The **STAT4** gene influences the genetic predisposition to systemic sclerosis phenotype


1Instituto de Parasitología y Biomedicina ‘Lopez-Neyra’ (CSIC), Granada, Spain, 2Department of Rheumatology and 3Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, 4Servicio de Medicina Interna, Hospital Valle de Hebron, Barcelona, Spain, 5Department of Rheumatology, Lund University Hospital, S-221 85 Lund, Sweden, 6Servicio de Medicina Interna, Hospital Universitario Central de Asturias, Oviedo, Spain, 7Servicio de Medicina Interna, Hospital Clínico Universitario, Granada, Spain, 8Department of Rheumatology and Clinical Immunology, Charité University Hospital, Berlin, Germany, 9Department of Rheumatology, Sint Maartenskliniek, Nijmegen, The Netherlands, 10Servicio de Medicina Interna, Hospital Virgen del Rocio, Sevilla, Spain, 11Servicio de Reumatología, Hospital Reina Sofía, Córdoba, Spain, 12Servicio Medicina Interna, Hospital Virgen de la Victoria, Málaga, Spain, 13Servicio de Reumatología, Hospital del Mar, Barcelona, Spain, 14Servicio de Medicina Interna, Hospital Carlos Haya, Malaga, Spain, 15Servicio de Reumatología, Hospital Xeral-Calde, Lugo, Spain, 16Servizio di Reumatologia ed Immunologia Clinica Spedali Civili, Brescia, Italy, 17Referral Center for Systemic Autoimmune Diseases, University of Milan, Milan, Italy, 18Department of Rheumatology, Newcastle, UK, 19Servicio de Inmunología, Hospital Virgen del Rocio, Sevilla, Spain and 20Department of Medicine, University of Washington, Seattle, WA, USA

Received December 3, 2008; Revised February 25, 2009; Accepted March 9, 2009

The aim of this study was to investigate the possible role of **STAT4** gene in the genetic predisposition to systemic sclerosis (SSc) susceptibility or clinical phenotype. A total of 1317 SSc patients [896 with limited cutaneous SSc (lcSSc) and 421 with diffuse cutaneous SSc (dcSSc)] and 3113 healthy controls, from an initial case–control set of Spanish Caucasian ancestry and five independent cohorts of European ancestry (The Netherlands, Germany, Sweden, Italy and USA), were included in the study. The rs7574865 polymorphism was selected as **STAT4** genetic marker. We observed that the rs7574865 T allele was significantly associated with susceptibility to lcSSc in the Spanish population (**P** = 1.9 × 10⁻⁵ odds ratio (OR) 1.61 95% confidence intervals (CI) 1.29–1.99), but not with dcSSc (**P** = 0.41 OR 0.84 95% CI 0.59–1.21). Additionally, a dosage effect was observed showing individuals with rs7574865 TT genotype higher risk for lcSSc (OR 3.34, **P** = 1.02 × 10⁻⁷ 95% CI 2.11–5.31). The association of the rs7574865 T allele with lcSSc was confirmed in all the replication cohorts with different effect sizes (OR ranging between 1.15 and 1.86), as well as the lack of association of **STAT4** with dcSSc. A meta-analysis to test the overall effect of the rs7574865 polymorphism showed a strong risk effect of the T allele for lcSSc susceptibility (pooled OR 1.54 95% CI 1.36–1.74; **P** < 0.0001). Our data show a strong and reproducible association of the **STAT4** gene with the genetic predisposition to lcSSc suggesting that this gene seems to be one of the genetic markers influencing SSc phenotype.

∗To whom correspondence should be addressed at: Instituto de Parasitología y Biomedicina ‘Lopez-Neyra’, Consejo Superior de Investigaciones Científicas, Parque Tecnológico Ciencias de la Salud, Avenida del Conocimiento s/n 18100-Armilla, Granada, Spain. Tel: +34 958181669; Fax: +34 958181632; Email: martin@ipb.csic.es

© The Author 2009. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oupjournals.org
INTRODUCTION

Systemic sclerosis (SSc) is one of the most disabling autoimmune conditions characterized by an extensive fibrotic process that affects multiple organs and tissues (1). Although the etiology of SSc is still poorly understood, it is widely accepted that the interaction of environmental factors with different individual genetic factors leads to SSc development (2).

To date, our knowledge of genetic factors contributing to SSc susceptibility or clinical phenotypes is very limited. Only genes within the major histocompatibility complex have been associated with SSc genetic predisposition or clinical manifestations in a consistent and reproducible fashion (2). Similar to other complex genetic disorders, several loci are expected to contribute to SSc genetic predisposition each with only moderate magnitude (3).

Genes implicated in the main pathogenic mechanisms of SSc are interesting candidates. One of central events responsible of the SSc development is the dysregulation of the immune system. The altered immune response in SSc is marked by an increased T cell activation and the secretion of pro-inflammatory mediators that contribute to the generation of fibrosis and endothelial alterations, hallmarks of SSc (4).

Several mechanisms regulate T cell activation and differentiation, being one of the most important the specific activation of gene transcription after cytokine stimulation (5). Signal transducers and activators of transcription (STAT) are a family of transcription factors that have been demonstrated to exert a fundamental role in driving T cell differentiation and signaling (6,7). Among the six described STAT proteins, STAT4 has acquired great interest. STAT4 is implicated in the differentiation of Th1 and also in the recently described Th17 cells, two of the T cell subsets that are implicated in SSc pathogenesis (4,8–10). In addition, the STAT4 gene has been associated with genetic predisposition to different autoimmune diseases (AIDs), such as rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), type 1 diabetes, inflammatory bowel disease and Sjögren’s syndrome, showing a tag single nucleotide polymorphism (SNP) of STAT4 intron 3 (rs7574865) the strongest effect (11–17).

In view of these findings, we considered the STAT4 gene as an excellent candidate gene and aimed to investigate its possible implication in the genetic predisposition to SSc susceptibility or clinical phenotype.

RESULTS

STAT4 is associated with limited cutaneous SSc in the Spanish population

First we conducted an association study in a case–control set of Spanish Caucasian ancestry. The distribution of the STAT4 rs7574865 genotypes and alleles in the Spanish controls was very similar to that reported previously in Caucasian populations [minor allele frequency (MAF) 0.21] and were observed to be in Hardy–Weinberg equilibrium (HWE) (Table 1) (11). After comparing the genotypes and alleles of the rs7574865 polymorphism between SSc patients with limited disease and healthy controls, we observed that the rs7574865 T allele was strongly associated with lcSSc \[ P = 1.9 \times 10^{-5} \text{ odds ratio (OR) 1.61 95\% confidence intervals (CI) 1.29–1.99} \] (Table 1). In addition, the TT genotype was significantly increased in lcSSc patients compared with controls (\( P = 1.02 \times 10^{-7} \) OR 3.34 95\% CI 2.11–5.31) (Table 1). Nevertheless, no evidence of association was observed when the distribution of STAT4 rs7574865 genotypes or alleles between patients with diffuse disease and healthy controls was compared (Table 1).

The increased frequency of the rs7574865 T allele and genotype among lcSSc patients was observed also when they were compared with dcSSc patients (T allele in lcSSc 29.5% versus 23.3% in dcSSc), although the differences did not reach statistical significance (Table 1).

A replication study in five independent Caucasian populations and a meta-analysis confirm that STAT4 is strongly associated with limited SSc

In view of the interesting findings observed in our Spanish population and in order to confirm the association of STAT4 with lcSSc, we conducted a large replication study including five independent populations with Caucasian ancestry. All analyzed control populations were in HWE for the STAT4 rs7574865 genetic variant. The frequency for the rs7574865 T allele in the control populations ranged between 0.21 and 0.25 as expected from previous studies (11).

Interestingly, we confirmed the significant association of the rs7574865 T allele with susceptibility to lcSSc in all cohorts tested including patients from the Netherlands (OR 1.8 95\% CI 1.32–2.47), Germany (OR 1.6 95\% CI 1.15–2.21), North America (OR 1.4 95\% CI 1.04–3.66) and Italy (OR 2.0 95\% CI 1.05–1.78) (Table 2). Although not statistical significant, we observed the same trend in the Swedish population (OR for the rs7574865 T allele 1.3). Probably, this was due to the lower frequency for both rs7574865 TT genotype and T allele observed in the Swedish lcSSc patients compared with that observed in the other populations analyzed (Table 2). Furthermore, the lack of association of STAT4 rs7574865 polymorphism with the diffuse cutaneous SSc subtype was confirmed in all five independent case–control sets (Table 2).

Additionally, we conducted a meta-analysis to test the overall effect of the rs7574865 T allele on lcSSc susceptibility. The estimation of the homogeneity between the Spanish, Dutch, German, North American, Italian and Swedish populations did not reveal significant differences between them. In consequence, we performed a combined analysis considering the data from the six case–control series using the Mantel–Haenszel test under fixed effects that again showed the strong risk effect of the STAT4 rs7574865 T allele for lcSSc genetic predisposition (pooled OR 1.54 95\% CI 1.36–1.74; \( P < 0.0001 \)) (Fig. 1A). Furthermore, in the pooled analysis, individuals carrying the TT genotype showed an increased risk for lcSSc susceptibility (pooled OR of 2.55 95\% CI 1.93–3.36, \( P < 0.0001 \)). The same trend was observed for the comparison of lcSSc with dcSSc that showed a significant increased frequency of the STAT4 rs7574865 T allele and TT genotype in the lcSSc patients group (Pooled OR for the T allele 1.27 95\% CI 1.04–1.54, \( P = 0.019 \) (Fig. 1B); Pooled OR for the TT genotype 1.7 95\% CI 1.10–2.60, \( P = 0.02 \)).

In addition, we performed a pooled analysis to further investigate the overall effect of the rs7574865 T allele on...
For the comparison of lcSSc versus dcSSc:
P-value T allele: 0.11 OR 1.36 95% CI 0.92–2.02.
P-value TT genotype: 0.11 OR 1.83 95% CI 0.75–4.42.

Table 2. Distribution of STAT4 rs7574865 genetic variant in five replication cohorts

<table>
<thead>
<tr>
<th>Population</th>
<th>rs7574865</th>
<th>Controls, n = 1296</th>
<th>Limited SSc, n = 242</th>
<th>P-value</th>
<th>OR 95% CI</th>
<th>Diffuse SSc, n = 90</th>
<th>P-value</th>
<th>OR 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dutch</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>813 (62.7)</td>
<td>130 (53.7)</td>
<td>0.008</td>
<td>0.68</td>
<td>(0.50–0.77)</td>
<td>54 (60.0)</td>
<td>0.61</td>
<td>0.88</td>
</tr>
<tr>
<td>GT</td>
<td>428 (33.0)</td>
<td>81 (33.5)</td>
<td>0.89</td>
<td>1.02</td>
<td>(0.76–1.37)</td>
<td>30 (33.3)</td>
<td>0.95</td>
<td>1.03</td>
</tr>
<tr>
<td>TT</td>
<td>55 (4.2)</td>
<td>31 (12.8)</td>
<td>1.02 x 10^-7</td>
<td>3.34</td>
<td>(2.11–5.31)</td>
<td>6 (6.7)</td>
<td>0.29</td>
<td>1.82</td>
</tr>
<tr>
<td><strong>Allele</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>2054 (79.2)</td>
<td>341 (70.5)</td>
<td>1.9 x 10^-5</td>
<td>0.62</td>
<td>(0.50–0.77)</td>
<td>138 (76.7)</td>
<td>0.41</td>
<td>0.84</td>
</tr>
<tr>
<td>T</td>
<td>538 (20.8)</td>
<td>143 (29.5)</td>
<td>1.61</td>
<td>1.29</td>
<td>(1.19–2.99)</td>
<td>42 (23.3)</td>
<td>1.18</td>
<td>0.82</td>
</tr>
</tbody>
</table>

**STAT4 is not implicated in SSc clinical manifestations**

The possible implication of STAT4 in SSc clinical manifestations was further investigated. First, we analyzed the Spanish cohort individuals in whom data regarding selective auto-antibodies status (306 SSc patients for anti-Scl-70 and 294 for ACA) were available. Data related to pulmonary function were available in 281 SSc patients for pulmonary fibrosis, 233 for DLCO and 279 for forced vital capacity (FVC). We observed no significant association between the rs7574865 genetic variant and the presence of SSc specific Antibodies (Scl70 or ACA) or pulmonary involvement (Table 3). This lack of association of STAT4 and SSc clinical features was confirmed in the five replication cohorts (data not shown).

**DISCUSSION**

SSc is characterized by skin fibrosis, vascular damage and immune system activation leading to a very heterogeneous pattern of clinical symptoms. Although, these three key
The STAT4 gene may be considered as a novel susceptibility gene for lcSSc.

Nevertheless, our data does not support evidence that the STAT4 gene exerts such a strong effect in dcSSc genetic predisposition. Certainly, the independent analyses of rs7574865 T allele and TT genotype and the meta-analysis of rs7574865 TT genotype in dcSSc compared with the healthy population point that this polymorphism has a negligible effect on dcSSc susceptibility. Only the meta-analysis of the rs7574865 T allele showed a trend for association with dcSSc. In this regard, considering that there is still no consensus on the levels of statistical significance for meta-analysis (19), the P-value (0.03) and the 95% intervals close to the null obtained for the meta-analysis of the rs7574865 T allele found in our study do not allow us to conclude an association of this SNP with dcSSc. In addition, the comparison of the overall distribution of rs7574865 T and TT genotype between lcSSc and dcSSc patients showed that this genetic marker confers a significant increased risk for SSc limited phenotype susceptibility. However, due to the lower proportion of dcSSc in European populations, the total number of dcSSc included in our study (n = 421) may not be enough to reach a high statistical power. On the other hand, it has been empirically demonstrated that meta-analyses with P-values between 0.01 and 0.05, similar to that obtained in our analyses for the dcSSc phenotype, do not have strong reliability (20). On this basis, the results from the present study suggest that the STAT4 gene does not seem to play a major role in dcSSc; however, further independent studies are needed to confirm this hypothesis.

STAT4 is an essential transcription factor for the regulation of the immune response. Upon the stimulation of cytokines such as IL-12, IL-23 or IL-17 STAT4 is activated and drives the expression of several pro-inflammatory mediators implicated in the differentiation and proliferation of Th1 and Th17 T cell subsets (8,21). Therefore, it is plausible that the elevated levels of IL-12, IL-23 and IL-17 observed in SSc patients lead to the activation of the STAT4 pathway (22,23). Then, a prolonged STAT4 increased activity due to different genetic variation in this gene might cause a sustained inflammatory response together with the expansion and infiltration of pro-inflammatory T cell subpopulations in skin and internal organs of SSc patients. However, further functional studies are necessary to elucidate the exact molecular mechanisms by which STAT4 is implicated in the pathogenesis of SSc and more precisely, how this transcription factor can lead to the development of the lcSSc.

In the past few years, using both genome-wide association studies and candidate gene association studies, genes such as PTPN22, CTLA4 or IL23R have been associated with genetic susceptibility to various autoimmune conditions, leading to the hypothesis that different AIDs may share common genetic factors and pathways (24–27).

The association of STAT4 with SSc reported here, together with previous findings that have shown recently the association of STAT4 with susceptibility to RA, SLE and Sjögren syndrome support the notion that this gene seems to be another common genetic factor for autoimmunity. Nevertheless, to understand how STAT4 influences the development of AIDs the next step is to conduct functional studies in
order to identify which is/are the real causing variants that may influence STAT4 activity or expression.

In summary, in this study we described for the first time a strong and reproducible association of the STAT4 gene with the genetic susceptibility to lcSSc. Our data shed light on the pathogenic mechanisms that may underlie SSC development and open a new opportunity for the treatment of this debilitating disease.

**MATERIAL AND METHODS**

**Patients**

A total of 1317 SSC patients and 3113 controls were included in the study. First we analyzed an initial case–control set of 332 SSC patients (242 with lcSSc and 90 with dcSSc) and 1296 healthy controls of Spanish Caucasian ancestry. Additionally, five independent replication cohorts were analyzed (Dutch: 101 lcSSc, 30 dcSSc and 893 controls; German: 153 lcSSc, 117 dcSSc and 227 controls; North American: 30 lcSSc, 53 dcSSc and 77 controls; Italian: 259 lcSSc, 153 dcSSc and 227 controls; Swedish: 111 lcSSc, 39 dcSSc and 285 controls).

All the patients fulfilled the 1980 American College of Rheumatology (ACR) classification criteria for SSC (28). In addition, patients were classified as having limited or diffuse SSC. When patients with SSC have cutaneous involvement proximal from elbows and knees they fulfilled definitions for limited scleroderma (29). Those SSC patients with cutaneous changes proximal from elbows and knees were classified as having diffuse SSC (30).

In addition, the following clinical data were collected for ascertainment of SSC clinical phenotype; age, gender, disease duration and presence of SSC specific auto-antibodies, anti-topoisomerase (Anti-Scl70) and anti-centromere (ACA). Lung involvement was assessed according to the international guidelines (31). Pulmonary fibrosis was assessed by a computed tomography scan. Restrictive syndrome and diffusion capacity of the lungs was defined as a FVC <75% of the predicted value and a diffusion capacity for carbon monoxide (DLCO) of less than 75% of predicted (based on age, sex, height and ethnic origin). The main clinical features of the SSC patients from all the analyzed case sets are summarized in Table 4.

The control population consisted in unrelated healthy individuals recruited in the same geographical region as SSC patients and matched by age, sex and ethnicity with the SSC patients groups.

The study was approved by local ethical committees from all the participating centers. Both patients and controls were included in the study after written informed consent.

**STAT4 genotyping**

DNA from patients and controls was obtained using standard methods. As STAT4 genetic marker, we selected the rs7574865 polymorphisms, since this is the tagger SNP of the haplotype block of STAT4 intron 3 associated with autoimmunity, and at the same time the genetic variant showing the strongest association with AIDs. Samples were genotyped for the rs7574865 polymorphism by Taqman 5’-allelic discrimination assay technology using a Pre-designed SNP Genotyping
Assays provided by Applied Biosystems (Part number: C_29882391_10, Foster City, CA, USA). The PCR reaction was performed in a total volume of 5 μl with the following amplification protocol: denaturation at 92°C for 10 min, followed by 40 cycles of denaturation at 92°C for 15 s and annealing and extension at 60°C for 1:00 min. Post-PCR, the genotype of each sample was automatically attributed by measuring the allele-specific fluorescence in the ABI Prism 7900 Sequence Detection System, using the SDS 2.3 software for allele discrimination (Applied Biosystems, Foster City, CA, USA).

All samples were genotyped in the same center to avoid genotyping inconsistencies and to verify the genotyping consistency, allelotype samples were genotyped twice showing 99% identical genotypes.

Statistical analysis
We tested HWE for each case–control set by using the program FINETI (http://ihg.gsdf.org/cgi-bin/hw/hwa2.pl). Significance was calculated by 2x2 contingency tables and Fisher’s exact test, to obtain P-values, OR and 95% CI by using Statcalc software (Epi Info 2002; Centers for Disease Control and Prevention, Atlanta, GA, USA). P-values below 0.05 were considered as statistically significant. The analysis of the combined data from all populations was performed using the Stats Direct software. First, homogeneity of OR among cohorts was calculated using Breslow-Day and Woolf Q methods. We then performed a calculation of the pooled OR under a fixed-effects model (Mantel–Haenszel meta-analysis) or random effects (DerSimonian-Laird) when necessary.

The estimation of the power of the study was performed using the Quanto v 0.5 software (Department of Preventive Medicine, University of Southern California, CA, USA). For the pooled analysis of lcSSc (n = 841) and considering a medium MAF of 0.25, our study reach a 93% power to detect the effect of a polymorphism at an OR of 1.3 similar to that observed for the rs7574865 T allele in previous studies (11,12). Under the same conditions, the estimation of the power for the pooled analysis of deSSc that included a total of 421 patients was 67%.

ACKNOWLEDGEMENTS
We like to thank Sofia Vargas for her invaluable contribution in the collection, isolation and storage of the DNA samples.

Conflict of Interest statement. None declared.

FUNDING
This work was supported by grants SAF2006-00398, and in part by Junta de Andalucía, grupo CTS-180. B.R. is supported by the I3P CSIC program funded by the ‘Fondo Social Europeo’. T.R.D.J.R. was sponsored by an unrestricted Grant from the ‘Landelijk Katholieke Reuma Centrum (LKR)’, and VENI and VIDI laureate from the Dutch association of Research (NWO).

REFERENCES