ELF1 is associated with systemic lupus erythematosus in Asian populations

Jing Yang1,†, Wanling Yang1,†, Natiya Hirankarn6, Dong Qing Ye9, Yan Zhang1, Hai-Feng Pan9, Chi Chiu Mok10, Tak Mao Chan2, Raymond Woon Sing Wong2, Mo Yin Mok2, Ka Wing Lee12, Sik Nin Wong11, Alexander Moon Ho Leung13, Xiang-Pei Li14, Yingyos Avihingsanon7, Pornpimol Rianthavorn8, Thavatchai Deekajorndej8, Kanya Suphapeetiporn8, Vorasuk Shotelersuk8, Larry Baum15, Patrick Kwan16, Tsz Leung Lee1, Marco Hok Kung Ho1, Pamela Pui Wah Lee1, Wilfred Hing Sang Wong1, Shuai Zeng1, Jing Zhang1, Chun-Ming Wong3, Irene Oi Lin Ng3, Maria-Mercè Garcia-Barceló4, Stacey S. Cherny5, Paul Kwong-Hang Tam4, Pak Chung Sham5, Chak Sing Lau2 and Yu Lung Lau1,*

†These authors contributed equally to this work.

∗To whom correspondence should be addressed at: Department of Paediatrics and Adolescent Medicine, Li Ka Shing Faculty of Medicine, Queen Mary Hospital, The University of Hong Kong, Hong Kong, China. Tel: +852 28554481; Fax: +852 28551123; Email: lauylung@hkucc.hku.hk

Advance Access published on November 2, 2010

Systemic lupus erythematosus (SLE) is an autoimmune disease with a strong genetic involvement. The susceptibility genes identified so far can only explain a small proportion of disease heritability. Through a genome-wide association in a Hong Kong Chinese cohort and subsequent replication in two other Asian populations, with a total of 3164 patients and 4482 matched controls, we identified association of ELF1 (E74-like factor 1) with SLE (rs7329174, OR 1.26, joint P = 1.47 × 10^{-8}). ELF1 belongs to the ETS family of transcription factors and is known to be involved in T cell development and function. Database analysis revealed transcripts making use of three alternative exons for this gene. Near equivalent expression levels of distinct transcripts initiated from alternative exons were detected in peripheral blood mononuclear cells from both SLE patients and healthy controls. Although a direct association of rs7329174 with the three forms of transcripts for this gene was not detected, these findings support an important role of ELF1 in SLE susceptibility and suggest a potentially tight regulation for the expression of this gene.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a prototype autoimmune disease characterized by autoantibody production and multiorgan damage. It mainly affects women of childbearing age and has population differences in both disease prevalence and severity (1,2). Genetic factors are known to
play key roles in the disease, with ~30 times increased risk for siblings of those affected, and further increased risk for monozygotic twins (3–5).

Through genome-wide association studies (GWAS) and other family- or population-based studies, a number of SLE susceptibility genes have been established through replications across populations. For example, after the initial report on the association of IRF5 with the disease (6), its association was then confirmed in other European populations (7–9) and in Asian populations (10–12). The association of STAT4 with SLE initially identified in Caucasians (13) was quickly confirmed in other populations such as Asians and Colombians (14–17). Association of ITGAM to SLE was confirmed in Chinese, although the risk alleles have very low frequency (18). Interestingly, two simultaneous Asian GWAS on SLE, including our own (19,20), identified association of ETS1 and WDFY4 with the disease, which had not been detected in earlier studies based on the populations of European ancestry, raising a possibility of population differences for disease association for these loci. PXK was found to be associated with SLE in Caucasians (21,22) but not in Asians (14,23).

ELF1 (E74-like factor 1) belongs to the ETS family of transcription factors that regulate the expression of a wide range of genes and play important roles in hematopoiesis, immune cell development and function, and angiogenesis (24,25). It acts as both an enhancer and a repressor to regulate the transcription of a variety of genes (26,27). Previously, we identified ETS1, a prototype ETS family transcription factor, as associated with SLE in Asian populations (19). Here, we report association of ELF1 with SLE in Chinese populations living in Hong Kong and Mainland, China, respectively, and a Thai population living in Bangkok. Database analysis and subsequent RT–PCR experiment identified interesting features of this gene, suggestive of tight regulation of its expression.

RESULTS

Table 1 shows SNPs in and around the ELF1 gene on chromosome 13q10 with significant disease association (P < 0.005) from the GWAS stage, which includes 612 SLE patients of Hong Kong Chinese and 1160 matched controls. This GWAS partially overlaps with our previously reported GWAS finding ETS1 and WDFY4 associated with SLE (19), but reflecting a doubling of SLE samples genotyped by the same platform. We observe a genomic inflation factor of 1.045 and a good match between cases and controls of the Hong Kong samples, analyzed by principal component introduced in EIGEN-STRAT (28), similar to what reported before (19). SNP rs7329174 showed the most significant association with the disease, and conditional logistic regression did not present convincing evidence for any independent contributions from other SNPs in this region, although this should be viewed with caution since tests on independence are usually under powered. Figure 1 shows linkage disequilibrium (LD) among these SNPs, ranging from low to moderate LD (r² = 0.06–0.44).

We have chosen rs7329174 for further replication in the remaining samples from Hong Kong and samples from another Chinese population in Anhui, central China, as well as samples from Bangkok, Thailand. Association analysis was done by comparing cases with controls matched geographically. Although we observed a smaller effect size and marginal significance in the Anhui samples and the samples from Bangkok, Thailand, nevertheless the same trend was observed in all three populations (Table 2). A marginal P-value was observed in a heterogeneity test of different ORs in the three cohorts (by the Breslow–Day test, P = 0.09), suggesting only minor differences, if any.

A marginal significant association with lupus nephritis was also observed for rs7329174 in a patient-only test, but the effect is only significant for the Hong Kong samples (OR = 1.25, P = 0.02). In addition to renal nephritis, we also saw a significant association of this SNP in a patient-only analysis to dsDNA antibodies production (OR = 1.27, P = 0.02). Analysis of other subphenotypes did not find significant differences through the patient-only analysis. A note of caution is that none of the subphenotype associations would survive correction by the number of subphenotypes examined and more studies would be needed to establish connections between subphenotypes for this disease and this locus.

In an effort to delineate the potential functional role of rs7329174, we performed bioinformatics analysis of databases from NCBI for transcripts of ELF1, especially focusing on the region where rs7329174 is located. According to Entrez_gene database, there are two alternative variants for the ELF1 gene (http://www.ncbi.nlm.nih.gov/gene/1997). Transcript NM_172373.3 contains an exon1 (shown as E1B in Fig. 1) that is 37 kb upstream of a constitutive exon2. NM_001145353.1 lacks exon1 completely, with otherwise the same sequence as NM_172373.3. We searched the expressed sequence tags (ESTs) and refseq_rna databases from NCBI for transcripts for this gene and did not find any supporting evidence for the transcription initiation reflected by NM_001145353.1. Instead, we identified transcripts containing two other alternative first exons in addition to finding support for the exon1 represented by NM_172373.3. One of these alternative first exons, shown as E1C in Figure 1, locates 570 bp upstream of the constitutive exon2 and ~1 kb downstream of SNP rs7329174. Another alternative first exon, E1A, is ~80 kb upstream of the constitutive exon2 and is close to WBP4 gene, which is on the other strand of the chromosome. All three alternative first exons are supported by multiple EST sequences derived from different tissues and cDNA libraries. Examples of these transcripts include, for example, BX640798 and CN312731 for E1A and CD522023 and DC296627 for E1C. According to the EST entries, the three alternative first exons seem to express with a similar prevalence.

Using peripheral blood mononuclear cell (PBMC) from SLE patients and healthy blood donors, we confirmed the prevalent expression of the three alternative exon1s (Fig. 1B). However, we did not find correlation between these alternative transcription initiations represented by these distinct exon1s and the rs7329174 genotypes. For the samples we have examined (~30 each in both cases and controls), transcripts containing the three alternative exon1s seem to have a similar expression level (Fig. 1B), although the sample size is considered small given the fact that expression experiments are usually affected by both experimental and individual variations.
Table 1. SNPs in and around ELF1 in GWAS (612 cases and 1160 controls) that showed significant association with SLE

<table>
<thead>
<tr>
<th>SNP</th>
<th>Position on chromosome 13 (NCBI Build 36)</th>
<th>A1 F_A</th>
<th>F_U</th>
<th>A2 OR (95% CI)</th>
<th>P-values</th>
<th>Logistic regression P-values</th>
<th>EIGAN-STRAT-corrected P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs7329174</td>
<td>40 456 110</td>
<td>G</td>
<td>0.283</td>
<td>0.219 A</td>
<td>1.41 (1.22–1.63)</td>
<td>7.18 × 10^{-6} N/A</td>
<td>2.25 × 10^{-5}</td>
</tr>
<tr>
<td>rs10507489</td>
<td>40 462 422</td>
<td>A</td>
<td>0.118</td>
<td>0.194 G</td>
<td>0.78 (0.66–0.92)</td>
<td>3.79 × 10^{-3} 0.093 (1.21 × 10^{-5})</td>
<td>0.00226</td>
</tr>
<tr>
<td>rs4942016</td>
<td>40 492 060</td>
<td>T</td>
<td>0.391</td>
<td>0.441 C</td>
<td>0.81 (0.71–0.92)</td>
<td>1.49 × 10^{-3} 0.29 (8.10 × 10^{-7})</td>
<td>0.00046</td>
</tr>
<tr>
<td>rs794329</td>
<td>40 532 954</td>
<td>A</td>
<td>0.243</td>
<td>0.190 G</td>
<td>1.37 (1.18–1.59)</td>
<td>4.77 × 10^{-5} 0.34 (0.002)</td>
<td>5.16 × 10^{-5}</td>
</tr>
<tr>
<td>rs2772179</td>
<td>40 560 028</td>
<td>C</td>
<td>0.424</td>
<td>0.471 T</td>
<td>0.83 (0.73–0.94)</td>
<td>3.37 × 10^{-3} 0.10 (1.75 × 10^{-5})</td>
<td>0.00077</td>
</tr>
</tbody>
</table>

A1, minor allele; A2, major allele; F_A, minor allele frequency in the affected; F_U, minor allele frequency in the unaffected.

*P-values by conditional logistic regression when controlling for the effect of rs7329174 (P-values for rs7329174 when controlling for the effect of this SNP in the same conditional logistic regression test).

**Association P-values on allelic test after EIGAN-STRAT correction on population substructure using the principal component method.

Figure 1. LD among SNPs in ELF1 associated with SLE and transcription of the gene from three alternative exon1s. (A) SNPs showed significant association in the GWAS stage, their LD to each other and their relative position to ELF1. Shown are also the three alternative exon1s designated as E1A, E1B and E1C for this gene. SNP rs7329174 is about 1 kb upstream of E1C. LD among the SNPs was calculated by Haploview and shown are the r² value between the SNPs. (B) Confirmation of the expression of all three alternative exon1s of this gene by RT–PCR from PBMC of both patients and controls. Shown are representative results from a control and case, respectively.

DISCUSSION

SNP rs7329174 has the highest G allele frequency in Asians, 23.3 and 30.7% in Chinese and Japanese, respectively, according to HapMap. It also has high minor allele frequency in Gujarati Indians in Houston, a HapMap phase III population (16.5% on G allele). However, it has low allele frequencies in other populations and is near monomorphic A in Caucasians (for HapMap CEPH data, there are 164 AA genotypes and 1 AG genotype). Neighboring SNPs that are in LD with rs7329174 (r² > 0.5), such as rs9532692, rs7335629, rs4264266, rs17593261 and rs3794329, are either monomorphic or have extremely low minor allele frequencies (<1%) in HapMap CEPH data. The allele frequency difference raises a possibility of differences in disease associations for this locus between populations. The major allele in humans, the ‘A’ allele, was observed in the same position in both Rhesus Macaque and Chimpanzee (Fig. 2), suggesting a possibility that the risk allele ‘G’ may be gained in humans.

Regulation of ELF1 gene expression seems a plausible mechanism, should SNP rs7329174 be the functional variant itself. Analysis of a database containing both gene expression and SNP genotyping data for the HapMap individuals did not show any correlation between rs7329174 genotype or its surrogates and the ELF1 expression level (29,30) (data not shown). The finding of the three equally used exon1s, however, does indicate that the expression of this gene is probably under the complex, multilayer regulation.

In mouse, ELF1 is expressed at high levels throughout thymocyte development and in all subsets of maturing thymocytes and T cells. In addition, high-level ELF1 expression was also observed in B lymphocytes and macrophages (31). Elf-1 binds and activates or represses a list of genes important to T cell development and function, such as CD4, which plays a central role in the selection, differentiation, survival and activation of Th cells (32,33). Elf-1 is also known to increase the promoter activity of CD3-Delta (34), TdT (35), IL3 (36) and GM-CSF (37). In B lymphocytes, Elf-1 binds to a B cell-specific regulatory element in the IgH enhancer and activates murine IgH expression (38). It is also thought to be the major regulator of the Ig alpha gene promoter (39) as well as the gene for Fc receptor for IgA (FCAR) (40).

IL-2 abnormality plays a central role in the activity and function of T cells in SLE. SLE patients were found to have decreased levels of IL-2 (41,42), as found in lupus mouse models (43–45). Elf-1 was known to bind to an IL2 enhancer and regulate its expression (26). Essential positive regulatory element was also identified in the promoter region of IL2Rα, which is regulated by Elf-1. Thus, Elf-1 may play a role in IL-2 production as well as its responses (46–48).
Elf-1 may also play a role in T cell receptor signaling. T cells from patients with SLE are characterized by decreased expression of CD247 (CD3 zeta chain) and increased expression of FcR\(^{\gamma}\)-chain (49). Elf-1 has been shown to enhance the expression of CD247 (50–52) and suppress the expression of FcR\(^{\gamma}\)-chain (53).

The LAT gene, which encodes a protein that is downstream of the signal transduction pathway following activation of the T cell receptor, is also a target gene of Elf-1 in transcription regulation (54). Interestingly, we also identified CD247 in SLE association using the same cohorts (Li et al., manuscript in submission), suggesting important roles of this pathway in SLE.

Our work for the first time identifies this T cell transcription factor to be associated with SLE, and the various transcript forms found for this gene suggest a tight control in its expression and function. Further studies are still needed to replicate the association of this gene in other populations to firmly establish the role of this transcription factor in SLE pathogenesis. The monomorphic nature of rs7329174 in Caucasian populations suggests a potential difference among major ethnic groups on this locus in SLE susceptibility.

### MATERIALS AND METHODS

#### Subjects

One thousand three hundred twenty-two SLE samples were collected from four hospitals in Hong Kong: Queen Mary Hospital, Tuen Mun Hospital, Queen Elizabeth Hospital and Pamela Youde Nethersole Eastern Hospital. The patients were all of self-reported Chinese ethnicity living in Hong Kong. Six hundred twenty of these patients were used in the GWAS stage (612 passed quality control, see below), and the rest was used in the replication of the findings. One thousand three hundred eighty SLE patients collected in Anhui were all self-reported Chinese ethnicity living in Anhui province, central China. They were recruited from the Department of Rheumatology at both Anhui Provincial Hospital and the First Affiliated Hospital of Anhui Medical University, located in Hefei, Anhui Province, central China. Four hundred sixty-two Thai patients with SLE attending the King Chulalongkorn Memorial Hospital, a tertiary referral center in Bangkok, were recruited through the Lupus Research Unit in the Department of Microbiology, Chulalongkorn University, Thailand. All subjects met the revised criteria of the American College of Rheumatology for SLE diagnoses (55).

For the diagnosis, renal nephritis was defined as proteinuria of \(0.5 \text{ g/day}\) or biopsy-proven lupus nephritis.

Controls used in the GWAS stage were from both healthy individuals and from other studies conducted in the University of Hong Kong, genotyped with the same platform. For the replication stage, Hong Kong controls were healthy blood donors kindly contributed by the Hong Kong Red Cross and were all of self-reported Chinese ethnicity living in Hong Kong. Controls from Anhui were selected from a pool of healthy blood donors recruited from Hefei City, with an effort to match for the age and sex of the corresponding SLE patients. Thai controls were recruited from unrelated voluntary healthy donors from the same ethnic background and geographic area as the Thai SLE patients.

The Hong Kong study was approved by the Institutional Review Board of the University of Hong Kong and Hospital Authority Hong Kong West Cluster, New Territory West Cluster and Hong Kong East Cluster. The study on Anhui and Thai samples was approved by the Institutional Review Board of Research Ethics Committee of Anhui Medical University and the Ethics Committee of the Faculty of Medicine, Chulalongkorn University, respectively. All patients gave informed consent for this study.

#### Genotyping

Six hundred twenty-one SLE patients were genotyped by Illumina 610-Quad Human Beadchip. Quality control on both individual samples and SNPs was conducted as described.

### Table 2. Replication of ELF1 association with SLE in additional samples and in other Asian populations

<table>
<thead>
<tr>
<th>rs7329174</th>
<th>#_case</th>
<th>#_control</th>
<th>F_A</th>
<th>F_U</th>
<th>P-value*</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GWAS (HK)</td>
<td>612</td>
<td>1160</td>
<td>0.2826</td>
<td>0.2186</td>
<td>7.18 × 10^6</td>
<td>1.41 (1.21–1.64)</td>
</tr>
<tr>
<td>Replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HK</td>
<td>710</td>
<td>684</td>
<td>0.2611</td>
<td>0.201</td>
<td>0.0001</td>
<td>1.40 (1.17–1.68)</td>
</tr>
<tr>
<td>AH</td>
<td>1380</td>
<td>1297</td>
<td>0.2301</td>
<td>0.2047</td>
<td>0.0123</td>
<td>1.16 (1.02–1.32)</td>
</tr>
<tr>
<td>Thai</td>
<td>462</td>
<td>951</td>
<td>0.2738</td>
<td>0.2440</td>
<td>0.044</td>
<td>1.17 (0.98–1.40)</td>
</tr>
<tr>
<td>Joint Analysis</td>
<td>3164</td>
<td>4482</td>
<td>1.47 × 10^8</td>
<td>1.26 (1.16–1.36)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

HK, Hong Kong samples; AH, Anhui samples; Thai, samples from Bangkok, Thailand.

*One-sided test was applied to the replication stage but not the joint analysis.

#### Figure 2

Sequence comparison among Human, Chimpanzee and Rhesus Macaque. The position of SNP rs7329174 is pointed by an arrow.

---

**Human_all_versions**

**Chimp_NW_001224513**

**Macaca_NW_001104434**

---
before (19). A total of 612 SLE patients and 1160 controls passed quality control and were analyzed on 513 108 SNPs. In the replication stage, SNP rs7329174 was genotyped by the TaqMan SNP genotyping method using assay-on-demand probes and primers (Applied Biosystems, Foster City, CA, USA). Around 100 samples were genotyped by both Illumina Beadchip and TaqMan and the results from the two platforms had 100% concordance for this SNP.

Analysis of ELF1 transcripts, RT–PCR and PCR

EST sequences from NCBI dbEST corresponding to this locus and cDNA sequences from refseq_rna were extracted by BLAST search using the RefSeq cDNA sequence for ELF1 (GenBank: NM_172373.3) as a template, and their corresponding genomic structure was analyzed by aligning these sequences to reference human genome sequence.

PBMCs extracted from EDTA-treated blood (patients) or Buffy coats (controls) were used for the extraction of total RNA. Total RNA from patients and controls (30 individuals each) was isolated using TRizol reagent (Invitrogen Corporation, Carlsbad, CA, USA) according to the manufacturer’s instruction. First-strand cDNA synthesis was performed using high-capacity cDNA reverse transcription kit (Applied Biosystems) according to the manufacturer’s instructions. Briefly, 2 μg of total RNA (10 μl) from each sample was added to a mixture of 2.0 μl of 10× RT buffer, 0.8 μl of 25× dNTP mix (100 mm), 1.0 μl of MultiScribe™ reverse transcriptase, 2 μl of 10× RT random primers, 1.0 μl of RNase inhibitor and 3.2 μl of nuclease-free water. The final reaction mix was kept at 25°C for 10 min, 37°C for 120 min and heated at 85°C for 5 s. The reaction was stopped by cooling down to 4°C until further experiment.

Transcripts containing alternative exon1s identified by the analysis of EST from NCBI were examined by RT–PCR using the following primers. Forward primers for the three alternative exon1 variants are: 5′-AAGAAGCCACTGAAGAGGCGTCA-3′ (E1A in Fig. 1); 5′-CAGACTCGTCATGGACACACCTCCCAATCTGCAATCTGAGAATCTC-3′ (E1B) and 5′-AAGAAGCCACTGAAGAGGCGTCA-3′ (E1C). A common reverse primer from the constitutive exon2 was used for all three reactions: 5′-AGATTCCAGTACGCAGGCAGC-3′. PCRs were conducted using Qiagen Hotstart Taq master-mix by the following conditions: 94°C for 30 s, 55°C for 30 s, 72°C for 30 s for 40 cycles (42 cycles for SLE patients), and an extension cycle for 11 min at 72°C.

ACKNOWLEDGEMENTS

We thank Drs Pornchai Kingwatanakul and Ankanee Channakul for their contributions on Thai paediatric patient care and sample collection.

Conflict of Interest statement. None declared.

FUNDING

This study was partially supported by the generous donation from Shun Tak District Min Yuen Tong of Hong Kong (to Y.L.L.). W.Y. thanks support from Research Grant Council of the Hong Kong Government (GRF HKU781079M). D.Q.Y. thanks grant support from the key program of National Natural Science Foundation of China (30830089). N.H. and Y.A thank support from the National Research Council of Thailand; K.S. and V.S. thank supports from Thailand Research Fund. Genotyping of samples serving as controls in this study was supported by Research Grants Council of Hong Kong GRF (HKU 7623/08M, HKU 7747/07M and HKU775608M). Y.Z., S.Z. and J.Z. are supported by Edward Sai-Kim Hotung Paediatric Education and Research Fund and University Postgraduate Studentship.

REFERENCES


