268.
- Chakravarty, E.F., Bush, T.M., Manzi, S., Clarke, A.E., and Ward, M.M. (2007). Prevalence of adult systemic lupus erythematosus in California and Pennsylvania in 2000: estimates obtained using hospitalization data. *Arthritis Rheum.* *56*, 2092–2094.
- Alarcón-Segovia, D., Alarcón-Riquelme, M.E., Cardiel, M.H., Caeiro, F., Massardo, L., Villa, A.R., and Pons-Estel, B.A.; Grupo Latinoamericano de Estudio del Lupus Eritematoso (GLADEL) (2005). Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort. *Arthritis Rheum.* *52*, 1138–1147.
- Han, J.-W., Zheng, H.-F., Cui, Y., Sun, L.-D., Ye, D.-Q., Hu, Z., Xu, J.-H., Cai, Z.-M., Huang, W., Zhao, G.-P., et al. (2009). Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat. Genet.* *41*, 1234–1237.
- He, C.F., Liu, Y.S., Cheng, Y.L., Gao, J.P., Pan, T.M., Han, J.W., Quan, C., Sun, L.D., Zheng, H.F., Zuo, X.B., et al. (2010). TNIP1, SLC15A4, ETS1, RasGRP3 and IKZF1 are associated with clinical features of systemic lupus erythematosus in a Chinese Han population. *Lupus* *19*, 1181–1186.
- Yang, W., Shen, N., Ye, D.-Q., Liu, Q., Zhang, Y., Qian, X.-X., Hirankarn, N., Ying, D., Pan, H.-F., Mok, C.C., et al.; Asian Lupus Genetics Consortium (2010). Genome-wide association study in Asian populations identifies variants in ETS1 and WDFY4 associated with systemic lupus erythematosus. *PLoS Genet.* *6*, e1000841.
- Zhong, H., Li, X.L., Li, M., Hao, L.X., Chen, R.W., Xiang, K., Qi, X.B., Ma, R.Z., and Su, B. (2011). Replicated associations of TNFAIP3, TNIP1 and ETS1 with systemic lupus erythematosus in a southwestern Chinese population. *Arthritis Res. Ther.* *13*, R186.
- Lee, H.-S., Kim, T., Bang, S.Y., Na, Y.J., Kim, I., Kim, K., Kim, J.-H., Chung, Y.-J., Shin, H.D., Kang, Y.M., et al. (2014). Ethnic specificity of lupus-associated loci identified in a genome-wide association study in Korean women. *Ann. Rheum. Dis.* *73*, 1240–1245.
- Leng, R.-X., Pan, H.-F., Chen, G.-M., Feng, C.-C., Fan, Y.-G., Ye, D.-Q., and Li, X.-P. (2011). The dual nature of Ets-1: focus to the pathogenesis of systemic lupus erythematosus. *Autoimmun. Rev.* *10*, 439–443.
- Wang, D., John, S.A., Clements, J.L., Percy, D.H., Barton, K.P., and Garrett-Sinha, L.A. (2005). Ets-1 deficiency leads to altered B cell differentiation, hyperresponsiveness to TLR9 and autoimmune disease. *Int. Immunol.* *17*, 1179–1191.
- Maller, J.B., McVean, G., Byrnes, J., Vukcevic, D., Palin, K., Su, Z., Howson, J.M.M., Auton, A., Myers, S., Morris, A., et al.; Wellcome Trust Case Control Consortium (2012). Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat. Genet.* *44*, 1294–1301.
- Kottyan, L.C., Zoller, E.E., Bene, J., Lu, X., Kelly, J.A., Rupert, A.M., Lessard, C.J., Vaughn, S.E., Marion, M., Weirauch, M.T., et al.; UK Primary Sjögren's Syndrome Registry; UK Primary Sjögren's Syndrome Registry (2015). The IRF5-TNPO3 association with systemic lupus erythematosus has two components that other autoimmune disorders variably share. *Hum. Mol. Genet.* *24*, 582–596.
- Rasmussen, A., Sevier, S., Kelly, J.A., Glenn, S.B., Aberle, T., Cooney, C.M., Grether, A., James, E., Ning, J., Tesiram, J., et al. (2011). The lupus family registry and repository. *Rheumatology (Oxford)* *50*, 47–59.
- Lessard, C.J., Adrianto, I., Kelly, J.A., Kaufman, K.M., Grun-dahl, K.M., Adler, A., Williams, A.H., Gallant, C.J., Anaya, J.M., Bae, S.C., et al.; Marta E. Alarcón-Riquelme on behalf of the BIOLUPUS and GENLES Networks (2011). Identification of a systemic lupus erythematosus susceptibility locus at 11p13 between PDHX and CD44 in a multiethnic study. *Am. J. Hum. Genet.* *88*, 83–91.
- Hochberg, M.C. (1997). Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* *40*, 1725.
- McKeigue, P.M., Carpenter, J.R., Parra, E.J., and Shriver, M.D. (2000). Estimation of admixture and detection of linkage in admixed populations by a Bayesian approach: application to African-American populations. *Ann. Hum. Genet.* *64*, 171–186.
- Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* *38*, 904–909.
- Hoggart, C.J., Parra, E.J., Shriver, M.D., Bonilla, C., Kittles, R.A., Clayton, D.G., and McKeigue, P.M. (2003). Control of confounding of genetic associations in stratified populations. *Am. J. Hum. Genet.* *72*, 1492–1504.
- Hoggart, C.J., Shriver, M.D., Kittles, R.A., Clayton, D.G., and McKeigue, P.M. (2004). Design and analysis of admixture mapping studies. *Am. J. Hum. Genet.* *74*, 965–978.
- Smith, M.W., Patterson, N., Lautenberger, J.A., Truelove, A.L., McDonald, G.J., Waliszewska, A., Kessing, B.D., Malasky, M.J., Scafe, C., Le, E., et al. (2004). A high-density admixture map for disease gene discovery in african americans. *Am. J. Hum. Genet.* *74*, 1001–1013.
- Halder, I., Shriver, M., Thomas, M., Fernandez, J.R., and Frudakis, T. (2008). A panel of ancestry informative markers for estimating individual biogeographical ancestry and admixture from four continents: utility and applications. *Hum. Mutat.* *29*, 648–658.

24. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* *81*, 559–575.
25. Marchini, J., Howie, B., Myers, S., McVean, G., and Donnelly, P. (2007). A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* *39*, 906–913.
26. Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* *21*, 263–265.
27. Barrett, J.C. (2009). Haploview: Visualization and analysis of SNP genotype data. *Cold Spring Harb Protoc* *2009*, ip71.
28. Stephens, M., and Balding, D.J. (2009). Bayesian statistical methods for genetic association studies. *Nat. Rev. Genet.* *10*, 681–690.
29. Altshuler, D.M., Gibbs, R.A., Peltonen, L., Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., Peltonen, L., et al.; International HapMap 3 Consortium (2010). Integrating common and rare genetic variation in diverse human populations. *Nature* *467*, 52–58.
30. Wijeratne, A.B., Manning, J.R., Schultz, J.J., and Greis, K.D. (2013). Quantitative phosphoproteomics using acetone-based peptide labeling: method evaluation and application to a cardiac ischemia/reperfusion model. *J. Proteome Res.* *12*, 4268–4279.
31. Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* *120*, 15–20.
32. Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* *158*, 1431–1443.
33. Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R., et al. (2010). The NIH Roadmap Epigenomics Mapping Consortium. *Nat. Biotechnol.* *28*, 1045–1048.
34. Rosenbloom, K.R., Sloan, C.A., Malladi, V.S., Dreszer, T.R., Learned, K., Kirkup, V.M., Wong, M.C., Maddren, M., Fang, R., Heitner, S.G., et al. (2013). ENCODE data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res.* *41*, D56–D63.
35. Betel, D., Wilson, M., Gabow, A., Marks, D.S., and Sander, C. (2008). The microRNA.org resource: targets and expression. *Nucleic Acids Res.* *36*, D149–D153.
36. Li, G., Fullwood, M.J., Xu, H., Mulawadi, F.H., Velkov, S., Vega, V., Ariyaratne, P.N., Mohamed, Y.B., Ooi, H.S., Tennakoon, C., et al. (2010). ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing. *Genome Biol.* *11*, R22.
37. Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X., Kundaje, A., Cheng, Y., et al. (2012). Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res.* *22*, 1798–1812.
38. Farh, K.K., Marson, A., Zhu, J., Kleinewietfeld, M., Housley, W.J., Beik, S., Shores, N., Whitton, H., Ryan, R.J., Shishkin, A.A., et al. (2015). Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* *518*, 337–343.
39. Hu, X., Kim, H., Stahl, E., Plenge, R., Daly, M., and Raychaudhuri, S. (2011). Integrating autoimmune risk loci with gene-expression data identifies specific pathogenic immune cell subsets. *Am. J. Hum. Genet.* *89*, 496–506.
40. Dörner, T., Giesecke, C., and Lipsky, P.E. (2011). Mechanisms of B cell autoimmunity in SLE. *Arthritis Res. Ther.* *13*, 243.
41. Wang, H., Xu, J., Ji, X., Yang, X., Sun, K., Liu, X., and Shen, Y. (2005). The abnormal apoptosis of T cell subsets and possible involvement of IL-10 in systemic lupus erythematosus. *Cell. Immunol.* *235*, 117–121.
42. John, S.A., Clements, J.L., Russell, L.M., and Garrett-Sinha, L.A. (2008). Ets-1 regulates plasma cell differentiation by interfering with the activity of the transcription factor Blimp-1. *J. Biol. Chem.* *283*, 951–962.
43. Moisan, J., Grenningloh, R., Bettelli, E., Oukka, M., and Ho, I.C. (2007). Ets-1 is a negative regulator of Th17 differentiation. *J. Exp. Med.* *204*, 2825–2835.
44. Hirschhorn, J.N., Lohmueller, K., Byrne, E., and Hirschhorn, K. (2002). A comprehensive review of genetic association studies. *Genet. Med.* *4*, 45–61.
45. Jakes, R.W., Bae, S.C., Louthrenoo, W., Mok, C.C., Navarra, S.V., and Kwon, N. (2012). Systematic review of the epidemiology of systemic lupus erythematosus in the Asia-Pacific region: prevalence, incidence, clinical features, and mortality. *Arthritis Care Res. (Hoboken)* *64*, 159–168.
46. Ramana, C.V., Chatterjee-Kishore, M., Nguyen, H., and Stark, G.R. (2000). Complex roles of Stat1 in regulating gene expression. *Oncogene* *19*, 2619–2627.