Polymorphisms at 16p13 are associated with systemic lupus erythematosus in the Chinese population

Zheng Zhang,1,2 Yilin Cheng,1,2 Xueya Zhou,3 Yang Li,1,2 Jinping Gao,1,2 Jianwen Han,1,2 Cheng Quan,1,2 Sumin He,1,2 Yongmei Lv,1,2 Dayan Hu,1,2 Kunju Zhu,1,2 Liangdan Sun,1,2 Sen Yang,1,2 Xuejun Zhang1,2

ABSTRACT

Background Chromosomal region 16p13 has been reported to harbour variants associated with several autoimmune diseases, including type 1 diabetes, rheumatoid arthritis and multiple sclerosis.

Objective To test whether variants in the 16p13 region are also associated with systemic lupus erythematosus (SLE) by performing a candidate locus study in the Chinese Han population.

Methods Tag single nucleotide polymorphisms (SNPs) encompassing 50 kb upstream and downstream of the 250 kb linkage disequilibrium block, previously implicated in several autoimmune diseases, were analysed in 1047 patients with SLE and 1205 controls. The SNP showing the strongest association with SLE was then replicated in an independent cohort of 1643 cases and 5930 controls.

Results and conclusions The association between SNP rs12599402 and SLE reached the genome-wide significance level (p<5 × 10−8). The SNP was likely to tag the same functional variant as previously reported in European populations. The results suggested that the chromosomal region at 16p13 contains common susceptibility genes for different immune-mediated disorders.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a prototypic systemic autoimmune disease characterised by autoantibody production, complement activation, immune complex deposition and tissue and organ damage.1 The mechanism leading to SLE is the exaggerated B-cell response caused by activated T cells or dendritic cells, or soluble mediators such as cytokines; T cells or natural killer cells can modify the profile of cytokine generation.2 Tissue deposition of antibodies or immune complex induces inflammation and subsequent injury of multiple organs.

It was observed that different immune-mediated diseases often share considerable overlap between risk loci.3 Recently, polymorphisms in a 250 kb linkage disequilibrium (LD) block on chromosome 16p13 were reported to be associated with several autoimmune disorders including type 1 diabetes, multiple sclerosis and Crohn’s disease.4–7 In this study, we aimed to test if common variants within 16p13 were also associated with SLE susceptibility.

PATIENTS AND METHODS

We included 2690 patients with SLE and 7135 ethnically and geographically matched healthy controls in this study. All affected individuals defined as cases were diagnosed using the revised criteria for the classification of SLE from the American College of Rheumatology. Clinical information was collected from the affected individuals through a full clinical check-up by physician specialists. All controls were clinically assessed to be without SLE, other related autoimmune disorders, or systemic disorders. In addition, our control cohort excluded individuals who had first-, second-, or third-degree relatives affected with SLE, based on the information obtained from questionnaires. All participants provided written informed consent. The study was approved by the ethics committee of Anhui Medical University, and was conducted according to the Declaration of Helsinki principles. EDTA-anticoagulated venous blood samples were collected from all participants. Genomic DNA was extracted from peripheral blood lymphocytes by standard procedures using Flexi Gene DNA kits (Qiagen, Shanghai, China) and was diluted to a working concentration of 50 ng/μl for genome-wide genotyping, and 15–20 ng/μl for the validation study.

The candidate gene study was conducted on the case–control cohort recruited for a genome-wide association study (GWAS) as previously described.8 After quality control, 1047 cases and 1205 controls were used in this study. From the GWAS dataset, we extracted single nucleotide polymorphism (SNP) genotype data spanning from 50 kb upstream to 50 kb downstream of the 250 kb LD block at chromosome 16p13. The SNP displaying the strongest signal of association with SLE in this region was chosen for replication. All cases and controls in the replication stage were genotyped using the Sequenom MassARRAY technology (Sequenom, San Diego, California, USA), according to the manufacturer’s instructions.

Statistical analysis

Data management and association analysis were performed with PLINK 1.06 software.9 In both the GWAS and the replication stage, single-marker association analysis was performed using logistic regression with gender as a covariate. The joint analysis of all combined samples was performed using logistic regression with gender and sample cohorts as covariates. Haplotype analysis was done.
using two or three adjacent markers containing the top associated SNP. Expression quantitative trait association analysis was performed using linear regression assuming additive genotype effects.

Case-only analysis (eg, presence of photosensitivitiy versus no photosensitivity) was performed to identify which subphenotypes were associated with the specific genotype. We also conducted a case—control analysis (eg, presence of photosensitivity in patients versus healthy controls) to examine the risk in different subphenotypes of SLE.

RESULTS

We analysed a subset of SNPs residing in the candidate locus from our GWAS dataset. After quality control 90 SNPs were left, capturing more than 90% of the common SNPs (minor allele frequency > 0.05) in HapMap JPT + CHB panel with LD $r^2 > 0.7$. The strongest association signal was seen at SNP rs12599402 ($p = 1.4 \times 10^{-3}$, OR = 0.79 with 95% CI 0.68 to 0.91, protective allele C). After conditioning on rs12599402, no single SNP was significantly associated with SLE ($p > 0.01$), suggesting rs12599402 alone accounts for the association signals observed at all other sites (online supplementary table 1).

Haplotypes carrying different rs12599402 alleles were significantly associated with SLE (online supplementary table 2). So rs12599402 is not the causal variant itself, but an LD proxy thereof. Consistent with a single causal variant, haplotypic effects were no longer significant after controlling for the effect of rs12599402. Then we genotyped SNP rs12599402 in an independent cohort (1643 SLE cases and 5930 healthy controls) using the Sequenom MassARRAY system (Sequenom iPLEX Assay). Genotype counts did not deviate from Hardy–Weinberg equilibrium. An association between the minor G allele and SLE was confirmed ($p = 2.47 \times 10^{-9}$, OR = 0.81, 95% CI 0.75 to 0.88). A combined analysis of rs12599402 resulted in $p = 1.34 \times 10^{-8}$; which met the criterion of $p < 5 \times 10^{-8}$ for the significance of a genome-wide association.10 The minor allele G is protective, with a frequency of 0.40 in controls (table 1).

To assess the putative functional effect of SNP rs12599402, we tested the correlation between genotypes and the normalised expression level of nearby genes (obtained from the GeneVar database) in the HapMap CHB+JPT panel. We found that the minor allele of rs12599402 was associated with a decreased expression level of DEX1 ($p = 0.03$) and an increased level of CIITA ($p=0.04$), but no association was seen with the expression level of CLEC16A (online supplementary table 3).

Information about disease manifestations and genotype data for SNP rs12599402 were available for 2690 cases and 7135 controls. We performed a stratification analysis on age at diagnosis and the 11 American College of Rheumatology criteria. Most patients with SLE had an association with a high level of antinuclear antibody (90%), while only about 6.47% patients had an association with a neurological disorder. The risk allele showed consistent association when comparing patients of different subphenotypes with shared controls. We also did not observe any significant difference of risk allele frequencies when comparing positive and negative subphenotypes using different criteria. (online supplementary table 4). So the variant is not likely to increase the SLE risk by exerting its effect on specific subphenotypes.

DISCUSSION

In this study, we detected and confirmed a significant association of SNP rs12599402 in intron 19 of the CLEC16A gene with SLE in the Chinese Han population. Conditional regression and haplotype analysis indicated that there is one single causal variant in this region tagged by rs12599402. In their large-scale replication study, Gateva et al found a suggestive association signal in the same locus (rs12708716, $p = 5.6 \times 10^{-3}$, OR = 1.16, risk allele A).11 The same allele of SNP rs12708716 or its LD proxy (but not rs12599402-A) was previously reported to be associated with an increased risk of type 1 diabetes,12 13 14 Crohn’s disease,7 Addison’s disease15 and recently, rheumatoid arthritis.16 The SNP rs12708716 was not directly included in the Illumina 610 k chip. To assess its effect on SLE, we first imputed the genotypes on our GWAS cohort from phased HapMap data (using MACH 1.0) then performed logistic regression on allele dosages adding sex as a covariate. We found the same risk allele A of rs12708716 led to an increased SLE risk in the Chinese population, with a similar effect size estimate but weaker evidence ($p = 0.01$, OR = 1.24). It is likely that both SNPs tag the same causal variant. The fact that studies of different ethnicities resulted in different top-associated markers can be explained by different LD patterns. To support this, we found that rs12599402 and rs12708716 were highly correlated with each other in the HapMap CEU panel ($r^2 = 0.78$), but only moderately correlated in the JPT+CHB panel ($r^2=0.38$). The risk alleles (A allele of rs12599402 and A allele of rs12708716) are coupled on the most frequent two-locus haplotype in both populations. To see if we could use the different LD patterns to narrow the range of the causal variant, we considered the intersection of SNPs tagged by rs12599402 in Chinese and by rs12708716 in Europeans. In the HapMap data, SNP rs12599402 is correlated with 15 other SNPs with LD $r^2 > 0.7$ in the JPT +CHB panel; rs12708716 tags 55 at the same threshold, 14 of which are shared by the two populations. The 14 shared LD tags span a 100 kb genomic region within the CLEC16A gene, so extensive linkage disequilibrium in both Chinese and European populations at this region precludes us from pinpointing the causal variant. If the genetic effect of the causal variant is transferable across multiple populations, then candidate gene studies in populations with weaker LD, like Africans, can help refine the set of candidate-causal variants.17

Although the causal variant can be confined to its surrounding LD block, its functional effect may extend well beyond the block. Recently, there has been mounting evidence that subtle effects of genetic variation on gene expression are probably pathogenic mechanisms for complex diseases.16 17 With this in mind, we explored the regulatory effects of the top associated SNP with nearby genes using a public Expression Quantitative Trait Loci (eQTL) dataset generated from transformed B-cell lines of the HapMap CHB+JPT panel. We found that the protective allele of rs12599402 was associated with a decreased expression level of DEX1, though the associations were not significant after correcting for multiple tests. Nica et al generated an updated eQTL dataset for the HapMap 3 CEU panel of 109

Table 1 Association results for the single nucleotide polymorphism rs12599402

<table>
<thead>
<tr>
<th>GWAS*</th>
<th>Replication*</th>
<th>Combined†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case/control (counts)</td>
<td>1047/1205</td>
<td>1843/5930</td>
</tr>
<tr>
<td>SLE MAF (%)</td>
<td>36.44</td>
<td>35.04</td>
</tr>
<tr>
<td>Control MAF (%)</td>
<td>41.54</td>
<td>39.79</td>
</tr>
<tr>
<td>p Value</td>
<td>1.40 × 10^{-3}</td>
<td>2.47 × 10^{-6}</td>
</tr>
<tr>
<td>OR (95% CI)</td>
<td>0.79 (0.68 to 0.91)</td>
<td>0.81 (0.75 to 0.88)</td>
</tr>
</tbody>
</table>

*All p values and ORs (95% CI) were adjusted by sex in the GWAS and replication.
†The combined p value and OR (95% CI) were adjusted by sex and sample cohorts.
GWAS, genome-wide association study; MAF, minor allele frequency; SLE, systemic lupus erythematosus.
unrelated individuals.\textsuperscript{10} Their data (table 1 of reference 18) also supported the co-localisation of the functional variant tagged by rs12708716 in Europeans with the \textit{cis}-eQTL of the \textit{DEXI} gene. However, little is known about this gene. \textit{CIITA} is also a likely candidate, as its protein product acts as a positive regulator of class II major histocompatibility complex gene transcription. The absence of HLA class II expression can lead to severe immunodeficiency, known as bare lymphocyte syndrome type II.\textsuperscript{11} We observed that the risk allele could lower the expression level of \textit{CIITA}, and thereby may contribute to the susceptibility to autoimmune diseases.

It was believed that the \textit{CLEC16A} gene within this locus is the culprit gene, because it is highly expressed in immune cells and has a predicted domain that can interact with pathogen recognition proteins.\textsuperscript{6} However, the top associated SNP was not found to alter the \textit{CLEC16A} expression. Hakonarson \textit{et al} also tested the association between an LD proxy of rs12708716 and autoimmune diseases.\textsuperscript{6} We found no non-synonymous or splice variants in Europeans with the rs12708716 in Europeans with the \textit{DEXI} gene, supported the co-localisation of the functional variant tagged by DEXI. However, little is known about this gene.

In conclusion, this study added SLE to the list of diseases associated with polymorphisms at the chromosome 16p15 locus. The underlying candidate variant is likely to be shared between populations, and mediate the susceptibilities to more than one autoimmune disorder. Therefore, our results lend support to the idea of a shared genetic mechanism in immune-mediated diseases.\textsuperscript{20}

Acknowledgements

We thank all the people who have so willingly participated in this study. We also thank Dr Greg Votcher from Tsinghua University who proofread the manuscript.

Funding

This work was funded by the National Natural Science Foundation (30972727) and the Anhui Skin Genetic Study Innovative Research Team Program (TD200701).

Competing interests

None.

Patient consent

Obtained.

Ethics approval

This study was conducted with the approval of the ethics committee of Anhui Medical University, and according to the Declaration of Helsinki principles.

Provenance and peer review

Not commissioned; externally peer reviewed.

REFERENCES


9. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PJ, Daly MJ, Shamu C, Pucknick A toolset for whole-genome association and population-based linkage analyses. \textit{Am J Hum Genet} 2007;80:559–75.


