Published in final edited form as:
Arthritis Rheum. 2009 April ; 60(4): 1085–1095. doi:10.1002/art.24387.

High density genotyping of STAT4 gene reveals multiple haplotypic associations with Systemic Lupus Erythematosus in different racial groups

Bahram Namjou[1], **Andrea L. Sestak**[1], **Don L. Armstrong**[2], **Raphael Zidovetzki**[2], **Jennifer A. Kelly**[1], **Noam Jacob**[3], **Voicu Ciobanu**[3], **Kenneth M. Kaufman**[1,4,5], **Joshua O. Ojwang**[1], **Julie Ziegler**[20], **Francesco Quismorio**[3], **Andreas Reiff**[3,6], **Barry L. Myones**[7], **Joel M. Guthridge**[1], **Swapan K. Nath**[1], **Gail R. Bruner**[1], **Ruth Mehrian-Shai**[3], **Earl Silverman**[8], **Marisa Klein-Gitelman**[9], **Deborah McCurdy**[10], **Linda Wagner-Weiner**[11], **James J. Nocton**[12], **Chaim Putterman**[13], **Sang-Cheol Bae**[14], **Yun Jung Kim**[14], **Michelle Petri**[15], **John D. Reveille**[16], **Timothy J. Vyse**[17], **Gary S. Gilkeson**[18], **Diane L. Kamen**[18], **Marta E. Alarcón-Riquelme**[19], **Patrick M. Gaffney**[1], **Kathy L Moser**[1], **Joan T. Merrill**[1], **R. Hal Scofield**[1], **Judith A. James**[1], **Carl D. Langefeld**[20], **John B. Harley**[1,4,5], and **Chaim O. Jacob**[3,*]

[1] Oklahoma Medical Research Foundation, 825 NE 13th Street, Oklahoma City, Oklahoma 73104, USA

[2] Department of Cell Biology and Neuroscience, University of California, Riverside CA 92521, USA

[3] The Lupus genetic Group, University of Southern California, 2011 Zonal Avenue, Los Angeles, California 90033, USA

[4] US Department of Veterans Affairs Medical Center, 921 NE 13th Street, Oklahoma City, Oklahoma 73104, USA

[5] University of Oklahoma Health Sciences Center, 1100 N Lindsey, Oklahoma City, Oklahoma 73104, USA

[6] Children's Hospital of Los Angeles, Los Angeles, CA 90027, USA

[7] Texas Children's Hospital, Baylor College of Medicine, Houston, TX 77030, USA

[8] Hospital for Sick Children, Toronto, OT Canada

[9] Children's Memorial Hospital and Northwestern University, Chicago IL 60614, USA

[10] Department of Pediatrics, UCLA, Los Angeles CA 90095, USA

[11] LaRabida Hospital and University of Chicago, Chicago, IL 60649, USA

[12] Medical College of Wisconsin, Milwaukee, WI 53226, USA

[13] Division of Rheumatology, Albert Einstein College of Medicine, Bronx, NY 10461

[14] Division of Rheumatology, Department of Internal Medicine and the Hospital for Rheumatic Diseases, Hanyang University, Seoul, Republic of Korea

[15] Johns Hopkins University School of Medicine, Division of Rheumatology, Baltimore, MD 21205

[16] University of Texas-Houston Health Science Center, Houston, TX 77030

[17] Imperial College London, Hammersmith Hospital, Du Cane Road, London W12 0NN UK

*Corresponding author: Chaim O. Jacob MD, PhD, University of Southern California (USC) School of Medicine, Department of Medicine, 2011 Zonal Ave, HMT #705, Los Angeles CA 90033, Jacob@usc.edu, Tel: 323 442-1822, Fax: 323 442-2874.

[18] Medical University of South Carolina, Charleston, SC, USA

[19] Department of Genetics and Pathology, Uppsala University, Uppsala, Sweden

[20] Wake Forest University Health Sciences, Medical Center Blvd., Winston-Salem, North Carolina 27157, USA

Abstract

Objective—Systemic lupus erythematosus (SLE) is the prototypic systemic autoimmune disorder with complex etiology and a strong genetic component. Recently, gene products involved in the interferon pathway have been under intense investigation in SLE pathogenesis. STAT1 and STAT4 are transcription factors that play key roles in the interferon and Th1 signaling pathways, making them attractive candidates for SLE susceptibility.

Methods—Fifty-six single-nucleotide polymorphisms (SNPs) across STAT1 and STAT4 genes on chromosome 2 were genotyped using Illumina platform as a part of extensive association study in a large collection of 9923 lupus cases and controls from different racial groups. DNA from patients and controls was obtained from peripheral blood. Principal component analyses and population based case-control association analyses were performed and the p values, FDR q values and Odds ratios with 95% confidence intervals (95% CIs) were calculated.

Results—We observed strong genetic associations with SLE and multiple SNPs located within the STAT4 gene in different ethnicities (Fisher combined $p= 7.02\times10^{-25}$). In addition to strong confirmation of the association in the 3rd intronic region of this gene reported previously, we identified additional haplotypic association across STAT4 gene and in particular a common risk haplotype that is found in multiple racial groups. In contrast, only a relatively weak suggestive association was observed with STAT1, probably due to the proximity to STAT4.

Conclusion—Our findings indicate that the STAT4 gene is likely to be a crucial component in SLE pathogenesis among multiple racial groups. The functional effects of this association, when revealed, might improve our understanding of the disease and provide new therapeutic targets.

Systemic lupus erythematosus (SLE) is a complex multi-organ autoimmune disorder with a strong genetic component characterized by breakdown of self-tolerance, which results in a wide range of immunological abnormalities including pathogenic immune complex formation, T and B lymphocyte dysregulation, and defective clearance of apoptotic materials.

The roles of various cytokines and their signaling molecules have gained importance in understanding the pathogenesis of SLE. Among these, the interferons, both type I and type II, have received particular attention. Accordingly, peripheral blood mononuclear cells from SLE patients show a pattern of upregulated IFN-induced genes (1–3) and this “interferon signature” correlates with disease severity markers (4). IFN-α treatment of individuals with viral infections or malignancies might result in SLE-like manifestations (5,6). Increase serum IFN-activity was found to be a heritable trait in families with SLE (7). Furthermore, a component of the interferon pathway, IRF5 has been established as an SLE susceptibility gene (8–10). Similarly, the participation of IFN-γ (the only type II IFN), has been inferred in human SLE, and confirmed in lupus mice (11–13).

Signal transmission from the interferons involves STAT1 and STAT4, which are members of the signal transducer and activators of transcription (STAT) family of transcriptional factors. These proteins are involved in essential cellular events such as differentiation, proliferation, and apoptosis following cytokine and growth factor signaling (14). T

By binding to their receptors, interferons and other cytokines trigger Jak kinases to phosphorylate and activate STAT proteins (15). Before activation, STAT proteins are cytosolic

and activation by tyrosine phosphorylation results in their homo- and hetero-dimerization through interactions involving their SH2 domains; STAT dimers then translocate to the nucleus, where they either directly bind to DNA or act together with other DNA-binding proteins in multiprotein transcription complexes to direct transcription of a large variety of gene products (14,15).

The human STAT genes have been identified in three chromosomal clusters: STAT1 and STAT4 on human chromosome 2 (q12-33), STAT2 and STAT6 on chromosome 12 (q13-14) and STAT3, STAT5a, and 5b on chromosome 17 (q11.2-22) (16).

STAT1 is activated both by IFNα/β and by IFN-γ signaling (17), which plays an important role in the activation of macrophages and in the defense response to pathogenic agents (18,19). STAT1 targets genes that can promote inflammation and induce apoptosis (17).

STAT4, identified through its homology to STAT1, was found to lie adjacent to the STAT1 gene at 2q32.2-2q32.3, containing 24 exons and spanning 122 Kb. STAT4 is activated by several cytokines including IL-12, IL-23 and IFNα, and stimulates the transcription of specific genes including IFNγ (20). Previous genome scans in SLE have revealed linkage to the 2q33 region (21,22). This fact, together with the extensive involvement of type I and type II IFNs in the pathogenesis of SLE, made the cluster of STAT1 and STAT4 on chromosome 2q an obvious candidate region for genetic predisposition to this autoimmune disease. Recently, Remmers *et al* (23) showed genetic association between STAT4 and RA and also association of one SNP (rs7574865) with SLE in Europeans. Furthermore, two recently published genome-wide association studies of SLE in European ancestry populations confirm the association with the STAT4 gene (24,25)

In this report, we describe the results of a fine mapping study in which we evaluated 56 single nucleotide polymorphisms (SNPs) spanning the STAT1 and STAT4 genes on chromosome 2 in a large collection of 9923 lupus cases and controls from different racial groups. This study is the largest study of these genes in SLE and the first to investigate the associations in populations with higher prevalence including African Americans and Hispanics. Our results confirm and significantly extend the previous association in multiple racial groups.

Materials and Methods

Recruitment and Biological Sample Collection

The present study included 9923 participants (4771 SLE cases and 5,152 controls) enrolled in the Lupus Genetics Studies at OMRF as described (26), in the Lupus Genetic Study Group at USC as described (27), in the PROFILE Study Group at UAB (28), and from additional collaborators to the studies. The demographics and numbers of samples are provided in Table 1. Among SLE cases, 769 independent cases were defined as childhood-onset according to the criterion that the diagnosis of SLE was made before the age of 13 by at least one pediatric rheumatologist participating in the study. All protocols were approved by the Institutional Review Boards at each respective institution. All patients met the revised 1997 ACR criteria for the classification of SLE (29). Ethnicity was self-reported and verified by parental and grandparental ethnicity, when known. Blood samples were collected from each participant, and genomic DNA was isolated and stored using standard methods.

Genotyping

Genotyping was performed using Illumina iSelect™ Infinium II Assays on the BeadStation™ 500GX system (Illumina, San Diego, CA) at the Lupus Genetics Studies unit of the Oklahoma Medical Research Foundation and at the University of Texas Southwestern DNA Microarray Core facility. Genotype data were only used from samples with a call rate greater than 90% of

the SNPs screened (98.05% of the samples). The average call rate for all samples was 97.18%. For analysis, only genotype data from SNPs with a call frequency greater than 90% in the samples tested and an Illumina GenTrain score greater than 0.7 were used. GenTrain scores measure the reliability of SNP detection based on the distribution of genotypic classes. In order to minimize sample misidentification, data from 91 SNPs that had been previously genotyped on 42.12% of the samples were used to verify sample identity. In addition, at least one sample previously genotyped was randomly placed on each Illumina Infinium BeadChip and used to track samples throughout the genotyping process.

Statistical Analyses

Testing for association was completed using the freely available programs SNPGWA (http://www.phs.wfubmc.edu/web/public_bios/sec_gene/downloads.cfm) and PLINK (30). For each SNP, missing data proportions for cases and controls, minor allele frequency and exact tests for departures from Hardy-Weinberg expectations were calculated. In addition to allelic test of association, the additive genetic model was used as the primary hypothesis of statistical inference. If the lack-of-fit (LOF) test for the additive model was significant (LOF $p<0.05$), then the minimum p-value from the dominant, additive or recessive models is reported. For recessive models, at least 30 individuals homozygous for the minor allele were required. Haploview version 4.0 (31) was used to estimate the linkage disequilibrium (LD) between markers and haplotype structures in different ethnicities. The deviation of the observed frequency of a haplotype from the expected is a quantity called the linkage disequilibrium and is commonly denoted by a capital D. D has the disadvantage of depending on the frequency of the alleles. The so called D′ is a common normalized measure of D by dividing it with the theoretical maximum for the observed allele frequencies.

Conditional haplotype analyses were conducted using WHAP program version 2.09 and conditional logistic regression for clinical or serological criteria were conducted using PLINK. African-American and Gullah populations (a relatively more homogenous group of African-American who live in the Low Country of South Carolina) have been analyzed separately. Combined p values were calculated from the per-ethnicity p values using the Fisher method. Q values were calculated using the q value package (available from http://cran.r-project.org) which implements the q value correction of False Discovery Rate (FDR) (32). Q values correspond to the proportion of false positives among the results. Thus, Q values less than 0.05 signify less than 5% of false positives and is taken as a measure of significance.

Stratification Analyses

To account for potential confounding substructure or admixture in these samples, principal component analyses (PCA) were performed (33) using all SNPs (numbering 20,506 genotyped on these subjects as part of a large effort to determine the genetic susceptibility in SLE) except those within the HLA region and known associations from the published genome scan (24). Four principal components were identified that explained a total of ~60% of the observed genetic variation. The PCA scores were used to identify individual that were genetically distant from the other samples and prone to introducing admixture bias. A total of 252 controls and 165 cases were so identified and removed from further analysis (European American: 124 controls and 89 cases; African American: 88 controls and 38 cases; Hispanic American 35 controls and 30 cases; Korean: 1 control; Gullah: 4 controls and 8 cases). After removing these genetic outliers, duplicates and relative samples, 4374 independent SLE cases and 4860 controls remained for analysis. All these subjects were also independent from SLEGEN study (24). We then performed genomic control analysis to calculate the inflation factor λ (Lambda) using all SNPs minus HLA region and previously identified genes (18,446 SNPs, 92% of original SNPs), which produced a $\lambda=1.13$ in European samples, $\lambda=1.03$ in Hispanics, $\lambda=1.08$

in African-Americans, λ=1.04 is Koreans and λ=1.02 in Gullah. Inflation factor is a measure that quantifies the degree to which population stratification increases the χ^2 test statistic.

Only the Hispanic sample required PCA as covariates in the logistic regression model to remove the final source of confounding via admixture to obtain the above inflation factor.

Results

Association of SLE with STAT4 SNPs

To determine if STAT variants associate with SLE, we genotyped 59 SNPs that span the STAT1 and STAT4 genes in our subjects. Fifty-six of these SNPs passed quality control standards and were subsequently used for analyses. The SNPs were evaluated in multiple racial groups (Table 1). To address the population stratification and admixture effect, all outliers have been removed, and results have been corrected based on principal component analyses.

Childhood-onset SLE presents a unique subgroup of patients for genetic study because an earlier disease onset, a more severe disease course, a greater frequency of family history of SLE, and a lesser effect of sex hormones in disease development (34,35) may imply involvement of different genetic factors relative to adult onset disease. Therefore, we initially analyzed these two groups separately. We had a total of 769 samples of the childhood-onset cases. As shown in Fig. 1, the p values for SNPs in STAT4 gene are well-correlated between adults and childhood onset cases, with overall correlation coefficient r =0.84, justifying the joint analysis of the two groups, as presented in all following results. We detected significant associations ($10^{-15}<p<10^{-5}$) in the STAT4 gene. As shown in Table 2, the greatest significance was observed with rs10168266 ($p=1.38\times10^{-15}$) in the Europeans with six other SNPs at $p<10^{-9}$ (rs7568275: 4.26×10^{-15}, rs7582694: 7.67×10^{-15}, rs10181656: 1.16×10^{-14}, rs3024886: 2.71×10^{-13}, rs10174238: 3.30×10^{-13}, rs3821236: 3.41×10^{-11}). Furthermore, all significant SNPs observed in the European population were also strongly significant in the Asian-Korean population, with the strongest significance observed with rs10168266 ((4.00×10^{-10}) (Table 2). Several SNPs were also found to be associated less strongly ($10^{-5}<p<10^{-3}$) in the Hispanic and African populations (Supplementary Table 1 and Fig. 2). In genotype based analyses, the best model of association for almost all of SNPs was additive model (supplementary Table 1). Several SNPs were also found to be associated less strongly ($10^{-5}<p<10^{-3}$) in the Hispanic and African populations (Supplementary Table 1 and Fig. 2). Seventeen SNPs were genotyped in the STAT1 gene, but only produced suggestive results ($0.0005<p<0.05$) (Fig. 2 and Supplementary Table 1). The classical Bonferroni correction for multiple testing is both too strict and inappropriate in studies such as the present one because it assumes that each test is independent, whereas in actuality a complex and unknown mutual dependence is present among SNPs of the same gene. Therefore, for multiple test correction we calculated the false discovery rate (FDR) q values (32) (Table 2, and supplementary Table 1).

Haplotype Structure across the STAT4 gene in different populations

Haplotype analyses in different racial groups identified multiple significant haplotypes (Fig. 3 and Table 3). Particularly in European, three major significant haplotypes have been detected spanning 73 Kb from 3rd intron to exon 17 of STAT4 gene (Fig 3 and Table 3). These three risk haplotypes are: Block1 (13 Kb) AAAG, spanning from exon 17 to intron 14 with $p=9.24\times10^{-14}$, block 2 (18kb) CATTTAAA spanning from intron 14 to exon 4 with $p=4.25\times10^{-14}$ and block 3 (32 Kb) GGCGAGCG located mostly on 3rd intron of STAT4 gene with $p= 1.69\times10^{-15}$. Parts of these three major haplotypes in European were also significant in other ethnical backgrounds with the same sequence (Table 3). Especially, an eight marker haplotype spanning 18 kb across (Block2) was strongly significant in Korean-Asian and

Hispanic population (Table 3). The frequency of this conserved haplotype was 39% in Korean-Asian, 23% in European and 40% in Hispanic patients (Table 3). Conditional analyses on this haplotype showed that SNP rs10168266 explained the whole association in this haplotype. In addition, in African-American, part of this haplotype (11 kb) were also significant with frequency of 70% in lupus cases compare to 65% in controls ($p=7.90\times10^{-3}$) (Table 3). An eight marker haplotype in intron 3 of STAT4 (Block 3, GGCGAGCG) were also significantly associated in the Europeans ($p=1.69\times10^{-15}$) (Table 3). In this haplotype there was no single SNP that could explain the whole association mainly because of the high LD between the most significant associated SNPs rs7568275, rs10181656, rs7582694, and rs10174238 (Fig. 3) located in this haplotype. In fact, conditional logistic regression showed that the GAGCG haplotype as a unit explained the association in this haplotype. Part of this haplotype was also significant in the Hispanic and Asian populations but in African-Americans different haplotype in this block was significant (Table 3). Because of relatively high LD ($D'=0.87$) between the two haplotypes: (CATTTAAA) and (GGCGAGCG) (Block 2 and 3), an extended haplotype consisting of these two was reconstructed (CATTTAAAGGCGAGCG) and reevaluated in European, Korean and Hispanic. Table 3 shows the frequency in case and controls of this extended haplotype which remained significant in multiple ethnicities (in European best $p=4.32\times10^{-14}$). Conditional analyses in this extended haplotype suggest that SNP rs10168266 is the best SNP that could explain the entire omnibus result. Furthermore, when, only responsible variants or units in each haplotype were combined, i.e. SNPs: (rs10168266, rs7568275, rs10181656, rs7582694 and rs10174238) SNP rs10168266, still had an independent effect after controlling for everything else ($p=0.009$).

As expected, the block structure in the STAT4 region generally was less cohesive across the entire region in the African American samples than in the Korean-Asian or the European and therefore the extended haplotype block cannot be reconstructed in the African population (Fig. 3).

Finally, one additional haplotype that was less significantly associated in European ($p=5.70\times10^{-6}$) and Korean-Asian population ($p=6.61\times10^{-8}$) was located at 3′ end of the STAT4 gene (C-terminal domain). This haplotype is almost 25 Kb removed from the highly significant associated SNPs and manifests a weak LD ($D'= 0.54$) with the first block in Figure 3. Conditional analysis with the block 1 haplotype suggests that this association is unlikely to have an independent effect ($p=0.63$).

STAT4 Associations with Clinical Subsets

We performed stratification analyses by gender, age of onset, the 11 ACR criteria, and presence of autoantibodies (anti-Ro, anti-La, anti-RNP, anti-dsDNA, anti-SM) for the SNPs that most likely explain the haplotypic association. After corrections for multiple comparisons, such analysis did not improve the significance of the results (p values) beyond what we have shown in Table 3; However, presence of anti-Ro antibodies in European, did improve the odds ratio to 1.74 (1.36–2.21) in the best associated SNP (rs10168266), while in Korean-Asian early age of onset (less than 13 years old) improved the odds ratio for this SNP to 2.44 (1.66–3.60).

Discussion

We have identified robust associations between the STAT4 (but not STAT1) gene and multiple SNPs in a large study of SLE cases and controls from different racial backgrounds (best combined $p=7.02\times10^{-25}$). To our knowledge this is the first study that targets both STAT4 and STAT1 genes with high resolution SNP genotyping in an extensive case-control study that includes high risk minority African-American, Hispanic, and Korean-Asian populations and a large cohort of childhood-onset SLE.

We found that the observed association with STAT1 is orders of magnitude weaker than that of STAT4 with respect to both p values and number of significant SNPs. This suggestive association with STAT1 most likely reflects the LD with STAT4 since STAT1 and STAT4 are only 25 kb apart. Indeed, we observed (D′=0.96) between rs10199181 in STAT1 among Europeans which produced the best suggestive result ($p = 1.40 \times 10^{-3}$) in this gene and the nearest marker on 3′ end of STAT4 gene rs3024896 ($p = 6.13 \times 10^{-6}$). However, this suggestive result disappeared completely p=0.60 upon conditional analyses. Although the genetic association between SLE and STAT1 cannot be formally excluded, our collection of almost 10,000 samples, which has a 99% power to find effects with odds ratio (OR) = 1.3 at $p = 10^{-8}$, and 84% power to find effects with OR =1.2 at $p = 10^{-6}$ for SNPs with a minor allele frequency of 0.3 and D′=1, makes it highly unlikely that STAT1 is a susceptibility gene for SLE.

Remmers *et al* (23) have demonstrated association of one STAT4 SNP (rs7574865) with SLE in Europeans. Although we did not directly type rs7574865 in our samples, the HapMap CEU data indicate that rs7574865, is a perfect proxy of rs7568275 in European Americans ($r^2=1$). The SNP rs7568275 was one of the top associated SNPs in our data with Fisher combined $p = 1.08 \times 10^{-22}$ (Table 2 and supplementary Table 1). Using HapMap CEU data in European, imputation method support this proxy association for SNP rs7574865 with $p = 6.41 \times 10^{-15}$ in our data. Both of these SNPs are located in the third intron of STAT4.

Lee *et al.* (36) replicated the association of STAT4 with RA in European and Korean patients. Three SNPs (rs10181656, rs13017460, and rs1517352) that were significantly associated with rheumatoid arthritis in Korean patients, were also significant in both European and Asian cases in our SLE subjects with the same associated alleles (Table 2 and supplementary Table 1). In addition, the minor allele frequencies observed between our study and the studies by Remmers *et al*. and Lee *at al*. are similar.

Also, Korman *et al*. (37) reported an association with rs7574865 and primary Sjogren's syndrome (PSS) in a study of 124 Caucasian PSS subjects and 1143 controls (p=0.01). PSS and SLE share overlapping autoantibody profiles (such as anti-Ro) and B lymphocyte hyperactivity, supporting the notion that related autoimmune diseases share common risk variants in STAT4.

Using a relatively dense map of SNPs within the STAT4 gene enabled us to construct haplotypes in the various racial groups. While we identified and strongly confirmed multiple SNP associations in 3rd intron of STAT4 gene, we also detected two additional haplotypes adjacent to this intron by dense SNP genotyping that spans from exon 4 to exon 17 of this gene. Most importantly, one of these haplotypes, an eight marker haplotype spanning 18 kb from exon 4 to exon 14 of the STAT4 gene was significant in multiple ethnicities. Conditional analyses on this haplotype suggest that rs10168266 can explain the whole association in this haplotype and, therefore, the genotyping data for this SNP can predict the risk or protective haplotype. Indeed, this SNP produced the best results in European, Korean-Asian, and Hispanic with the combined Fisher p value of $p = 1.12 \times 10^{-24}$ in our study (Table 2 and supplementary Table 1). The best model of association for this SNP (and most other highly significant markers) was an additive model (European $p = 7.80 \times 10^{-16}$) (supplementary Table 1). This intronic SNP is located between exons 5 and 6 of STAT4 gene and is 28 kb apart from the published rs7574865 at 3rd intron with the estimated LD=0.90 and $r^2=0.62$ between them in European population. However as mentioned previously they are located on different but adjacent haplotypes (block 2 and block 3 respectively) (Figure 3).

Functionally, STAT 4 is the main transcriptional regulatory molecule for IL-12 and, as such, is pivotal to the development of a fully functioning Th1 immune response (38). Polarization

of the immune response to Th1 vs Th2 has profound in vivo consequences. In general, Th1-type immune res ponses are characteristic of cell-mediated immunity, while Th2-type responses are associated with help for B-cell antibody production (humoral immunity) and allergic phenomena (39). Indeed, STAT4-deficient mice are protected from the effects of Th1-cell mediated autoimmune diseases. In models of experimental allergic encephalomyelitis (EAE) (40), experimental arthritis (41), colitis (42), myocarditis (43), and diabetes in the NOD mouse (44), STAT4-deficient mice display less disease, decreased parameters of inflammation, and reduced secretion of IFNγ. Although IFNγ-deficiency does mimic STAT4 deficiency in an arthritis model (41), IFNγ-deficient mice are not protected from EAE (45), myocarditis (43), or colitis (42). Thus, although IFNγ is an important STAT4-induced immune mediator, STAT4 regulates other genes independent of IFNγ that are crucial for the development of such diseases. By contrast, in other autoimmune diseases that are not Th1 mediated, STAT4-deficient mice are not protected from autoimmune diseases that are not Th1-mediated, such as myasthenia gravis or Graves' disease (46,47). Moreover, STAT4-deficiency caused more severe SLE-like disease in NZM 2328 and NZM 2410 mice than in corresponding wild-type control mice (48,49).

Since other cytokines (IL-23, IFNα) in addition to IL-12, can activate STAT4 (20), the phenotype of STAT4-deficient mice, in the different mouse models, may actually be a composite effect of defects in IL-12, IL-23, and IFNα signaling, in addition to other yet unidentified cytokine pathways (20). It should also be stressed that the different effects of STAT4 deficiency in animal models of arthritis compared to those in SLE point to the real possibility that the causal STAT4 polymorphisms in SLE and RA differ from each other.

A relatively small replication study in Swedish SLE patients suggest a significant correlation between the European risk allele and production of anti-dsDNA autoantibodies (50) while a second study in North Americans of European ancestry suggest a strong correlation between SNP rs7574865 and anti-dsDNA Abs and an even stronger association with nephritis (51). These subphenotype associations should be interpreted with caution given that the same SNP alleles are associated with RA patients that do not have anti-dsDNA autoAbs and do not develop nephritis. In this regards, the *in vivo* findings that STAT4 deficient lupus mice develop accelerated nephritis despite decreased levels of anti-dsDNA autoAbs (48,49) might be highly relevant. Subphenotype analyses of our data, in a much larger collection of SLE subjects then these previous studies, do not support a stronger association between STAT4 risk alleles and autoAb levels or presence of nephritis. For example, rs1068266, that produces the best association with nephritis in our European cases (OR=1.59) was not statistically significant from the results obtained with the same SNP in Europeans without nephritis (OR=1.46).

Based on the nature of the SNPs implicated in the present study, it is premature to suggest a molecular mechanism that would explain the gene's association with SLE. Although we used a relatively dense SNP map, the possibility of other polymorphisms that may be responsible for the exact disease mechanism still exists and must await a complete resequencing of the STAT4 gene in SLE patients. Nevertheless, our findings here should significantly advance our understanding and establish new key steps in the pathogenesis of the disease. Furthermore, the unambiguous establishment of STAT4 as a susceptibility gene for SLE provides justification for the developments of therapeutic approaches targeting this molecule or other molecules within its biochemical pathway.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The cooperation of patients and normal control individuals involved in this study is gratefully acknowledged. This work was supported in part by NIH grant RO1AR445650 and ALR grant 52104 to COJ and by the USC FCE. At the Oklahoma Medical Research Foundation (OMRF), the work was supported by the NIH (AR42460, RR015577, AI31584, AR12253, AR48940, DE015223, RR020143, AI062629, AI24717, AI07633, and AR62277), Lupus Foundation of America, the Alliance for Lupus Research, and the U.S. Department of Veterans Affairs. At the University of Alabama at Birmingham, the work was supported by NIH grants P01-AR49084, P60-AR48095, T32-AR07450, M01-RR00032. Members of the PROFILE Study Group include GS Alarcon, E Brown, JC Edberg, BJ Fessler, RP Kimberly, G McGwin Jr, M Petri, R Ramsey-Goldman, J Reveille, LM Vila.

References

1. Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proc Natl Acad Sci U S A 2003;100:2610–5. [PubMed: 12604793]
2. Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, Banchereau J, et al. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. J Exp Med 2003;197:681–5. [PubMed: 12642600]
3. Crow MK, Kirou KA, Wohlgemuth J. Microarray analysis of interferon-regulated genes in SLE. Autoimmunity 2003;36:481–90. [PubMed: 14984025]
4. Baechler EC, Gregersen PK, Behrens TW. The emerging role of interferon in human systemic lupus erythematosus. Curr Opin Immunol 2004;16:801–7. [PubMed: 15511676]
5. Ronnblom LE, Alm GV, Oberg KE. Possible induction of systemic lupus erythematosus by interferon-alpha treatment in a patient with a malignant carcinoid tumour. J Intern Med 1990;227:207–210. [PubMed: 1690258]
6. Gota C, Calabrese L. Induction of clinical autoimmune disease by therapeutic interferon-alpha. Autoimmunity 2003;36:511–518. [PubMed: 14984028]
7. Niewold TB, Hua J, Lehman TJ, Harley JB, Crow MK. High serum IFN-alpha activity is a heritable risk factor for systemic lupus erythematosus. Genes Immun 2007;8:492–502. [PubMed: 17581626]
8. Graham RR, Kozyrev SV, Baechler EC, Reddy MV, Plenge RM, Bauer JW, 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. Nat Genet 2006;38:550–5. [PubMed: 16642019]
9. Cunninghame Graham DS, Manku H, Wagner S, Reid J, Timms K, Gutin A, et al. Association of IRF5 in UK SLE families identifies a variant involved in polyadenylation. Hum Mol Genet 2007;16:579–91. [PubMed: 17189288]
10. Graham RR, Kyogoku C, Sigurdsson S, Vlasova IA, Davies LR, Baechler EC, et al. Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. Proc Natl Acad Sci U S A 2007;104:6758–63. [PubMed: 17412832]
11. Jacob CO, van Der Meide P, McDevitt HO. In vivo treatment of (NZB X NZW)F1 lupus-nephritis with monoclonal antibody to interferon gamma. J Exp Med 1987;166:798–802. [PubMed: 3114409]
12. Ozmen L, Roman D, Fountoulakis M, Schmid G, Ryffel B, Garotta G. Experimental therapy of SLE: the treatment of NZB/W with mouse soluble interferon-gamma receptor. Eur J Immunol 1987;25:6–12. [PubMed: 7843255]
13. Haas C, Ryffel B, Le Hir M. IFN-gamma receptor deletion prevents autoantibody production and glomerulonephritis in lupus-prone (NZBxNZW)F1 mice. J Immunol 1998;160:3713–8. [PubMed: 9558072]
14. Levy DE, Darnell JE Jr. STATs transcriptional control and biological impact. Nat Rev Mol Cell Biol 2002;3:651–662. [PubMed: 12209125]
15. Murray PJ. The JAK-STAT signaling pathway: input and output integration. J Immunol 2007;178:2623–2629. [PubMed: 17312100]
16. Ihle JN. The Stat family in cytokine signaling. Curr Opin Cell Biol 2001;13:211–7. [PubMed: 11248555]
17. Kim HS, Lee MS. STAT1 as a key modulator of cell death. Cellular Signaling 2007;19:454–465.

18. Meraz MA, White JM, Sheehan KC, Bach EA, Rodig SJ, Dighe AS, et al. Targeted disruption of the Stat1 gene in mice reveals unexpected physiologic specificity in the JAK-STAT signaling pathway. Cell 1996;84:431–42. [PubMed: 8608597]
19. Chapgier A, Boisson-Dupuis S, Jouanguy E, Vogt G, Feinberg J, Prochnicka-Chalufour A, et al. Novel STAT1 alleles in otherwise healthy patients with mycobacterial disease. PLoS Genet 2006;2:e131. [PubMed: 16934001]
20. Kaplan MH. STAT4: a critical regulator of inflammation in vivo. Immunol Res 2005;31:231–42. [PubMed: 15888914]
21. Moser KL, Neas BR, Salmon JE, Yu H, Gray-McGuire C, Asundi N, et al. Genome scan of human systemic lupus erythematosus: evidence for linkage on chromosome 1q in African-American pedigrees. Proc Natl Acad Sci U S A 1998;95:14869–74. [PubMed: 9843982]
22. Cantor RM, Yuan J, Napier S, Kono N, Grossman JM, Hahn BH, et al. Systemic lupus erythematosus genome scan: support for linkage at 1q23, 2q33,16q12-13, and 17q21-23 and novel evidence at 3p24, 10q23-24, 13q32, and 18q22-23. Arthritis Rheum 2004;50:3203–10. [PubMed: 15476245]
23. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007;357:977–986. [PubMed: 17804842]
24. Harley JB, Alarcón-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet 2008;2:204–10. [PubMed: 18204446]
25. Hon G, Graham RR, Modrek B, Taylor KE, Ortman W, et al. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. NEJM 2008;358:900–909. [PubMed: 18204098]
26. Kaufman KM, Kelly JA, Herring BJ, Adler AJ, Glenn SB, Namjou B, et al. Evaluation of the genetic association of the PTPN22 R620W polymorphism in familial and sporadic systemic lupus erythematosus. Arthritis Rheum 2006;54:2533–2540. [PubMed: 16868974]
27. Jacob CO, Reiff A, Armstrong DL, Myones BL, Silverman E, Klein-Gitelman M, et al. Identification of novel susceptibility genes in childhood-onset systemic lupus erythematosus using a uniquely designed candidate gene pathway platform. Arthritis Rheum 2007;56:4164–73. [PubMed: 18050247]
28. Alarcon GS, McGwin G Jr, Petri M, Ramsey-Goldman R, Fessler B, Vila LM, et al. Time to renal disease and end-stage renal disease in PROFILE: a multiethnic lupus cohort. PLoS Med 2006;3:e396. [PubMed: 17076550]
29. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40:1725. [PubMed: 9324032]
30. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a toolset for whole-genome association and population-based linkage analysis. American Journal of Human Genetics 2007;81:559–75. [PubMed: 17701901]
31. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263–5. [PubMed: 15297300]
32. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 2003;100:9440–5. [PubMed: 12883005]
33. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904–9. [PubMed: 16862161]
34. Cassidy, JT. Textbook of Pediatric Rheumatology. Cassidy, JT.; Petty, RE., editors. Elsevier Saunders; Philadelphia: 1996. p. 329-406.
35. Lehman, TJA. Dubois' Lupus Erythematosus. Wallace, DJ.; Hahn, BH., editors. Lippincott Williams & Wilkins; Philadelphia: 2002. p. 863-884.
36. Lee HS, Remmers EF, Le JM, Kastner DL, Bae SC, Gregersen PK. Association of STAT4 with rheumatoid arthritis in the Korean population. Mol Med 2007;13:455–60. [PubMed: 17932559]
37. Korman BD, Alba MI, Le JM, Alevizos I, Smith JA, Nikolov NP, et al. Variant form of STAT4 is associated with primary Sjögren's syndrome. Genes Immun 2008;9:267–70. [PubMed: 18273036]
38. Kaplan MH, Sun YL, Hoey T, Grusby MJ. Impaired IL-12 responses and enhanced development of Th2 cells in Stat-4-deficient mice. Nature 1996;382:174–177. [PubMed: 8700209]

39. Mosmann TR, Sad S. The expanding universe of T-cell subsets: Th1, Th2 and more. Immunol Today 1996;17:138–146. [PubMed: 8820272]
40. Chitnis T, Najafian N, Benou C, Salama AD, Grusby MJ, Sayegh MH, et al. Effect of targeted disruption of STAT4 and STAT6 on the induction of experimental autoimmune encephalomyelitis. J Clin Invest 2001;108:739–747. [PubMed: 11544280]
41. Finnegan A, Grusby MJ, Kaplan CD, O'Neill SK, Eibel H, Koreny T, et al. IL-4 and IL-12 regulate proteoglycan-induced arthritis through Stat-dependent mechanisms. J Immunol 2002;169:3345–3352. [PubMed: 12218156]
42. Simpson SJ, Shah S, Comiskey M, de Jong YP, Wang B, Mizoguchi E, et al. T cell-mediated pathology in two models of experimental colitis depends predominantly on the interleukin 12/Signal transducer and activator of transcription (Stat)-4 pathway, but is not conditional on interferon gamma expression by T cells. J Exp Med 1998;187:1225–1234. [PubMed: 9547334]
43. Afanasyeva M, Wang Y, Kaya Z, Stafford EA, Dohmen KM, Sadighi Akha AA, et al. Interleukin-12 receptor/STAT4 signaling is required for the development of autoimmune myocarditis in mice by an interferon-gamma-independent pathway. Circulation 2001;104:3145–3151. [PubMed: 11748115]
44. Yang Z, Chen M, Ellett JD, Fialkow LB, Carter JD, McDuffie M, et al. Autoimmune diabetes is blocked in Stat4-deficient mice. J Autoimmun 2004;22:191–200. [PubMed: 15041039]
45. Ferber IA, Brocke S, Taylor-Edwards C, Ridgway W, Dinisco C, Steinman L, et al. Mice with a disrupted IFN-gamma gene are susceptible to the induction of experimental autoimmune encephalomyelitis (EAE). J Immunol 1996;156:5–7. [PubMed: 8598493]
46. Wang W, Ostlie NS, Conti-Fine BM, Milani M. The susceptibility to experimental myasthenia gravis of STAT6−/− and STAT4−/− BALB/c mice suggests a pathogenic role of Th1 cells. J Immunol 2004;172:97–103. [PubMed: 14688314]
47. Land KJ, Moll JS, Kaplan MH, Seetharamaiah GS. Signal transducer and activator of transcription (Stat)-6-dependent, but not Stat4-dependent, immunity is required for the development of autoimmunity in Graves' hyperthyroidism. Endocrinology 2004;145:3724–3730. [PubMed: 15117875]
48. Jacob CO, Zang S, Li L, Ciobanu V, Quismorio F, Mizutani A, et al. Pivotal role of Stat4 and Stat6 in the pathogenesis of the lupus-like disease in the New Zealand mixed 2328 mice. J Immunol 2003;171:1564–1571. [PubMed: 12874250]
49. Singh RR, Saxena V, Zang S, Li L, Finkelman FD, Witte DP, et al. Differential contribution of IL-4 and STAT6 vs STAT4 to the development of lupus nephritis. J Immunol 2003;170:4818–4825. [PubMed: 12707364]
50. Sigurdsson S, Nordmark G, Garnier S, Grundberg E, Kwan T, Nilsson O, et al. A risk haplotype of STAT4 for systemic lupus erythematosus is over-expressed, correlates with anti-dsDNA and shows additive effects with two risk alleles of IRF5. Hum Mol Genet 2008;18:2868–76. [PubMed: 18579578]
51. Taylor KE, Remmers EF, Lee AT, Ortmann WA, Plenge RM, et al. Specificity of the STAT4 genetic association for severe disease manifestations of systemic lupus erythematosus. PloS Genetics 2008;4:e1000084. [PubMed: 18516230]

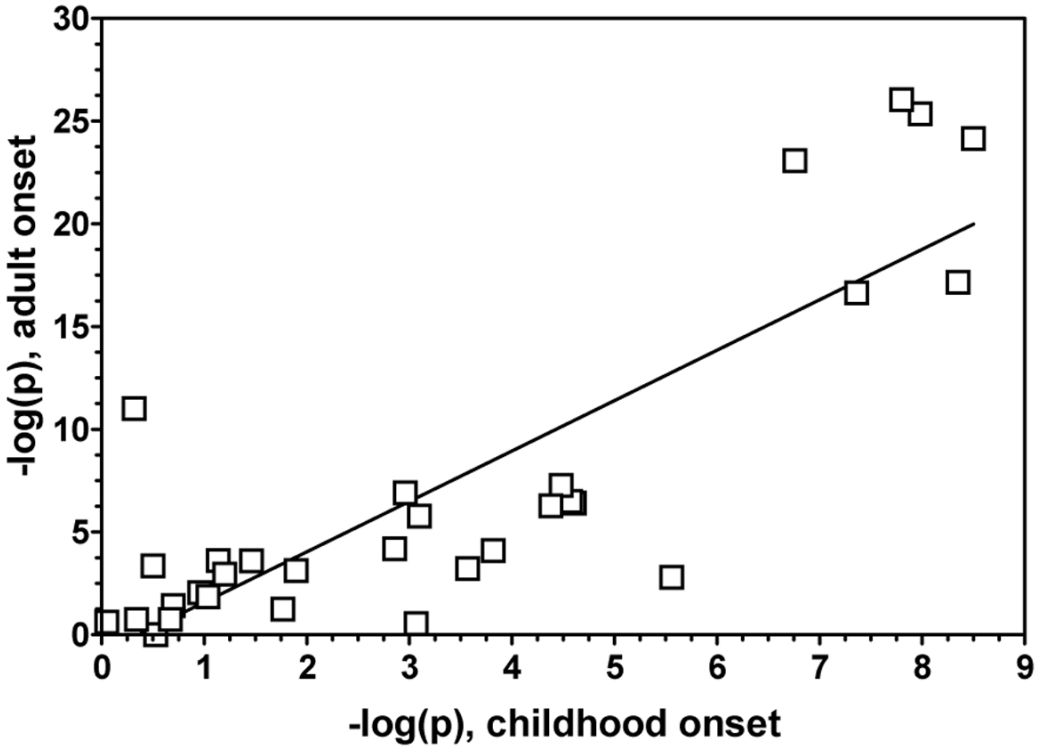

Figure 1.
Correlation plot of the separate analyses of the childhood- and adult-onset SLE cases. The p values are combined across the ethnicities using the Fisher method. The correlation coefficient is r = 0.84.

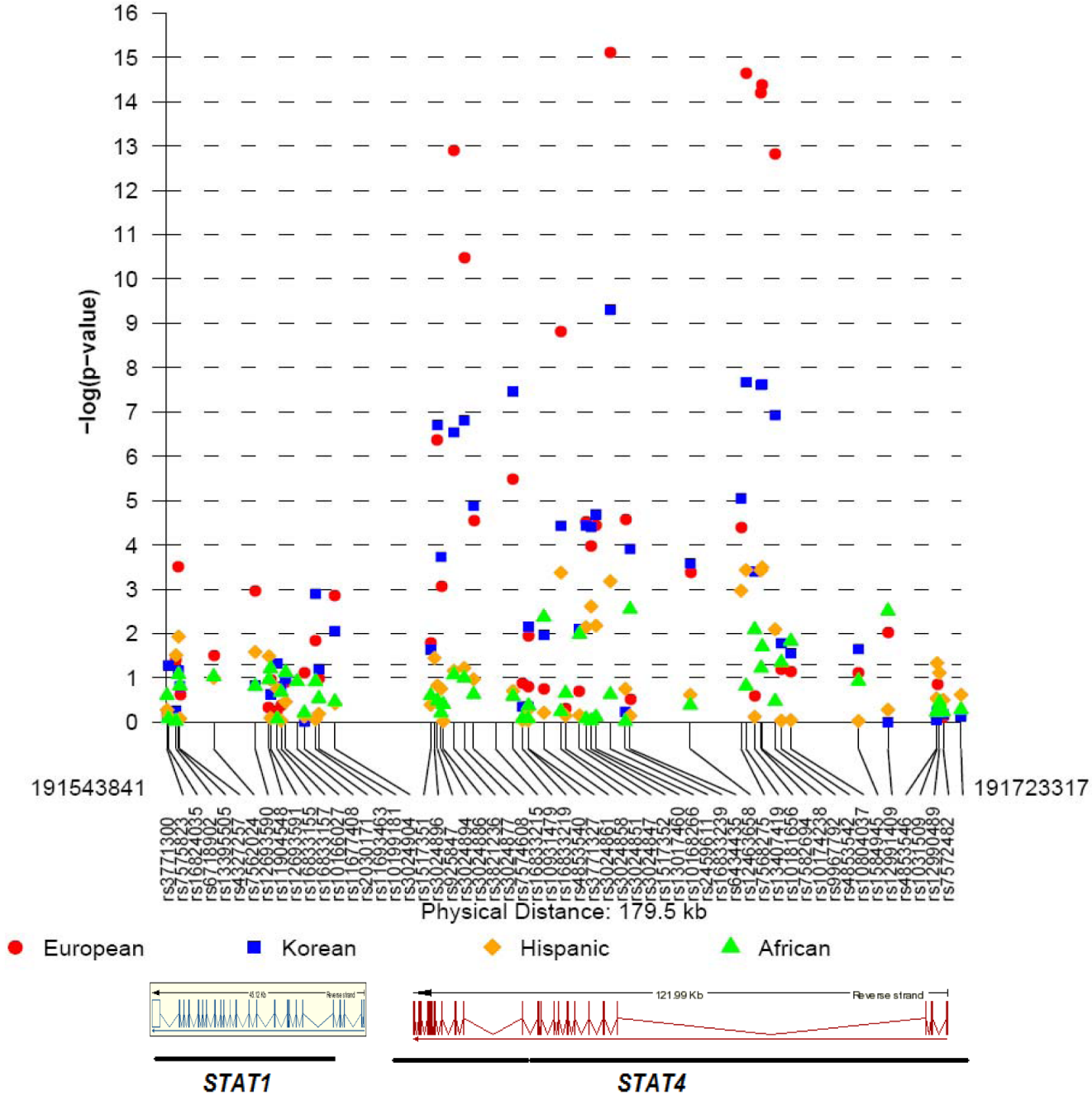

Figure 2.

Association of STAT1-STAT4 SNPs with SLE in four racial groups. The significance of the association is shown for 56 SNP. The position of STAT1 and STAT4 genes are indicated in the bottom part with vertical bars representing the locations of the exons.

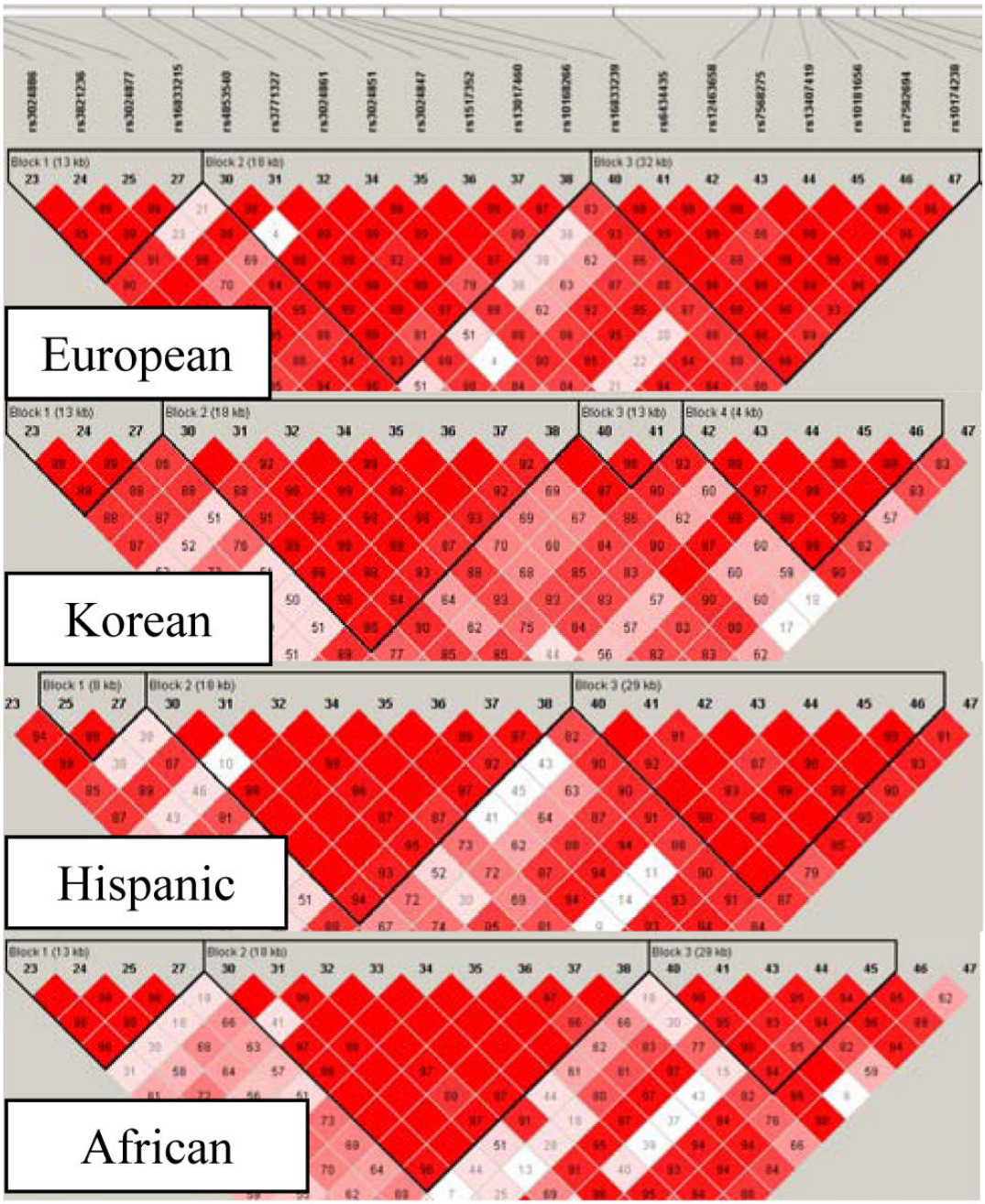

Figure 3.

Schematic representation of multiple significant haplotype block structure in STAT4 gene (spanning 73 Kb from 3rd intron to exon 17) among different racial groups.

Blocks connecting SNP pairs are shaded according to the strength of the linkage disequilibrium between the SNPs, from 0.0 (white) to 1.0 (bright red) as measured by the disequilibrium coefficient D′.

Table 1

The demographic distribution of 9923 independent lupus cases and controls.

	European Case/Control	Korean-Asian Case/Control	African American Case/Control	Gullah* Case/Control	Hispanic Case/Control
Female	2334/2092	621/730	649/565	134/118	604/214
Male:	249/1007	40/51	58/298	19-Dec	70/58
Total	5682 (2583/3099)	1442 (661/781)	1570 (707/863)	283 (146/137)	946 (674/272)

*The Gullah are African Americans who live in the Low Country of South Carolina and genetically show a much lower admixture rate with non African populations than other African Americans.

Table 2

Case Control Allelic Association results with SNP markers selected for STAT4 gene in different racial groups. Significant allelic association results ($p<10^{-5}$) and corresponding FDR q values are listed. (See supplementary Table 1 for more detail).

SNP	Position	Minor Allele	MAF cases	MAF controls	Major Allele	OR (95% CI)	χ^2	P Value (Allelic)	FDR q Value	Ethnicity
rs10168266	191644049	A	0.251	0.184	G	1.49 (1.35–1.65)	66.08	1.38E-15	1.90E-13	European
rs10168266	191644049	A	0.400	0.289	G	1.64 (1.40–1.91)	39.32	4.00E-10	6.87E-09	Korean
rs7568275	191674697	G	0.295	0.225	C	1.45 (1.32–1.59)	63.17	4.26E-15	2.93E-13	European
rs7568275	191674697	G	0.432	0.329	C	1.54 (1.32–1.79)	32.40	1.27E-08	1.71E-07	Korean
rs7582694	191678365	C	0.296	0.227	G	1.44 (1.32–1.58)	61.98	7.67E-15	3.51E-13	European
rs7582694	191678365	C	0.433	0.331	G	1.53 (1.32–1.79)	32.08	1.49E-08	1.71E-07	Korean
rs10181656	191678124	G	0.296	0.227	C	1.44 (1.31–1.58)	61.15	1.16E-14	3.98E-13	European
rs10181656	191678124	G	0.434	0.331	C	1.53 (1.32–1.78)	32.13	1.45E-08	1.71E-07	Korean
rs3024886	191608694	A	0.252	0.191	G	1.44 (1.31–1.59)	55.44	2.71E-13	7.45E-12	European
rs3024886	191608694	A	0.509	0.415	G	1.49 (1.28–1.73)	25.30	4.86E-07	3.93E-06	Korean
rs10174238	191681279	G	0.301	0.235	A	1.41 (1.29–1.55)	54.21	3.30E-13	7.56E-12	European
rs10174238	191681279	G	0.397	0.303	A	1.52 (1.30–1.78)	28.28	1.07E-07	1.05E-06	Korean
rs3821236	191611003	A	0.244	0.187	G	1.40 (1.26–1.54)	45.90	3.41E-11	6.69E-10	European
rs3821236	191611003	A	0.510	0.414	G	1.50 (1.29–1.75)	26.42	2.72E-07	2.34E-06	Korean
rs3024861	191632851	T	0.301	0.247	A	1.32 (1.21–1.45)	36.47	2.90E-09	4.43E-08	European
rs16833215	191622044	G	0.536	0.434	A	1.53 (1.31–1.78)	29.58	5.26E-08	5.56E-07	Korean
rs16833215	191622044	G	0.339	0.295	A	1.23 (1.12–1.34)	22.43	3.66E-06	2.79E-05	European
rs3024896	191604961	A	0.468	0.372	G	1.49 (1.28–1.73)	26.97	2.07E-07	1.89E-06	Korean
rs3024896	191604961	A	0.222	0.185	G	1.26 (1.14–1.39)	21.06	6.13E-06	4.43E-05	European
rs3024877	191613134	A	0.333	0.294	G	1.21 (1.10–1.32)	17.15	3.67E-05	1.96E-04	European
rs3024847	191638535	T	0.441	0.399	A	1.19 (1.10–1.29)	18.17	3.18E-05	1.96E-04	European

Table 3

Significant haplotypic associations in different racial group. The haplotype blocks correspond to Figure 3 (spanning from 3rd intron to exon 17).

Ethnicity	Block #	Sequence	Frequency Case, Control	χ^2	P value
European	1	AAAG	0.253, 0.190	55.52	9.24E-14
	2	C**ATTTAA**A	0.230, 0.169	57.04	4.25E-14
	3	GGCGAGCG	0.290, 0.220	63.39	1.69E-15
* European Combined	2&3	CATTTAAAGGCGAGCG	0.224, 0.163	57.01	4.32E-14
Korean-Asian	1	AAG	0.508, 0.408	28.40	9.85E-08
	2	C**ATTTAA**A	0.386, 0.277	38.10	6.70E-10
	3	GG	0.887, 0.837	14.77	1.00E-04
	4	CGAGC	0.430, 0.327	31.87	1.634E-08
Korean-Asian Combined	2&3&4	CATTTAAAGGCGAGCG	0.298, 0.208	30.15	3.98E-08
Hispanic	1	AG	0.553, 0.490	5.48	1.92E-02
	2	C**ATTTAA**A	0.400, 0.294	16.27	5.49E-05
	3	GGCGAGC	0.430, 0.323	16.36	5.23E-05
Hispanic Combined	2&3	CATTTAAAGGCGAGCG	0.378, 0.288	11.97	5.00E-04
African- American	1	AAAG	0.301,0.265	4.83	2.70E-02
	2	**ATTTAA**	0.701, 0.659	7.06	7.90E-03
	3	AACCC	0.072, 0.102	7.91	4.00E-03

* An extended common haplotype block consist of block 2 and block 3, from 16 SNPs rs4853540, rs3771327, rs3024861, rs3024851, rs3024847, rs1517352, rs13017460, rs10168266, rs16833239, rs6434435, rs12463658, rs7568275, rs13407419, rs10181656, rs7582694, rs10174238 (Fig. 2) were significant in European, Korean-Asian and Hispanic.