ARTHRITIS & RHEUMATOLOGY
Vol. 67, No. 11, November 2015, pp 2966–2977
DOI 10.1002/art.39275

Gene-Based Meta-Analysis of Genome-Wide Association Study Data Identifies Independent Single-Nucleotide Polymorphisms in *ANXA6* as Being Associated With Systemic Lupus Erythematosus in Asian Populations

Jing Zhang,[1] Lu Zhang,[2] Yan Zhang,[2] Jing Yang,[2] Mengbiao Guo,[2] Liangdan Sun,[3] Hai-Feng Pan,[3] Nattiya Hirankarn,[4] Dingge Ying,[2] Shuai Zeng,[2] Tsz Leung Lee,[2] Chak Sing Lau,[2] Tak Mao Chan,[2] Alexander Moon Ho Leung,[5] Chi Chiu Mok,[6] Sik Nin Wong,[6] Ka Wing Lee,[7] Marco Hok Kung Ho,[2] Pamela Pui Wah Lee,[2] Brian Hon-Yin Chung,[2] Chun Yin Chong,[2] Raymond Woon Sing Wong,[2] Mo Yin Mok,[2] Wilfred Hing Sang Wong,[2] Kwok Lung Tong,[8] Niko Kei Chiu Tse,[8] Xiang-Pei Li,[9] Yingyos Avihingsanon,[4] Pornpimol Rianthavorn,[4] Thavatchai Deekajorndej,[4] Kanya Suphapeetiporn,[4] Vorasuk Shotelersuk,[4] Shirley King Yee Ying,[8] Samuel Ka Shun Fung,[8] Wai Ming Lai,[8] Maria-Mercè Garcia-Barceló,[2] Stacey S. Cherny,[2] Pak Chung Sham,[2] Yong Cui,[3] Sen Yang,[3] Dong Qing Ye,[3] Xue-Jun Zhang,[3] Yu Lung Lau,[10] and Wanling Yang[2]

Objective. **Previous genome-wide association studies (GWAS), which were mainly based on single-variant analysis, have identified many systemic lupus erythematosus (SLE) susceptibility loci. However, the genetic architecture of this complex disease is far from being understood. The aim of this study was to investigate whether using a gene-based analysis may help to identify novel loci, by considering global evidence of association from a gene or a genomic region rather than focusing on evidence for individual variants.**

Methods. **Based on the results of a meta-analysis of 2 GWAS of SLE conducted in 2 Asian cohorts, we**

Supported in part by the Shun Tak District Min Yuen Tong of Hong Kong. Drs. Y. L. Lau and W. Yang's work was supported by the S. K. Yee Medical Foundation (general award). Drs. Y. L. Lau, W. Yang, and Chung's work was also supported by the Research Grant Council of the Hong Kong Government (grants GRF 17125114, HKU783813M, HKU781709M, HKU 784611M, and HKU 770411M). Dr. X.-J. Zhang's work was supported by the Chinese Ministry of Education of China (grant IRT-1046). Drs. X.-J. Zhang and Cui's work was supported by the National Key Basic Research Program of China (grant 2014CB541901), the Program for New Century Excellent Talents in University (grant NCET- 12-0600), and the Key Project of the Chinese Ministry of Education (grant 213018A). Dr. S. Yang's work was supported by the National Natural Science Foundation of China (grants 81171505 and 30972727) and the State Key Basic Research Program 973 of China (pre-project grant 2011CB512103). Dr. Ye's work was supported by the Key Program of the National Natural Science Foundation of China (grant 30830089). Dr. Sun's work was supported by the State Key Basic Research Program 973 of China (pre-project grant 2012CB722404). Dr. Hirankarn's work was supported by the Ratchadaphiseksomphot Endowment Fund of Chulalongkorn University (grant for International Research Integration: Chula Research Scholar). Drs. Suphapeetiporn and Shotelersuk's work was supported by the Ratchadapiseksomphot Endowment Fund of Chulalongkorn University (grant RES560530177-HR) and the Thailand Research Fund (grant RTA5680003).

[1]Jing Zhang, PhD: Queen Mary Hospital and The University of Hong Kong, Hong Kong, China, and Eye and ENT Hospital of Fudan University, Shanghai, China; [2]Lu Zhang, MPhil, Yan Zhang, PhD, Jing Yang, PhD, Mengbiao Guo, MPhil, Dingge Ying, PhD, Shuai Zeng, PhD, Tsz Leung Lee, MBBS, Chak Sing Lau, MD, Tak Mao Chan, MD, FRCP, Marco Hok Kung Ho, MBBS, Pamela Pui Wah Lee, MD, Brian Hon-Yin Chung, MD, Chun Yin Chong, MBBS, Raymond Woon Sing Wong, MBBS, Mo Yin Mok, MD, FRCP, FRCPA, Wilfred Hing Sang Wong, MMedSc, Maria-Mercè Garcia-Barceló, PhD, Stacey S. Cherny, PhD, Pak Chung Sham, BM, BCh, MA, Msc, PhD, MRCPsych, Wanling Yang, PhD: Queen Mary Hospital and The University of Hong Kong, Hong Kong, China; [3]Liangdan Sun, MD, PhD, Hai-Feng Pan, MD, Yong Cui, MD, PhD, Sen Yang, MD, PhD, Dong Qing Ye, MD, Xue-Jun Zhang, MD, PhD: Anhui Medical University, China, Hefei, China; [4]Nattiya Hirankarn, MD, PhD, Yingyos Avihingsanon, MD, Pornpimol Rianthavorn, MD, Thavatchai Deekajorndej, MD, Kanya Suphapeetiporn, MD, PhD, Vorasuk Shotelersuk, MD: Chulalongkorn University, Bangkok, Thailand; [5]Alexander Moon Ho Leung, FRCP: Queen Elizabeth Hospital, Hong Kong, China; [6]Chi Chiu Mok, MD, Sik Nin Wong, MBBS: Tuen Mun Hospital, Tuen Mun, New Territories, Hong Kong, China; [7]Ka Wing Lee, MBBS: Pamela Youde Nethersole Eastern Hospital, Hong Kong, China; [8]Kwok Lung Tong, MBBS, FRCP, FHKAM(Med), FHKCP, Niko Kei Chiu Tse, MBBS, Shirley King Yee Ying, MBChB, Samuel Ka Shun Fung, MBBS, FRCP, FHKCP, FHKAM, Wai Ming Lai,

performed an in-depth gene-based analysis followed by replication in a total of 4,626 patients and 7,466 control subjects of Asian ancestry. Differential allelic expression was measured by pyrosequencing.

Results. **More than one-half of the reported SLE susceptibility loci showed evidence of independent effects, and this finding is important for understanding the mechanisms of association and explaining disease heritability. *ANXA6* was detected as a novel SLE susceptibility gene, with several single-nucleotide polymorphisms (SNPs) contributing independently to the association with disease. The risk allele of rs11960458 correlated significantly with increased expression of *ANXA6* in peripheral blood mononuclear cells from heterozygous healthy control subjects. Several other associated SNPs may also regulate *ANXA6* expression, according to data obtained from public databases. Higher expression of *ANXA6* in patients with SLE was also reported previously.**

Conclusion. **Our study demonstrated the merit of using gene-based analysis to identify novel susceptibility loci, especially those with independent effects, and also demonstrated the widespread presence of loci with independent effects in SLE susceptibility genes.**

To date, more than 50 loci have been shown to have a robust association with systemic lupus erythematosus (SLE) (1–23). However, the genetic basis of SLE is far from being understood (24). Several limitations may have contributed to the phenomenon of "missing heritability." First, genetic variants may have different allelic frequencies or linkage disequilibrium (LD) structures across major ethnicities or even in different populations of the same ethnicity. Therefore, the pattern of disease association might be different between different populations (25), making it difficult to replicate previous findings and to predict risk. It is also possible that different variants in a gene having no LD with each other are associated with the underlying disease in different populations, similar to what was shown for *IRF8* (15). Thus, efforts to replicate findings in different populations that focused on a single common single-nucleotide polymorphism (SNP) may miss the association signal.

MBBS: Princess Margaret Hospital, Hong Kong, China; [9]Xiang-Pei Li, MD: Anhui Provincial Hospital, Hefei, China; [10]Yu Lung Lau, MD: Queen Mary Hospital and The University of Hong Kong, Hong Kong, China, and The University of Hong Kong-Shenzhen Hospital, Shenzhen, China.

Address correspondence to Yu Lung Lau, MD, or Wanling Yang, PhD, Department of Paediatrics and Adolescent Medicine, Li Ka Shing Faculty of Medicine and Queen Mary Hospital, The University of Hong Kong, 21 Sassoon Road, Hong Kong, China. E-mail: lauylung@hku.hk or yangwl@hku.hk.

Submitted for publication July 1, 2014; accepted in revised form July 2, 2015.

The second limitation is that exploration of independent effects within an individual locus has been insufficient. Independent contributors are often overlooked in the subsequent replication steps, and the validation of their effects may require much larger sample sizes than are available in most studies. Additionally, coverage of many genomic regions in previously available SNP chips was also limited, thus reducing the chances of capturing independent contributors to disease susceptibility. However, there was strong support for the notion of multiple independent effects from a single locus contributing to disease susceptibility. A recent study identified multiple independent effects in risk loci for several complex diseases and provided evidence that "the variance explained by the multiple effects of independent contributors in a locus was much higher than the variance explained by the single reported SNP effect" (26). The same scenario might be applicable to SLE. In our recent study of SLE in Asian populations (15), 2 of 5 loci showed evidence of independent effects, and the detection of independent contributing variants near an established locus was also reported recently (27).

Gene- or region-based association analysis is an approach that may improve the power of genome-wide association studies (GWAS). Generally, this approach is undertaken to test whether a set of SNPs in a given gene or region shows evidence of association with a trait of interest. Gene-based analysis has several attractive advantages (28). First, by focusing on the association signal from the entire gene rather than the individual SNPs, the analysis could detect associated genes with multiple independent contributors of small effect size that could have been missed by a single variant–based analysis. In addition, gene-based analysis can alleviate the problem of insufficient chip coverage that was encountered in earlier GWAS (29). The effect of an untyped causal SNP might be captured more efficiently by detecting the joint effect of multiple markers that are in LD with it. Such an approach also reduces the multiple-testing burden of GWAS, because it requires correction for only ~20,000 genes rather than the hundreds of thousands of SNPs tested in a typical GWAS. Furthermore, the characteristics of a gene, such as its position, sequence, and function, are highly consistent across human populations. Therefore, gene-based analysis is more likely to yield consistent results across major ethnicities and also provides an opportunity to better understand genetic findings at the functional level (25,30).

Several gene-based association methods have been proposed, such as regression-based tests (25) and integration of SNP-based test statistics from a gene

(31,32). In the regression-based tests, all SNPs in a gene (except those that are redundant) are entered as predictor variables simultaneously. This approach might be hampered by low statistical power due to high dimensionality. In addition, the requirements of raw genotype data and heavy computational demand also limited the application of regression-based tests on a genome-wide scale. Recently, Li et al (28) proposed "a rapid gene-based association test that uses extended Simes procedure (GATES)" as an efficient approach to gene-based testing by integrating the association *P* values for all the contributing SNPs within a gene and its upstream and downstream regions, obtaining an overall *P* value for statistical association within the gene. The statistical power of the GATES method was comparable with that of other alternative gene-based analysis methods (28). This method has been successfully applied in studies aimed at identifying variants conferring risk of schizophrenia in Han Chinese (33).

In the current study, based on meta-analysis of 2 GWAS of SLE in 2 Chinese populations (15), we performed an in-depth gene-based analysis using the GATES method and followed up the findings by performing a replication study in a total of 4,626 patients and 7,466 controls in several Asian cohorts. The results of our analysis confirmed most of the SLE susceptibility genes identified previously at the single-SNP level and showed the widespread presence of independent effects among these genes. In addition, we identified *ANXA6* as a novel SLE susceptibility gene with multiple independently contributing variants. Our results provide new insight into the genetic architecture of this prototypic autoimmune disease.

PATIENTS AND METHODS

Subjects. The samples used in this study overlapped with those used in our previous study (15). Briefly, samples were collected from Hong Kong (612 patients and 2,193 controls in the GWAS stage; 1,027 patients and 1,706 controls in the replication panel), Anhui Province, China (1,044 patients and 1,201 controls in the GWAS stage; 1,463 patients and 1,398 controls in the replication panel), and Bangkok, Thailand (480 patients and 968 controls in the replication stage). Patients in the Hong Kong cohort were attending 5 hospitals in Hong Kong: Queen Mary Hospital, Tuen Mun Hospital, Queen Elizabeth Hospital, Pamela Youde Nethersole Eastern Hospital, and Princess Margaret Hospital. They were all of self-reported Chinese ethnicity and were living in Hong Kong. Controls for the Hong Kong cohort were healthy blood donors at the Hong Kong Red Cross (for the Hong Kong replication panel) and individuals from other GWAS studies performed at the University of Hong Kong, genotyped on the same platform at the same time (GWAS stage). Samples for the Anhui GWAS were collected from patients who attended several hospitals in 2 geographic regions (central and southern China), and the corresponding control subjects were clinically assessed and determined not to have SLE, other autoimmune disorders, systemic disorders, or a family history of autoimmune diseases. All of the SLE patients in the Anhui replication panel were of self-reported Chinese ethnicity, were living in Anhui Province, and were attending the Departments of Rheumatology at Anhui Provincial Hospital and the First Affiliated Hospital of Anhui Medical University in Hefei, Anhui Province, China. Control subjects were chosen from a pool of healthy blood donors recruited from Hefei, Anhui province, with an effort to match for the age and sex of the corresponding SLE patients. The Thai patients with SLE attended the King Chulalongkorn Memorial Hospital, Bangkok. The Thai control subjects were recruited from among unrelated voluntary healthy blood donors with the same ethnic background and geographic location as the Thai patients. All patients recruited for this study had medical records documenting fulfillment of the American College of Rheumatology 1982 revised criteria for the classification of SLE (34). The studies were approved by the respective institutional review boards of all the institutions listed above, and all subjects provided informed consent.

Genotyping. The 2 GWAS from Hong Kong and Anhui were conducted as previously described (15), using an Illumina Human610-Quad BeadChip array. Further replication on selected SNPs was performed using a TaqMan SNP genotyping method with Applied Biosystems Assays-on-Demand probes and primers (catalog nos. C__8723142_10 for rs1561535, C__29329619_10 for rs6707773, C__25934588_10 for rs3815700, C__11836916_10 for rs4958893, C__31986133_10 for rs11960458, C__29349750_10 for rs6896621, C__7513840_20 for rs868531, C__15758511_10 for rs2303028, and C__3169127_20 for rs10036748). The accuracy of genotyping was confirmed by direct sequencing of polymerase chain reaction (PCR) products for 96 randomly chosen samples, which showed 100% concordance. Genotyping concordance between the Human610-Quad BeadChip and the TaqMan SNP genotyping methods was also checked on randomly selected samples, and the 2 methods showed complete concordance.

Association analysis. An association analysis and meta-analysis were performed as previously described (15). The quality control process included removing SNPs with a low call rate (<90%), a low frequency of minor alleles (<1%), violation of Hardy-Weinberg equilibrium in controls ($P \leq 10^{-4}$), removing samples with a low call rate (<90%), and identifying hidden relationships detected using Plink version 1.07 (32). We then applied 2 principal component (PC) corrections to the whole-genome association statistics on the 2 GWAS data sets. Two major PCs (PC1 and PC2) were added as covariates to adjust for population stratification. PC1 and PC2 were calculated with EigenSoft 5.0 (https://github.com/DReichLab/EIG) using whole-genome data, and the SNPTEST program (https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html) was used for the association analysis. Next, we performed a meta-analysis using METAL software (35). Imputation using IMPUTE2.1.1 was carried out in the Hong Kong and Anhui subjects from the GWAS, using SNP genotypes from 286 Asians in the 1000 Genomes Project (June 2011 data release) as the reference. Imputed SNPs with an information score of >0.9 were included for further analyses. The same quality control criteria were applied to process imputed SNPs.

Joint analysis of association, taking into consideration the differences between cohorts, was performed using the

Cochran-Mantel-Haenszel test, and homogeneity of the effect size between diverse cohorts and different stages of the study was examined by Breslow-Day test, both of which were installed in Plink. Tests of independent effects toward disease association for SNPs in a single locus were performed using logistic regression with Plink, adjusted by the effect of the other SNPs in the same locus, including cohort as a covariate.

Gene-based and region-based analyses. GATES (28), implemented in KGG software (36), was used to perform gene-based analysis. SNPs were mapped onto genes according to gene coordinate information from the NCBI GRCh37 database, and SNPs within 10 kb upstream and 10 kb downstream of each gene were also included. A SNP that was located in the overlapping region of 2 genes was assigned to both genes. The r^2 values from the HapMap Han Chinese in Beijing (CHB) and Japanese in Tokyo, Japan (JPT) samples were used to adjust for marker dependency and SNPs in high LD ($r^2 \geq 0.8$ in CHB + JPT were merged).

The construction of gene-based association *P* values was performed as described previously (28). Briefly, letting $P_{(1)}, \ldots, P_{(m)}$ be the ascending *P* values of *m* SNPs within a gene, the overall *P* value for the gene is calculated as follows:

$$P_G = Min\left(\frac{m_e p_{(j)}}{m_{e(j)}}\right),$$

where m_e is the effective number of independent *P* values among the *m* SNPs and $m_{e(j)}$ is the effective number of independent *P* values among the top *j* SNPs.

The value of m_e is estimated to be equal to

$$M - \sum_{i=1}^{M} [I(\lambda i > 1)(\lambda i - 1)]\lambda i > 0,$$

where $I(x)$ is an indicator function and λi is the i^{th} eigenvalue of the *P* value correlation coefficient matrix $[\rho_{i,j}]$ of SNP-based statistic tests. When the SNPs were all independent, the eigenvalues were set at 1, so that m_e was equal to the number of SNPs. When all of the SNPs were in absolute LD, the first eigenvalue was equal to the number of SNPs and the rest were equal to 0, so that $m_e = 1$. For intermediate situations, the *P* value correlation coefficient ρ could be accurately approximated by a 6-order polynomial function of the pairwise allelic correlation coefficient r, and P_G will thus have an approximate uniform (0,1) distribution.

After correction for multiple testing, the significance level for Benjamini and Hochberg (1995) false discovery rate test (37) to control the error rate at 0.05 on the whole genome was 2.0×10^{-4} for the 23,411 genes tested. Briefly, *P* values are ranked in ascending order and denoted by $P_{(1)} \ldots P_{(m)}$, and then the Benjamini and Hochberg procedure controls the false discovery rate ($\alpha = 0.05$). The procedure for a given α is as follows: find the largest κ such that $P(\kappa) \leq \kappa/m\ \alpha$; then reject all $H_{(i)}$ for $I = 1, \ldots, \kappa$. Genes with *P* values less than 2.0×10^{-4} were presented.

We designed a modification that would allow consideration of all intergenic regions in GATES. We applied the sliding window algorithm with a window size of 200 kb and 100-kb overlap between 2 neighboring windows to allow GATES to treat all of the intergenic regions as if real genes were located within. This modification allowed us to analyze a total of 43,729 autosomal regions (including both real and "artificial" genes) using GATES for analysis of association.

Allele-specific transcription quantification. Briefly, 43 healthy individuals heterozygous for rs11960458 were recruited for an analysis of the relative *ANXA6* messenger RNA levels in the 2 alleles (T and C), using pyrosequencing. DNA detection was used as an internal control for normalization of expression. Total RNA obtained from peripheral blood mononuclear cells (PBMCs) from each individual were first treated with DNase. Both complementary DNA (cDNA) and genomic DNA from the same individuals were amplified by PCR, and the PCR amplicons were then purified with a QIAquick PCR Purification Kit (Qiagen). The sequencing primers were designed with Pyrosequencing Assay Design Software version 1.0. Reactions were performed on a Biotage PSQ 96MA machine, and allele quantification was analyzed by PSQ MA 2.1 software. T:C allele detection was performed for both DNA and cDNA, with the ratio for DNA being used as a means of normalization on the detection efficiency. A paired Student's *t*-test was adopted to compare expression from the T and C alleles for this gene.

RESULTS

Gene-based association analysis. As reflected by the results after performing PC analysis to adjust for admixture, no significant differences between 2 Chinese populations were apparent in terms of population structure. There was no substantial difference between SNP-based *P* values before and after adjustment for admixture, as the majority of differences (99.83%) were within one order of magnitude (see Supplementary Figure 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39275/abstract).

After adjustment for admixture, SNP-based *P* values were subjected to gene-based analysis by GATES. A total of 324,519 SNPs (60.44%) were mapped to 1 or more of the 23,411 genes. The inflation factor for gene-based association was 1.047.

We examined the Q–Q plot, comparing the expected and observed gene-based *P* values by removing genes in known susceptibility loci (4–6,9,10,12–16,18–20,23,27). After removing these genes, we still observed an excess of association signal (see Supplementary Figure 2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39275/abstract), suggesting that additional genes are associated with the disease. After correction for multiple testing, 83 genes reached gene-based genome-wide significance ($P < 2.0 \times 10^{-4}$), of which 41 genes were located in non–major histocompatibility complex (non-MHC) regions. Twenty-five of the 41 genes within 18 loci were located in regions for which SLE associations were already established in earlier studies based on single-variant analysis (Tables 1 and 2). Some of these genes are close to each other and are located in the same LD block and thus were assigned to the same locus, such as *LRRC18* and *WDFY4*, *HIC2* and *UBE2L3*, and *BLK* and *FAM167A*.

Table 1. Gene-based association analysis confirming association of the known SLE susceptibility loci with evidence of independent effects*

Chromosome, gene	Candidate independent SNPs	P_{adj}	P_{unadj}	Reported independent SNPs (refs.)	Marker dependency†
1, *TNFSF4* ($P = 1.79 \times 10^{-8}$) (best reported SNP rs2205960 [$P = 6.97 \times 10^{-14}$])	rs1418190	7.10×10^{-5}	1.30×10^{-5}	rs10798269 (4, 13)	$r^2 = 0.83$
	rs1418191	1.90×10^{-4}	5.98×10^{-5}	–	–
	rs1539259	2.08×10^{-4}	6.28×10^{-5}	–	–
	rs1578624	2.52×10^{-4}	8.08×10^{-5}	–	–
	rs704840	1.91×10^{-2}	9.77×10^{-12}	rs704840 (12)	Direct tagging
2, *STAT4* ($P = 1.0 \times 10^{-19}$) (best reported SNP rs7574865 [$P = 6.22 \times 10^{-21}$])	rs16833239	2.83×10^{-5}	5.02×10^{-10}	–	–
	rs7594501	3.56×10^{-5}	6.96×10^{-10}	–	–
	rs3771327	6.08×10^{-5}	2.95×10^{-8}	–	–
	rs1400656	7.16×10^{-5}	5.08×10^{-9}	–	–
	rs7601754	1.33×10^{-4}	2.43×10^{-10}	rs7601754 (13, 37)	Direct tagging
	rs7572482	9.25×10^{-4}	0.004297	–	–
	rs10931481	1.14×10^{-3}	1.07×10^{-6}	–	–
	rs1517352	2.09×10^{-2}	6.59×10^{-6}	–	–
	rs6740131	3.19×10^{-2}	6.19×10^{-5}	–	–
TET3 ($P = 2.25 \times 10^{-6}$) (best reported SNP rs6705628 [$P = 3.85 \times 10^{-7}$])	rs4852324	8.71×10^{-3}	3.29×10^{-8}	rs4852324 (15)	Direct tagging
	rs940296	8.64×10^{-3}	2.10×10^{-8}	rs4852324 (15)	$r^2 = 0.92$
4, *BANK1* ($P = 4.10 \times 10^{-5}$) (best reported SNP rs4522865 [$P = 9.26 \times 10^{-7}$])	rs10516487	2.27×10^{-2}	2.76×10^{-5}	rs10516487 (6, 12)	Direct tagging
6, *UHRF1BP1* ($P = 4.26 \times 10^{-12}$) (best reported SNP rs13205210 [$P = 5.51 \times 10^{-12}$])	rs3734266	2.38×10^{-2}	7.17×10^{-13}	rs11755393 (38)	Direct tagging
7, *IRF5*($P = 2.88 \times 10^{-9}$) (best reported SNP rs4728142 [$P = 7.80 \times 10^{-9}$])	rs729302	1.48×10^{-6}	8.84×10^{-10}	rs729302 (12, 13)	Direct tagging
10, *LRRC18_WDFY4* ($P = 1.47 \times 10^{-6}$) (best reported SNP rs1913517 [$P = 2.52 \times 10^{-6}$])	rs877819	2.03×10^{-4}	2.21×10^{-8}	rs877819 (13, 39)	Direct tagging
	rs2943244	1.56×10^{-2}	1.48×10^{-6}	–	–
	rs2663052	3.83×10^{-2}	1.27×10^{-7}	rs7097397 (12)	$r^2 = 0.81$
12, *CDKN1B* ($P = 6.48 \times 10^{-6}$) (best reported SNP rs10845606 [$P = 6.05 \times 10^{-6}$])	rs12822507	4.26×10^{-3}	0.000203	rs12822507 (15)	Direct tagging
	rs11055008	1.17×10^{-2}	6.16×10^{-7}	–	–
16, *CLEC16A* ($P = 4.29 \times 10^{-6}$) (best reported SNP rs12599402 [$P = 1.04 \times 10^{-7}$])	rs17673553	1.91×10^{-2}	0.00065	–	–
22, *HIC2_UBE2L3* ($P = 2.72 \times 10^{-7}$) (best reported SNP rs463426 [$P = 0.000289$])	rs5754217	7.51×10^{-3}	2.99×10^{-5}	rs5754217 (13, 40)	Direct tagging
	rs4821112	8.65×10^{-3}	3.76×10^{-5}	–	–

* The P values for the best reported single-nucleotide polymorphisms (SNPs) in each individual locus were derived from meta-analysis of the 2 systemic lupus erythematosus (SLE) genome-wide association studies (GWAS). P_{adj} = P values for each candidate independent SNP, adjusted for the effect of the best reported SNP in each individual locus and potential cohort differences (using cohort as a covariate) ($P_{adj} < 0.05$ was set as the cutoff for screening candidate SNPs); P_{unadj} = meta-analysis P values for the candidate independent SNPs based on the 2 SLE GWAS ($P_{unadj} < 0.001$ was set as the cutoff for screening candidate independent SNPs).
† Indicates linkage disequilibrium between each candidate independent SNP and the reported independent SNP, shown as the r^2 value in the HapMap Han Chinese in Beijing and Japanese in Tokyo, Japan populations.

Ten of the 18 known SLE susceptibility loci showed evidence of multiple independent effects (adjusted P [P_{adj}] < 0.05 and unadjusted P [P_{unadj}] < 0.001) (Table 1). The independent effects within all of these loci except *CLEC16A* have already been confirmed via direct or indirect (i.e., surrogate variants having high LD with them [$r^2 > 0.8$]) tagging of at least 1 independent contributor in previous replication studies (4,6,12,13,15,27,38–41). Therefore, the results of gene-based analysis were proven to be consistent with what has been reported in the literature. The roles of other potentially independent variants in the reported loci that did not reach genome-wide significance require confirmation by conducting replication analysis in additional samples.

The results also suggested that 16 novel non-MHC genes within 13 loci might be associated with SLE. Three of them showed potential evidence of independent effects (see Supplementary Table 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39275/abstract). Thus, we per-

Table 2. Gene-based association analysis confirming association of the known SLE susceptibility locus without evidence of independent effects

Chromosome, gene	P	Best reported SNP	P*
3, *CD80*	1.33×10^{-4}	rs6804441	9.24×10^{-6}
5, *TNIP1*	6.75×10^{-5}	rs10036748	3.10×10^{-6}
6, *TNFAIP3*	8.81×10^{-13}	rs2230926	2.20×10^{-13}
8, *BLK_FAM167A*	1.81×10^{-12}	rs7812879	2.88×10^{-13}
11, *ETS1*	5.22×10^{-5}	rs6590330	1.37×10^{-7}
12, *SLC15A4*	3.55×10^{-8}	rs1385374	2.58×10^{-9}
13, *ELF1*	5.86×10^{-5}	rs7329174	6.21×10^{-6}
22, *YDJC*	2.61×10^{-7}	rs2298428	1.31×10^{-7}

* Meta-analysis values of the best single-nucleotide polymorphisms (SNPs) in each individual locus from the 2 systemic lupus erythematosus (SLE) genome-wide association studies.

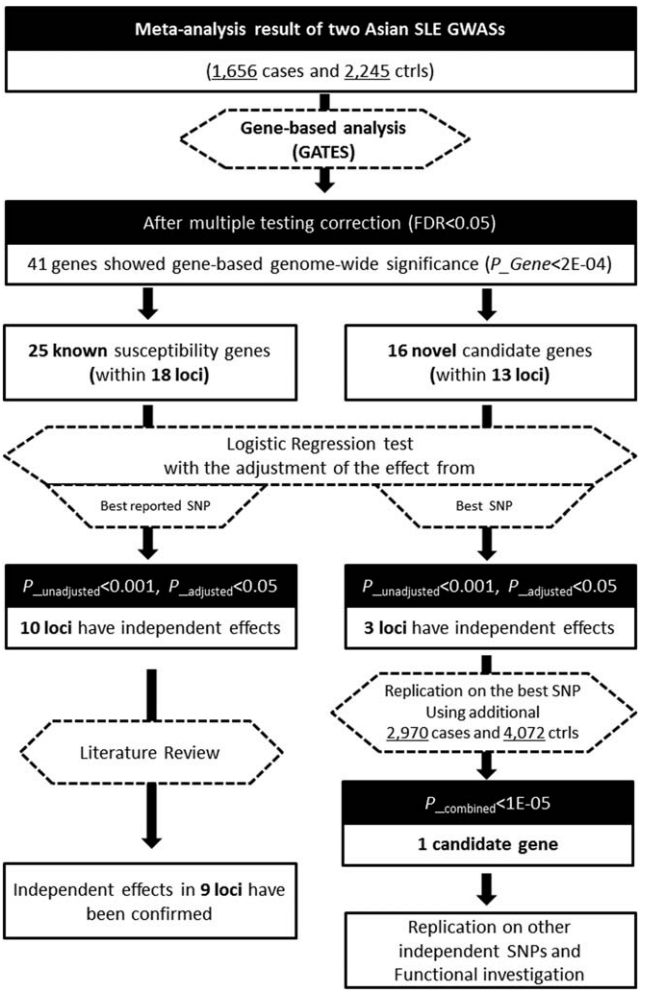

Figure 1. Flow chart of the experimental process and criteria for selection of the single-nucleotide polymorphisms (SNPs). SLE = systemic lupus erythematosus; GWAS = genome-wide association study; FDR = false discovery rate.

formed replication on the most significant SNPs in these 3 loci, using 2,970 additional SLE patients and 4,072 controls collected from Hong Kong, Anhui, and Thailand. A joint analysis of the results for all 3 cohorts showed that an intronic variant in *ANXA6* (rs4958893) was significantly associated with SLE (combined $P = 3.12 \times 10^{-7}$) (see Supplementary Table 2, available on the *Arthritis & Rheumatology* web site at http://online library.wiley.com/doi/10.1002/art.39275/abstract), and this locus was followed up further. Figure 1 shows a flow chart of the steps taken in the current study Figure 1.

Region-based association analysis. In order to detect association signals located in the intergenic regions, we designed a modification that would allow consideration of all of the intergenic regions in a gene-based analysis (see Patients and Methods for details). After correcting for multiple testing, we observed that 38 additional intergenic regions reached genome-wide significance. Seventeen of these regions were located in or near the MHC region, and the others were mostly in high LD with the known loci or with candidate loci identified by the analysis that considered only the real gene regions, such as those where *ETS1*, *TNFSF4*, *CDKN1B*, and *ANXA6* are located (for additional details, see Supplementary Table 3, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39275/abstract). Interestingly, with such modification, we further identified association signals from the *MIR146A*, *PRDM1*, and *IKZF1* regions that were missed when the intergenic regions were not systematically considered.

Association of the *ANXA6* locus with SLE. When we examined the meta-analysis results in the *ANXA6* region, 11 SNPs showed a meta-analysis *P* value (P_{meta}) of <0.001 (Figure 2A) (also see Supplementary Table 4, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39275/abstract) and LD patterns among them revealed that they could be grouped into 5 clusters (see Supplementary Figure 3, available on the *Arthritis & Rheumatology* web site at http://online library.wiley.com/doi/10.1002/art.39275/abstract). Imputation followed by meta-analysis were performed to obtain information on additional SNPs in this region. However, none of the imputed SNPs showed *P* values that were of greater significance than those of the interrogated SNPs. Additionally, evidence supporting the independent functional potential of these SNPs in disease pathogenesis was not adequate to warrant additional replication. For example, none of the imputations resulted in amino acid substitutions (see Supplementary Table 5, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39275/abstract) by SCAN annotation (http://www.scandb.org/).

RegulomeDB annotations (http://www.regulomedb.org/) were also investigated to identify SNPs in or near regulatory elements, and several such imputed SNPs were shown to have potential regulatory functions. However, these SNPs were all shown to have high LD with the genotyped SNPs (see Supplementary Table 5, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39275/abstract). Therefore, 4 additional SNPs, namely rs11960458, rs6896621, rs868531, and rs2303028, all of which were interrogated during the GWAS stage, were selected for further replication to identify independent contributors to disease risk within *ANXA6*. In order to exclude the possibility that the association of *ANXA6* with SLE was dependent on the effect of *TNIP1,* a neighboring gene that was previously

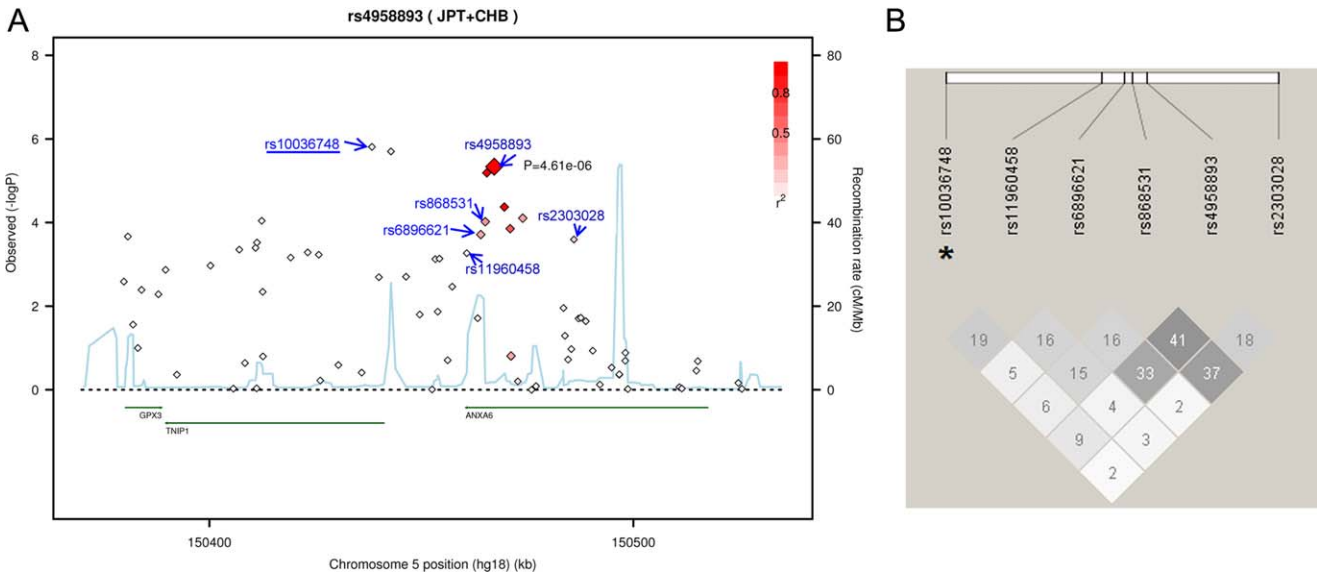

Figure 2. Systemic lupus erythematosus–associated single-nucleotide polymorphisms (SNPs) in the *TNIP1–ANXA6* region and linkage disequilibrium (LD) patterns among these SNPs, as observed in the current study. **A,** SNPs plotted according to chromosome position (NCBI build 36.3) and meta-analysis *P* values. SNP rs4958893, an intronic variant in *ANXA6*, had the strongest signal, and the colors of the other SNPs reflect their LD with rs4958893 (based on pairwise r^2 values from the meta-analysis data). The reported SNP in *TNIP1* (rs10036748) is underlined. Estimated recombination rates from HapMap Han Chinese in Beijing (CHB) and Japanese in Tokyo, Japan (JPT) samples (shown in the boxes) are plotted to reflect the local LD structure. The plot was generated using SNAP (www.broadinstitute.org/mpg/snap). **B,** Correlation coefficient (r^2) values for individuals in the HapMap 3 (release no. 2) CHB and JPT populations. The **asterisk** indicates the reported SNP in *TNIP1*.

reported to be associated with SLE (13), SNP rs10036748 was also included for replication in additional samples.

Joint analysis of the data for all 3 cohorts revealed association of both rs11960458 and rs6896621 with SLE (combined $P = 5.66 \times 10^{-7}$ and combined $P = 6.72 \times 10^{-6}$, respectively) (Table 3). The association of SNP rs10036748 in *TNIP1* was also replicated in the current study (combined $P = 1.94 \times 10^{-10}$). For SNPs rs4958893, rs11960458, rs686921, and rs10036748, the same trends were observed in all the replication cohorts (see Supplementary Figure 4, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39275/abstract). Importantly, independence tests showed that the association of SNPs rs4958893, rs11960458, and rs6896621 remained significant after adjustment for the effect of rs10036748 in *TNIP1* (Table 3). Furthermore, there were also independent effects among these 4 SNPs, because each SNP remained significant after adjustment for the effects of the other 3 SNPs (Table 3). Consistent with the results of independence tests, LD between these SNPs was low to moderate ($r^2 < 0.4$) (Figure 2B). In addition, the extended association down to *GPX3* observed from the data might be due to the *TNIP1* association, as indicated by the LD pattern between rs10036748 and the SNPs in this region (see Supplementary Figure 5, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39275/abstract). SNPs in *GPX3* were in moderate to high LD with rs10036748, and their *P* values were no longer significant when the effect from rs10036748 was adjusted by logistic regression (see Supplementary Table 6, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39275/abstract).

***ANXA6* expression profile and expression quantitative trail locus (eQTL) analysis.** In order to detect potential functions of the SNPs in transcription regulation, we queried the publicly available eQTL database GHS Express, which assessed the transcriptome of circulating monocytes in 1,490 unrelated individuals (42). We observed that rs4958893 was in absolute LD ($r^2 = 1$) with rs868641, a *cis* eQTL for *ANXA6* ($P = 3.41 \times 10^{-11}$). Additionally, the regulatory effect of rs868641 on *ANXA6* expression in monocytes was also confirmed by a recent study of cell type–specific eQTLs (43). SNP rs11960458 was also in high LD ($r^2 = 0.94$) with another *cis* eQTL for *ANXA6* (rs4958891; $P = 1.41 \times 10^{-10}$). Detailed results of the eQTL analyses are shown in Supplementary Tables 7 and 8 (available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39275/abstract).

Table 3. Replication results of SNPs in *ANXA6* and *TNIP1**

	MAF														
	Hong Kong			Anhui			Thailand			Combined	Combined		Conditional	Conditional
SNP	Patients	Controls	*P*	Patients	Controls	*P*	Patients	Controls	*P*	*P*	OR (95% CI)	P_{het}	*P*†	*P*‡
rs10036748	0.22	0.26	1.03×10^{-4}	0.20	0.25	1.82×10^{-7}	0.31	0.33	3.50×10^{-1}	1.94×10^{-10}	0.81 (0.76–0.86)	0.2703	–	2.14×10^{-7}
rs4958893	0.27	0.31	2.12×10^{-4}	0.30	0.33	8.52×10^{-4}	0.33	0.36	1.89×10^{-1}	3.12×10^{-7}	0.86 (0.81–0.91)	0.8223	2.32×10^{-64}	0.009352
rs11960458	0.48	0.45	4.29×10^{-4}	0.48	0.45	1.16×10^{-3}	0.34	0.32	1.41×10^{-1}	5.66×10^{-7}	1.15 (1.09–1.21)	0.9518	8.51×10^{-8}	0.002212
rs6896621	0.33	0.35	4.70×10^{-3}	0.32	0.36	2.29×10^{-4}	0.40	0.41	6.32×10^{-1}	6.72×10^{-6}	0.88 (0.83–0.93)	0.4675	6.92×10^{-15}	8.15×10^{-5}
rs868531	0.46	0.42	1.62×10^{-3}	0.44	0.42	1.86×10^{-2}	0.29	0.33	5.56×10^{-2}	0.001939	1.09 (1.03–1.15)	0.00826	0.0646	–
rs2303028	0.42	0.40	2.34×10^{-2}	0.38	0.37	2.08×10^{-1}	0.29	0.34	1.83×10^{-2}	0.1095	1.05 (0.99–1.11)	0.00748	0.3932	–

* The Hong Kong cohort comprised 1,639 patients and 3,899 controls. The Anhui cohort comprised 2,507 patients and 2,599 controls. The Thailand cohort comprised 480 patients and 968 controls. MAF = minor allele frequency; OR = odds ratio; 95% CI = 95% confidence interval; P_{het} = *P* value for heterogeneity.
† Adjusted for the effect of rs10036748 from all of the samples, with cohort treated as a covariate.
‡ Each of the 4 single-nucleotide polymorphisms (SNPs) was adjusted for the effect of the other 3 SNPs.

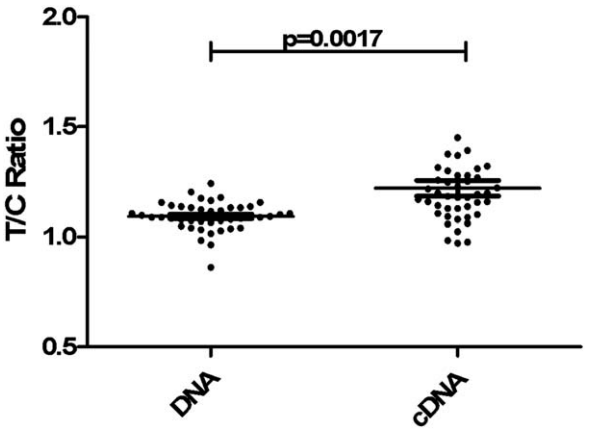

Figure 3. Differential allelic expression of *ANXA6* based on single-nucleotide polymorphism (SNP) rs11960458 in peripheral blood mononuclear cells (PBMCs) from healthy individuals. Differential allelic expression of *ANXA6* in PBMC cDNA and DNA prepared from 43 healthy individuals heterozygous for rs11960458 was measured by pyrosequencing. The T:C ratios for both DNA and cDNA are shown. Bars show the median and 95% confidence interval.

Interestingly, SNP rs11960458 is located in the 3′-untranslated region (3′-UTR) of *ANXA6*, and we speculated that it may affect the expression level of *ANXA6* or is in high LD with a variant that does. Therefore, we examined differential allelic expression of *ANXA6* based on rs11960458 in PBMCs from healthy Hong Kong Han Chinese individuals heterozygous for the SNP (n = 43). The risk allele (T) was significantly correlated with a higher level of *ANXA6* expression ($P = 0.0017$) (Figure 3). These results suggested that rs11960458, or a SNP or SNPs in high LD with it, might play a regulatory role and may predispose the risk allele carriers to SLE through increased expression of *ANXA6*.

DISCUSSION

In this study, using gene-based testing of the results of a meta-analysis in 2 Asian GWAS, we confirmed multiple previously identified SLE susceptibility genes and showed the widespread presence of independent effects among known SLE susceptibility genes. Importantly, we identified *ANXA6* as a novel SLE susceptibility gene that contained several independent contributors. *ANXA6* did not contain a single SNP that could potentially reach genome-wide significance based on the current sample size, the effect size, and allele frequency. Thus, the genetic association would not have been recognized without application of an alternative approach such as gene-based association analysis.

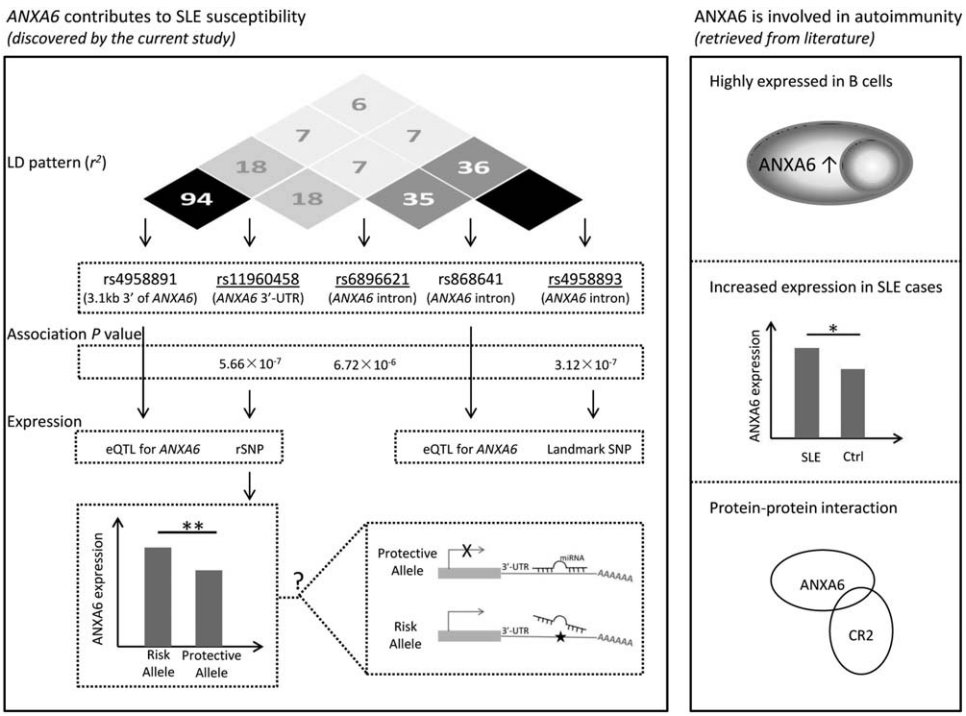

Figure 4. Summary of the association of *ANXA6* with systemic lupus erythematosus (SLE) and potential link between *ANXA6* and autoimmunity. * = $P < 0.05$; ** = $P < 0.01$. LD = linkage disequilibrium; eQTL = expression quantitative trail locus.

We also identified rs11960458 as a regulatory SNP; its risk allele correlated with higher expression of *ANXA6*, which provides a functional interpretation for the genetic association. The expression of *ANXA6* was high in immune-related cells, and increased expression of *ANXA6* was also observed in patients with SLE (see Supplementary Figures 6 and 7, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39275/abstract). In addition, a previous study demonstrated that annexin A6 interacts with human complement receptor 2 (CR2/CD21) (44). CR2 is the receptor for complement C3Dd and Epstein-Barr virus (EBV) binding on B cells and T cells and plays a central role in immune responses to foreign antigens and the development of autoimmunity to nuclear antigens in SLE (45). EBV has been considered to be an important environmental trigger for SLE (46,47). Additionally, genetic variations in CR2 were shown to be associated with SLE (48,49). Thus, the potential functional link between annexin A6 and CR2 supported the genetic findings of *ANXA6* as a novel SLE susceptibility gene.

Furthermore, 2 variants in *ANXA6* (SNPs rs4958893 and rs11960458) were in high LD with eQTLs that affect *ANXA6* expression. SNP rs11960458, which is located in the 3′-UTR of *ANXA6*, showed evidence of differential allelic expression, with the risk allele correlating with higher expression of *ANXA6*. These data indicate that multiple mechanisms may be involved in the association of this locus with disease. Interestingly, rs11960458 was also mapped onto a microRNA binding site of *ANXA6* in miRBase (http://www.mirbase.org/). Destruction of the microRNA binding site (as for miR-1234-3p and miR-4479) by this SNP might confer increased expression/stability, although the detailed mechanisms still require further investigation. A summary of the current findings and the potential link between *ANXA6* and autoimmunity is shown in Figure 4.

The current study also has several limitations, and the exploration of SLE susceptibility genes is by no means complete. Genes such as *ARID5B, ITGAM*, and *DRAM1*, whose associations have been identified in our previous studies, did not show strong evidence of association when analyzed at the gene level. One of the underlying reasons is the definition of gene boundary. As reported by the ENCODE (Encyclopedia of DNA Elements) collaborators, the majority of the regulatory elements are located close to the gene, although distant regulatory elements do exist (50). In the current study, the gene boundary was defined as 10 kb upstream and 10 kb downstream of each gene. The range of the gene boundary, which determines the number of SNPs to be analyzed for the gene, may affect the results of gene-based association studies. It is also important to recognize that there is a tradeoff for the gain in power when this method is used, because it may mask or dilute individual SNP effects when the SNPs in a gene are considered together; as mentioned above, this is especially true for large genes, such as *ARID5B*, *DRAM1*, and *ITGAM*. The lack of a convincing signal in *ITGAM* might also be attributable to the low frequency of risk alleles in this gene in Asian populations (minor allele frequency $<5\%$) (11).

In addition, a gene-based analysis can cover only SNPs within or near a gene, and some intergenic SNPs may also be associated with complex diseases and could be missed by this approach. In order to solve this problem, we designed an in-house algorithm to consider all the intergenic regions larger than 10 kb in a gene-based association analysis. We observed that most of the intergenic regions that showed significant association signals were those in high LD with the known association loci or candidate loci that were identified in the analysis considering only real genes. Interestingly, with such modification, we successfully identified the association signals from *MIR146A*, *PRDM1*, and *IKZF1*, which were not prominent by the gene-based method using a strict definition of genes. Application of this region-based approach might be useful in studies of other complex diseases. Information regarding this modification is available on the web site for the University of Hong Kong laboratory (http://paed.hku.hk/genome).

Recently, a resequencing-based study was conducted in 24,892 patients with 6 autoimmune disease phenotypes and 17,019 control subjects (51). The investigators concluded that the missing heritability for common autoimmune diseases "may be a result of many common-variant loci of weak effect." These results support the merit of using a gene-based approach to explore independent effects in susceptibility loci for complex diseases. Additionally, with the support of many integrated bioinformatics databases, application of a gene-based approach could be further extended to explore novel biologic pathways and protein–protein interaction networks that may play an important role in complex diseases.

ACKNOWLEDGMENTS

We thank Winnie Lau and her team for collecting samples and clinical records for the Hong Kong patients. We also thank PK and AC for their contributions to Thai pediatric patient care and sample collection.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Drs. Y. L. Lau and W. Yang had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. J. Zhang, Y. L. Lau, W. Yang.

Acquisition of data. J. Zhang, L. Zhang, Y. Zhang, J. Yang, Guo, Sun, Pan, Hirankarn, D. Ying, Zeng, T. L. Lee, C. S. Lau, Chan, Leung, C. C. Mok, S. N. Wong, K. W. Lee, Ho, P. P. W. Lee, Chung, Chong, R. W. S. Wong, M. Y. Mok, W. H. S. Wong, Tong, Tse, Li, Avihingsanon, Rianthavorn, Deekajorndej, Suphapeetiporn, Shotelersuk, Fung, Lai, Garcia-Barceló, Cherny, Sham, Cui, S. Yang, Ye, X.-J. Zhang, Y. L. Lau, W. Yang.

Analysis and interpretation of data. J. Zhang, L. Zhang, Y. Zhang, Guo, W. Yang.

REFERENCES

1. 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. New Engl J Med 2007;357: 977–86.
2. Cunninghame Graham DS, Graham RR, Manku H, Wong AK, Whittaker JC, Gaffney PM, et al. Polymorphism at the TNF superfamily gene TNFSF4 confers susceptibility to systemic lupus erythematosus. Nat Genet 2008;40:83–9.
3. Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, et al. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. Nat Genet 2008; 40:1059–61.
4. Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet 2008;40:204–10.
5. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, et al. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. New Engl J Med 2008; 358:900–9.
6. Kozyrev SV, Abelson AK, Wojcik J, Zaghlool A, Linga Reddy MV, Sanchez E, et al. Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nat Genet 2008;40:211–6.
7. Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W, et al. Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. Nat Genet 2008;40:1062–4.
8. Nath SK, Han S, Kim-Howard X, Kelly JA, Viswanathan P, Gilkeson GS, et al. A nonsynonymous functional variant in integrin-α(M) (encoded by ITGAM) is associated with systemic lupus erythematosus. Nat Genet 2008;40:152–4.
9. Gateva V, Sandling JK, Hom G, Taylor KE, Chung SA, Sun X, et al. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. Nat Genet 2009;41:1228–33.
10. Jacob CO, Zhu J, Armstrong DL, Yan M, Han J, Zhou XJ, et al. Identification of IRAK1 as a risk gene with critical role in the pathogenesis of systemic lupus erythematosus. Proc Natl Acad Sci U S A 2009;106):6256–61.
11. Yang W, Zhao M, Hirankarn N, Lau CS, Mok CC, Chan TM, et al. ITGAM is associated with disease susceptibility and renal nephritis of systemic lupus erythematosus in Hong Kong Chinese and Thai. Hum Mol Genet 2009;18:2063–70.
12. Yang W, Shen N, Ye DQ, Liu Q, Zhang Y, Qian XX, et al. Genome-wide association study in Asian populations identifies variants in ETS1 and WDFY4 associated with systemic lupus erythematosus. PLoS Genet 2010;6:e1000841.
13. Han JW, Zheng HF, Cui Y, Sun LD, Ye DQ, Hu Z, et al. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. Nat Genet 2009;41:1234–7.
14. Yang J, Yang W, Hirankarn N, Ye DQ, Zhang Y, Pan HF, et al. ELF1 is associated with systemic lupus erythematosus in Asian populations. Hum Mol Genet 2011;20:601–7.
15. Yang W, Tang H, Zhang Y, Tang X, Zhang J, Sun L, et al. Meta-analysis followed by replication identifies loci in or near CDKN1B, TET3, CD80, DRAM1, and ARID5B as associated with systemic lupus erythematosus in Asians. Am J Hum Genet 2013;92:41–51.
16. Okada Y, Shimane K, Kochi Y, Tahira T, Suzuki A, Higasa K, et al. A genome-wide association study identified AFF1 as a susceptibility locus for systemic lupus eyrthematosus in Japanese. PLoS Genet 2012;8:e1002455.
17. Lee YH, Bae SC, Choi SJ, Ji JD, Song GG. Genome-wide pathway analysis of genome-wide association studies on systemic lupus erythematosus and rheumatoid arthritis. Mol Biol Rep 2012;39:10627–35.
18. Cunninghame Graham DS, Morris DL, Bhangale TR, Criswell LA, Syvanen AC, Ronnblom L, et al. Association of NCF2, IKZF1, IRF8, IFIH1, and TYK2 with systemic lupus erythematosus. PLoS Genet 2011;7:e1002341.
19. Lessard CJ, Adrianto I, Kelly JA, Kaufman KM, Grundahl KM, Adler A, et al. Identification of a systemic lupus erythematosus susceptibility locus at 11p13 between PDHX and CD44 in a multiethnic study. Am J Hum Genet 2011;88:83–91.
20. Lessard CJ, Adrianto I, Ice JA, Wiley GB, Kelly JA, Glenn SB, et al. Identification of IRF8, TMEM39A, and IKZF3-ZPBP2 as susceptibility loci for systemic lupus erythematosus in a large-scale multiracial replication study. Am J Hum Genet 2012;90: 648–60.
21. Deng Y, Zhao J, Sakurai D, Kaufman KM, Edberg JC, Kimberly RP, et al. MicroRNA-3148 modulates allelic expression of toll-like receptor 7 variant associated with systemic lupus erythematosus. PLoS Genet 2013;9:e1003336.
22. Kaufman KM, Zhao J, Kelly JA, Hughes T, Adler A, Sanchez E, et al. Fine mapping of Xq28: both MECP2 and IRAK1 contribute to risk for systemic lupus erythematosus in multiple ancestral groups. Ann Rheum Dis 2013;72:437–44.
23. Armstrong DL, Zidovetzki R, Alarcon-Riquelme ME, Tsao BP, Criswell LA, Kimberly RP, et al. GWAS identifies novel SLE susceptibility genes and explains the association of the HLA region. Genes Immun 2014;15:347–54.
24. Surakka I, Isaacs A, Karssen LC, Laurila PP, Middelberg RP, Tikkanen E, et al. A genome-wide screen for interactions reveals a new locus on 4p15 modifying the effect of waist-to-hip ratio on total cholesterol. PLoS Genet 2011;7:e1002333.
25. Neale BM, Sham PC. The future of association studies: gene-based analysis and replication. Am J Hum Genet 2004;75:353–62.
26. Ke X. Presence of multiple independent effects in risk loci of common complex human diseases. Am J Hum Genet 2012;91: 185–92.
27. Zhang J, Zhang Y, Yang J, Zhang L, Sun L, Pan HF, et al. Three SNPs in chromosome 11q23.3 are independently associated with systemic lupus erythematosus in Asians. Hum Mol Genet 2014;33:524−33.
28. Li MX, Gui HS, Kwan JS, Sham PC. GATES: a rapid and powerful gene-based association test using extended Simes procedure. Am J Hum Genet 2011;88:283–93.
29. Wu MC, Kraft P, Epstein MP, Taylor DM, Chanock SJ, Hunter DJ, et al. Powerful SNP-set analysis for case-control genome-wide association studies. Am J Hum Genet 2010;86:929–42.
30. Cheung CL, Sham PC, Xiao SM, Bow CH, Kung AW. Meta-analysis of gene-based genome-wide association studies of bone mineral density in Chinese and European subjects. Osteoporos Int 2012;23:131–42.

31. Curtis D, Vine AE, Knight J. A simple method for assessing the strength of evidence for association at the level of the whole gene. Adv Appl Bioinform Chem 2008;1:115–20.
32. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81:559–75.
33. Wong EH, So HC, Li M, Wang Q, Butler AW, Paul B, et al. Common variants on Xq28 conferring risk of schizophrenia in Han Chinese. Schizophre Bull 2014;40:777–86.
34. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25:1271–7.
35. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010; 26:2190–1.
36. Li MX, Sham PC, Cherny SS, Song YQ. A knowledge-based weighting framework to boost the power of genome-wide association studies. PloS One 2010;5:e14480.
37. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 1995;57:289–300.
38. Yang W, Ng P, Zhao M, Hirankarn N, Lau CS, Mok CC, et al. Population differences in SLE susceptibility genes: STAT4 and BLK, but not PXK, are associated with systemic lupus erythematosus in Hong Kong Chinese. Genes Immun 2009;10:219–26.
39. Zhang Y, Yang W, Mok CC, Chan TM, Wong RW, Mok MY, et al. Two missense variants in UHRF1BP1 are independently associated with systemic lupus erythematosus in Hong Kong Chinese. Genes Immun 2011;12:231–4.
40. Zhao H, Yang W, Qiu R, Li J, Xin Q, Wang X, et al. An intronic variant associated with systemic lupus erythematosus changes the binding affinity of Yinyang1 to downregulate WDFY4. Genes Immun 2012;13:536–42.
41. Agik S, Franek BS, Kumar AA, Kumabe M, Utset TO, Mikolaitis RA, et al. The autoimmune disease risk allele of UBE2L3 in African American patients with systemic lupus erythematosus: a recessive effect upon subphenotypes. J Rheumatol 2012; 39:73–8.
42. Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, et al. Genetics and beyond: the transcriptome of human monocytes and disease susceptibility. PloS One 2010;5:e10693.
43. Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A, et al. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. Nat Genet 2012;44:502–10.
44. Frade R, Gauffre A, Hermann J, Barel M. EBV/C3d receptor (CR2) interacts by its intracytoplasmic carboxy-terminal domain and two distinct binding sites with the p53 anti-oncoprotein and the p68 calcium-binding protein. J Immunol 1992;149:3232–8.
45. Asokan R, Hua J, Young KA, Gould HJ, Hannan JP, Kraus DM, et al. Characterization of human complement receptor type 2 (CR2/CD21) as a receptor for IFN-α: a potential role in systemic lupus erythematosus. J Immunol 2006;177:383–94.
46. Tsokos GC, Magrath IT, Balow JE. Epstein-Barr virus induces normal B cell responses but defective suppressor T cell responses in patients with systemic lupus erythematosus. J Immunol 1983; 131:1797–801.
47. Kang I, Quan T, Nolasco H, Park SH, Hong MS, Crouch J, et al. Defective control of latent Epstein-Barr virus infection in systemic lupus erythematosus. J Immunol 2004;172:1287–94.
48. Wu H, Boackle SA, Hanvivadhanakul P, Ulgiati D, Grossman JM, Lee Y, et al. Association of a common complement receptor 2 haplotype with increased risk of systemic lupus erythematosus. Proc Nat Acad Sci U S A 2007;104:3961–6.
49. Boackle SA, Holers VM, Chen X, Szakonyi G, Karp DR, Wakeland EK, et al. Cr2, a candidate gene in the murine Sle1c lupus susceptibility locus, encodes a dysfunctional protein. Immunity 2001;15:775–85.
50. Consortium EP, Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, et al. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57–74.
51. Hunt KA, Mistry V, Bockett NA, Ahmad T, Ban M, Barker JN, et al. Negligible impact of rare autoimmune-locus coding-region variants on missing heritability. Nature 2013;498:232–5.