Association of the *IKZF1* 5’ UTR variant rs1456896 with lupus nephritis in a northern Han Chinese population

Y-M Zhang*, X-J Zhou*, F-J Cheng, Y-Y Qi, P Hou, M-H Zhao, H Zhang

Renal Division, Peking University First Hospital, Peking University Institute of Nephrology, Key Laboratory of Renal Disease, Ministry of Health of China, and Key Laboratory of Chronic Kidney Disease Prevention and Treatment, Peking University, Ministry of Education, Beijing, China

**Objectives:** Polymorphisms of IKAROS family zinc finger 1 (*IKZF1*) have been found to be associated with systemic lupus erythematosus (SLE) by genome-wide association studies (GWAS). The aim of the current study was to investigate the association between *IKZF1* functional variants and lupus nephritis (LN) in a northern Han Chinese population and analyse their relationship with clinical and pathological phenotypes in LN.

**Method:** The association between *IKZF1* functional variants and LN was analysed for the lead variant rs1456896 with both GWAS and expression quantitative trait loci (eQTL) top hits in 500 LN patients and 500 healthy controls. Replication was conducted in an independent cohort comprising 798 LN patients and 704 healthy controls. Using the ENCODE (Encyclopedia of DNA Elements) databases, functional annotations and differential gene expression data were evaluated.

**Results:** A significant association between the single nuclear polymorphism (SNP) rs1456896 and susceptibility to LN was observed in the two different cohorts (p = 9.32 × 10⁻³ and p = 3.00 × 10⁻²) and reinforced in combination (p = 1.36 × 10⁻³). In silico analysis indicates that rs1456896 is a regulatory variant and lower mRNA expressions of *IKZF1* were observed in both peripheral blood mononuclear cells (PBMCs) and renal biopsies from SLE patients compared to normal controls. Although patients with the protective genotype AA of rs1456896 seemed to have more pronounced clinical manifestations and a lower ratio of histological classes III and IV, no significant associations between rs1456896 genotypes and sub-phenotypes of LN were detected.

**Conclusions:** Our results suggest that the rs1456896 A allele is associated with protective susceptibility to LN. However, this association did not seem to be implicated in the disease and histopathological severity of LN in the current population.
associations are needed, and such endeavours should be centred on both the association signal and the function signal. In the present study, we focused on a lead variant with both GWAS and expression quantitative trait loci (eQTL) top hits and its associations with LN.

**Method**

**Study population**

To identify the association of *IKZF1* variants with LN, 500 LN patients (mean age 31.9 ± 11.2 years; 423 females) and 500 healthy blood donors (mean age 40.0 ± 8.6 years; 140 females), who were of Han Chinese origin in Beijing, were enrolled from the renal division of Peking University First Hospital. For comparison, a second population containing 798 LN patients (mean age 32.8 ± 12.2 years; 672 females) and 704 healthy blood donors (mean age 32.7 ± 8.9 years; 219 females) were enrolled. All the patients with LN met the American College of Rheumatology (ACR) revised criteria for SLE (14) and diagnoses were confirmed by renal biopsy using light microscopy, immunofluorescence, and electron microscopy.

The ethics review committee of Peking University First Hospital approved the study and all participants provided written informed consent.

**SNP selection and genotyping**

For SNP selection, a region spanning about 200 kb (chr7: 50304461–50492271, hg19) encompassing *IKZF1* as the RefSeq gene and its regulatory regions was searched in the updated HaploReg database (HaploReg v4.1, www.broadinstitute.org/mammals/haploreg/haploreg.php). Among the 2041 SNPs within this region, 15 variants had NHGRI-EBI GWAS hits, 72 variants had GRASP QTL hits, and 13 variants had both NHGRI-EBI GWAS hits and GRASP QTL hits. Among these 13 SNPs, rs1456896 had the most GRASP QTL hits (321 hits). rs1456896 was previously reported to be the top signal associated with inflammatory bowel disease (IBD) by GWAS (15, 16), suggesting it may be a lead variant within the region with regard to both genetic association and functional significance. Thus, rs1456896 was selected and genotyped by TaqMan allele discrimination assays (Applied Biosystems, Foster City, CA, USA) to search for the association between *IKZF1* variants and LN. A flow chart of the genetic analysis of *IKZF1* functional variants is shown in Figure 1.

**Bioinformatic analysis**

The variant effect of rs1456896 on regulatory motifs was investigated using the HaploReg v4.1 database and was quantified as the difference of logarithms of odds scores (LOD(alternative)–LOD(reference). A negative score suggested a relatively higher affinity for the reference sequence while a positive score indicated a relatively higher affinity for the alternative. DNA features and regulatory elements of the region that contains rs1456896 were identified using the RegulomeDB database (http://regulome.stanford.edu/), and the scoring standard for RegulomeDB is provided in Supplementary Table S1. In addition, regulatory annotations on rs1456896 were searched in the rSNPBase database (http://rsnp.psych.ac.cn/).

To explore whether rs1456896 had an eQTL effect, the gene expression profiling of Epstein–Barr virus (EBV)-transformed lymphoblastoid cell lines of 160 unrelated Asian individuals (HapMap3) from the Gene Expression Variation project (GENEVAR project, www.sanger.ac.uk/humgen/genevar/) was used. In addition, the gene differential expression analysis of *IKZF1* was performed in SLE patients and compared with healthy controls from large-scale genome-wide gene expression analyses conducted in peripheral blood mononuclear cells (PBMCs) (E-GEOD-50772) and renal biopsies (E-GEOD-32591) by using the ArrayExpress Archive database (www.ebi.ac.uk/arrayexpress/).

**Statistical analyses**

The genotype frequency of rs1456896 was tested for Hardy–Weinberg equilibrium separately in cases and
controls. Allele frequencies were compared between LN cases and controls by using the χ² test, and logistic regression analysis adjusted by sex and age was adopted to control the age and gender bias. In addition, the association of different degrees of severity and outcomes of LN with different genotypes of rs1456896 was determined. The results of the measurements are expressed as mean ± standard deviation (sd) and the χ² test or one-way analysis of variance (ANOVA) was used. Statistical analyses were performed using SPSS version 16.0 (SPSS Inc, Chicago, IL, USA). A two-tailed p value of less than 0.05 was considered statistically significant.

Results

Association of rs1456896 with LN

As shown in Table 1, the rs1456896 minor allele A was found to be significantly associated with LN in our discovery population [odds ratio (OR) 0.69, 95% confidence interval (CI) 0.52–0.91, p = 9.32 × 10⁻³]. The association was further replicated in another Chinese population (OR 0.84, 95% CI 0.72–0.98, p = 3.00 × 10⁻³) and was enhanced with combined data (OR 0.80, 95% CI 0.70–0.92, p = 1.36 × 10⁻³). Logistic regression analyses adjusted for sex and age also suggested that the protective allele A of rs1456896 was significantly associated with LN compared to controls (in the discovery population: OR 0.66, 95% CI 0.49–0.89, p = 6.94 × 10⁻³; in the replication population: OR 0.77, 95% CI 0.60–1.00, p = 4.97 × 10⁻²; in the combined population: OR 0.75, 95% CI 0.62–0.90, p = 2.32 × 10⁻³).

Furthermore, associations between different genotypes of rs1456896 and disease severity and outcomes, including age of onset, gender, proteinuria, estimated glomerular filtration rate (eGFR), serum creatinine level, C3 level, the SLE Disease Activity Index (SLEDAI) score, percentage of crescent, different histological types, response to treatment, and development of end-stage renal disease (ESRD), were detected in 279 LN patients who were routinely followed up for at least 1 year, with a mean follow-up time of 56.58 months. Among these patients, response to treatment was defined as a urine protein level that decreased by 50% or was below 1 g/24 h after treatment, and the development of ESRD was defined as dialysis or death. As shown in Supplementary Table S2, the patients with the protective genotype AA of rs1456896 seemed to have later onset, a smaller male percentage, lower proteinuria levels, higher eGFR levels, lower serum creatinine levels, lower SLEDAI scores, a lower ratio of histological classes III and IV, and a higher treatment remission rate. However, no significant associations between rs1456896 genotypes and sub-phenotypes of LN were detected.

Regulatory effects prediction

The HaploReg v4.1 database predicted a binding site motif that spans the 5' untranslated region (UTR) of the IKZF1 region for binding by the transcription factor hepatocyte nuclear factor (HNF)4. The difference between the LOD scores for A/G alleles (reference) was −10.1, predicting a lower affinity of HNF4 to the regulatory motif for the variant A (protective) allele relative to the G allele. We further located the set of SNPs in strong linkage disequilibrium (LD) (r² ≥ 0.8) with rs1456896 in an Asian population, and among the 21 SNPs, six were located in regions of promotor histone marks, 11 were in regions with enhancer histone marks, 12 were in DNAse regions, three were in protein binding regions, 17 were in motif changed regions, and four had NHGRI/EBI GWAS hits and/or GRASP QTL hits (Supplementary Table S3). In the RegulomeDB database, rs1456896 scored 4 (TF binding + DNase peak) and was predicted to be within the regions of protein binding, chromatin structure, and histone modifications. In the rSNPBase database, the eQTL effect of rs1456896 has been detected. Furthermore, five SNPs in strong LD (r² > 0.8) with rs1456896 were regarded as regulatory SNPs (rSNPs): one variant was involved in proximal and distal transcriptional regulation, one variant was involved in distal transcriptional regulation and showed an eQTL effect, while the other three variants were involved in distal transcriptional regulation. The evidence above suggests that rs1456896 is located in a regulatory region.

Gene expression analysis

To validate the cis-eQTL effect of rs1456896, the gene expression data of an Asian population (n = 160) in the

Table 1. Association of the SNP rs1456896 in the IKZF1 gene* with lupus nephritis (LN) in a Han Chinese population.

|                      | Discovery population | Replication population | Total population |
|----------------------|----------------------|------------------------|------------------|
| No. of cases/controls| 500/500              | 798/704                | 1298/1204        |
| MAF (%) (cases/controls) | 25.45/33.17         | 28.76/32.43            | 27.89/32.59      |
| p (allele model)     | 9.32 × 10⁻³         | 3.00 × 10⁻²            | 1.36 × 10⁻³      |
| OR (95% CI)          | 0.69 (0.52–0.91)    | 0.84 (0.72–0.98)       | 0.80 (0.70–0.92) |

SNP, single nucleotide polymorphism; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval.

*Chromosome 7; position hg19; minor allele A.

None of the genotypes in the controls or patients showed significant deviation from Hardy–Weinberg equilibrium.

Logistic regression analyses adjusted by sex and age also suggested that the protective allele A of rs1456896 was significantly associated with LN compared to controls.

www.scandrheumatol.dk
GENEVAR project was used. As shown in Supplementary Table S4, although the rs1456896 A allele seemed to be associated with a higher expression level of IKZF1, no significant association was identified.

We also ascertained whether IKZF1 was expressed differently in SLE patients and healthy controls using the ArrayExpress Archive database. We observed that IKZF1 mRNA expression was significantly down-regulated in SLE PBMCs (mean ± sd normalized fluorescence intensity 6784.31 ± 1334.21 among 61 SLE patients vs. 9040.20 ± 773.33 among 20 normal donor controls, \(p = 2.85 \times 10^{-4}\)) and tubulointerstitial samples for LN (4.17 ± 0.10 among 32 LN patients vs. 4.28 ± 0.14 among 15 pre-transplant living donor controls, \(p = 5.00 \times 10^{-3}\)) (Figure 2).

![Figure 2. Gene differential expression analyses of IKZF1 in systemic lupus erythematosus (SLE) patients and controls.](image)

**Discussion**

In the present study, the association between rs1456896 of IKZF1 and LN has been detected and replicated in an independent population. Various regulatory effects of rs1456896 were predicted and a significant lower expression level of IKZF1 has been detected in PBMCs and renal biopsy samples of SLE patients compared with controls, providing important indications for the role of IKZF1 in LN. The results of recent genetic studies have largely supported the notion of shared genetics in immune-related diseases, including SLE (17–19). The polymorphisms of IKZF1 have been reported to be associated with several immune-related diseases, such as IBD (15, 16), type 1 diabetes (T1DM) (20), childhood acute lymphoblastic leukaemia (C-ALL) (21), and SLE (10, 12, 13). IKZF1 encodes a transcription factor that belongs to the family of zinc-finger DNA-binding proteins associated with chromatin remodeling. It is reported to be involved in the pathogenesis of autoimmune diseases by regulating lymphocyte differentiation. SLE is a prototypic, systemic autoimmune disease that has a strong genetic component. There were studies suggested that Ikaros could regulate the transcription of STAT4, which was reported to be associated with SLE by several GWAS, by combining with its 5’ sequence (22, 23). In addition, interferon (IFN) is thought to be closely associated with SLE, especially with regard to skin and renal involvement. In a mouse model, the low expression level of IKZF1 was attributed to the induced number of plasmacytoid dendritic cells (pDCs) and IFN production (24). Ikaros was found to be essential for DC activation and T-helper (Th1)-cell differentiation, and IKZF1 might play an important role in IFN-\(\gamma\) production by inhibiting T-bet, a special transcription factor for Th1 cells in IFN-producing regulation (25).

In the present study, a significant difference was detected with the protective allele (A allele) of the SNP rs1456896 located in the 5’ UTR of IKZF1. As SLE is widely regarded as a polygenic disease, the functional variant rs1456896 in IKZF1 is likely to play some role in SLE/LN but is probably not the only causal factor in the disease. Thus, we failed to observe significant differences regarding clinical findings, pathological features, and renal outcomes in our LN patients (n = 279) with different genotypes of rs1456896. In in-silico analysis, the regulatory effects of rs1456896 have been predicted. In addition, individuals with the rs1456896 AA genotype seemed to have a higher expression level of IKZF1 in an Asian population in the GENEVAR project, although no significant association has been detected, which may be due to the small sample size and low allele frequency of rs1456896 #A allele. Consistent with the previous report conducted in Chinese SLE patients (26), lower IKZF1 expression levels have been observed in PBMC and tubulointerstitial samples of SLE patients by genome-wide gene expression analyses. These bioinformatics analyses may shed more light on the role of IKZF1 in the pathogenesis of LN. However, more widespread replications with larger populations and functional studies are needed in the future.

In summary, a relationship between the rs1456896 A allele and susceptible protection of LN was established in our study, which may be useful for improvements in the understanding of the pathogenesis of the disease.
Acknowledgements

This work was supported by grants from the Beijing Natural Science Foundation (7152148), the National Natural Science Foundation of China (81570629), the Chinese Society of Nephrology (150200030591), the Beijing Municipal Science and Technology for the Outstanding PhD Program (2012000110), the National Science Fund of China to the Innovation Research Group (81021004), and the National Key Technology R&D Programme (2013BAI09B14).

References

1. Lim LSH, Lee SJ, Feldman BM, Gladman DD, Pullenayegum E, Uerly E, et al. A systematic review of the quality of prognosis studies in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2014;66:1536–41.
2. Mohan C, Puttermann C. Genetics and pathogenesis of systemic lupus erythematosus and lupus nephritis. Nat Rev Nephrol 2015;11:329–41.
3. Alarcón-Segovia D, Alarcón-Riquelme ME, Cardiel MH, Caeiro F, Massardo L, Villa AR, et al. Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort. Arthritis Rheum 2005;52:1138–47.
4. Kuo CF, Grainge MJ, Valdes AM, See LC, Luo SF, Yu KH, et al. Familial aggregation of systemic lupus erythematosus and coaggregation of autoimmune diseases in affected families. JAMA Intern Med 2015;175:1518–26.
5. Deapen D, Escalante A, Weinrib L, Horwitz D, Bachman B, Roy-Barnman P, et al. A revised estimate of twin concordance in systemic lupus erythematosus. Arthritis Rheum 1992;35:311–18.
6. Wang JH, Nichogiannopoulou A, Wu L, Sun L, Sharpe AH, Bigby M, et al. Selective defects in the development of the fetal and adult lymphoid system in mice with an Ikaros null mutation. Immunity 1996;5:537–49.
7. Kirstetter P, Thomas M, Dierich A, Kastner P, Chan S. Ikaros is critical for B cell differentiation and function. Eur J Immunol 2002;32:720–30.
8. Winandy S, Wu L, Wang JH, Georgopoulos K. Pre-T cell receptor (TCR) and TCR-controlled checkpoints in T cell differentiation are set by Ikaros. J Exp Med 1999;190:1039–48.
9. Wojcik H, Griffiths E, Stagg S, Hagman J, Winandy S. Expression of a non-DNA-binding Ikaros isoform exclusively in B cells leads to autoimmunity but not leukemogenesis. Eur J Immunol 2007;37:1022–32.
10. 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.
11. He CF, Liu YS, Cheng YL, Gao JP, Pan TM, Han JW, et al. TNIP1, SLC15A4, ETS1, RasGRP3 and IKZF1 are associated with clinical features of systemic lupus erythematosus in a Chinese Han population. Lupus 2010;19:1181–6.
12. Cunnininghame Graham DS, Morris DL, Bhangale TR, Criswell LA, Syvanen AC, Ronnblom L, et al. Association of nec2, ikaros1, ifih1, ifih1, and tyk2 with systemic lupus erythematosus. PLoS Genet 2011;7:e1002341.
13. Wang C, Ahlford A, Järvinen TM, Nordmark G, Eloranta ML, Gunnarsson I, et al. Genes identified in Asian SLE GWASs are also associated with SLE in Caucasian populations. Eur J Hum Genet 2013;21:994–9.
14. 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.
15. Franke A, McGovern DPB, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn’s disease susceptibility loci. Nat Genet 2010;42:1118–25.
16. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 2012;491:119–24.
17. Hom G, Graham RR, Modrek B, Taylor KE, Orthmann W, Garnier S, et al. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. N Engl J Med 2008;358:900–9.
18. Zhou XJ, Nath SK, Qi YY, Cheng FJ, Yang HZ, Zhang Y, et al. Brief Report: identification of MTRM3 as a novel susceptibility gene for lupus nephritis in northern Han Chinese by shared-gene analysis with IgA nephropathy. Arthritis Rheumatol 2014;66:2842–8.
19. Zhou XJ, Cheng FJ, Zhu L, Lv JC, Qi Y, Hou P, et al. Association of systemic lupus erythematosus susceptibility genes with IgA nephropathy in a Chinese cohort. Clin J Am Soc Nephrol 2014;9:788–97.
20. Swafford AD, Howson JMM, Davison LJ, Wallace C, Smyth DJ, Schuilenburg H, et al. An allele of IKZF1 (Ikaros) conferring susceptibility to childhood acute lymphoblastic leukemia protects against type 1 diabetes. Diabetes 2011;60:1041–4.
21. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Oliver B, Sheridan E, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. Nat Genet 2009;41:1006–10.
22. Yap WH, Yeoh E, Tay A, Brenner S, Venkatesh B. STATA4 is a target of the hematopoietic zinc-finger transcription factor Ikaros in T cells. FEBS Lett 2005;579:4470–8.
23. Good SR, Thiue VT, Mathur AN, Yu Q, Suttersky GL, Yeh N, et al. Temporal induction pattern of STAT4 target genes defines potential for Th1 lineage-specific programming. J Immunol 2009;183:3839–47.
24. Allman D, Dalod M, Asselin-Paturel C, Delale T, Robbins SH, Schuilenburg H, et al. An allele of IKZF1 (Ikaros) conferring susceptibility to childhood acute lymphoblastic leukemia protects against type 1 diabetes. Diabetes 2011;60:1041–4.
25. Thomas RM, Chen C, Chundra N, Ma L, Taylor J, Pearce EL, et al. Ikaros silences T-bet expression and interferon-gamma production during T helper 2 differentiation. J Biol Chem 2010;285:2455–53.
26. Hu W, Sun L, Gao J, Li Y, Wang P, Cheng Y, et al. Down-regulated expression of IKZF1 mRNA in peripheral blood mononuclear cells from patients with systemic lupus erythematosus. Rheumatol Int 2011;31:819–22.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Supplementary Table S1. The scoring standard of the RegulomeDB database.
Supplementary Table S2. Association analyses of rs1456896 genotypes and disease severity and outcomes of patients with lupus nephritis (LN).
Supplementary Table S3. HaploReg v4.1 annotations for variants in strong linkage disequilibrium with rs1456896 in an Asian population.
Supplementary Table S4. Correlation analyses of rs1456896 genotypes and IKZF1 gene expression using gene expression profiling of EBV-transformed lymphoblastoid cell lines from Asian population in HapMap3.

Please note that the editors are not responsible for the content or functionality of any supplementary material supplied by the authors. Any queries should be directed to the corresponding author.