Journal of Autoimmunity 34 (2010) 155–162

Contents lists available at ScienceDirect

Journal of Autoimmunity

journal homepage: www.elsevier.com/locate/jautimm

Association of the *C8orf13-BLK* region with systemic sclerosis in North-American and European populations

Pravitt Gourh [a,1], Sandeep K. Agarwal [a,*,1], Ezequiel Martin [b,1], Dipal Divecha [a], Blanca Rueda [b], Haley Bunting [a], Shervin Assassi [a], Gene Paz [a], Sanjay Shete [c], Terry McNearney [d], Hilda Draeger [e], John D. Reveille [a], T.R.D.J. Radstake [f], Carmen P. Simeon [g], Luis Rodriguez [h], Esther Vicente [i], Miguel A. Gonzalez-Gay [j], Maureen D. Mayes [a], Filemon K. Tan [a], Javier Martin [b,1], Frank C. Arnett [a,1]

[a] *Division of Rheumatology and Clinical Immunogenetics, Department of Internal Medicine, University of Texas Health Science Center at Houston (UTHSC-H), Houston, TX, USA*
[b] *Instituto de Parasitología y Biomedicina, CSIC, Granada, Spain*
[c] *Department of Epidemiology, Division of Cancer Prevention and Population Sciences, The University of Texas, MD Anderson Cancer Center, Houston, TX, USA*
[d] *University of Texas Medical Branch at Galveston, Galveston, TX, USA*
[e] *University of Texas Health Science Center at San Antonio, San Antonio, TX, USA*
[f] *Dept of Rheumatology, Radboud University Nijmegen Medical Center, The Netherlands*
[g] *Dept of Internal Medicine, Hospital Valle de Hebron, Barcelona, Spain*
[h] *Dept of Rheumatology, Hospital Clínico San Carlos, Madrid, Spain*
[i] *Dept of Rheumatology, Hospital de la Princesa, Madrid, Spain*
[j] *Dept of Rheumatology, Hospital Xeral-Calde, Lugo, Spain*

ARTICLE INFO

Article history:
Received 7 July 2009
Received in revised form
21 August 2009
Accepted 29 August 2009

Keywords:
Scleroderma
Systemic sclerosis/SSc
Polymorphism/SNP
BLK
C8orf13
Anti-Topoisomerase I
Anti-centromere
Genetics
Autoantibody
rs13277113
rs2736340

ABSTRACT

Objective: Genetic studies in the systemic sclerosis (SSc), an autoimmune disease that clinically manifests with dermal and internal organ fibrosis and small vessel vasculopathy, have identified multiple susceptibility genes including HLA-class II, *PTPN22*, *IRF5*, and *STAT4* which have also been associated with other autoimmune diseases, such as systemic lupus erythematosus (SLE). These data suggest that there are common autoimmune disease susceptibility genes. The current report sought to determine if polymorphisms in the *C8orf13-BLK* region (chromosome 8p23.1-B lymphoid tyrosine kinase), which is associated with SLE, are associated also with SSc.
Methods: Two variants in the *C8orf13-BLK* region (rs13277113 & rs2736340) were tested for association with 1050 SSc cases and 694 controls of North Americans of European descent and replicated in a second series 589 SSc cases and 722 controls from Spain.
Results: The "T" allele at rs2736340 variant was associated with SSc in both the U.S. and Spanish case–control series ($P = 6.8 \times 10^{-5}$, OR 1.27, 95% CI 1.1–1.4). The "A" allele at rs13277113 variant was associated with SSc in the U.S. series only ($P = 3.6 \times 10^{-4}$, OR 1.32, 95% CI 1.1–1.6) and was significant in the combined analyses of the two series ($P = 2.0 \times 10^{-3}$; OR 1.20, 95% CI 1.1–1.3). Both variants demonstrated an association with the anti-centromere antibody ($P = 2.2 \times 10^{-6}$ and $P = 5.5 \times 10^{-4}$, respectively) and limited SSc ($P = 3.3 \times 10^{-5}$ and $P = 2.9 \times 10^{-3}$, respectively) in the combined analysis. Peripheral blood gene expression profiles suggest that B-cell receptor and NFκB signaling are dysregulated based on the risk haplotype of these variants.
Conclusion: We identify and replicate the association of the *C8orf13-BLK* region as a novel susceptibility factor for SSc, placing it in the category of common autoimmune disease susceptibility genes.

1. Introduction

Scleroderma (systemic sclerosis, SSc) is a chronic, multisystem autoimmune disease clinically characterized by progressive fibrosis of the skin and internal organs. Pathologically, SSc exhibits three cardinal features: inflammation and autoimmunity, vasculopathy and excessive extracellular matrix production and deposition. How

* Corresponding author. Division of Rheumatology and Clinical Immunogenetics, 6431 Fannin St., MSB 5.278, Houston, TX 77030, USA. Tel.: +1 713 500 6900; fax: +1 713 500 0580.
E-mail address: sandeep.k.agarwal@uth.tmc.edu (S.K. Agarwal).
[1] These authors contributed equally to this manuscript.

0896-8411/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jaut.2009.08.014

the disease process is triggered, remains to be established, but current paradigms point towards immune dysregulation as a central process in the pathogenesis of SSc.

Multiple lines of evidence, in patients as well as animal models of dermal fibrosis, point to dysregulation of the immune system and autoimmunity in the pathogenesis of SSc. Alterations in T-cell profiles and cytokines have been demonstrated [1]. In addition, the B-cell lineage is implicated [1,2]. CD19 deficient mice have reduced dermal fibrosis after subcutaneous bleomycin administration [3]. The presence of multiple SSc-associated autoantibodies, most commonly anti-centromere (ACA), anti-topoisomerase Q1 antibodies (ATA), and anti-RNA polymerase III antibodies (ARA), has been well-described in SSc patients and provides indirect evidence for the importance of the B-cell lineage in SSc [4–6]. Interestingly, the SSc-associated autoantibodies correlate with distinct clinical subsets characterized by extent of cutaneous involvement and pattern of organ involvement [7,8]. For example, pulmonary arterial hypertension is more common in patients with ACA, pulmonary fibrosis is more common in patients with ATA, and scleroderma renal crisis is more common in patients with ARA [7].

Recently, several genetic polymorphisms have been associated with SSc in susceptibility [9–15]. Many of these genes are susceptibility factors for other autoimmune diseases suggesting that they are common autoimmune disease susceptibility genes. These data support the paradigm of common dysregulated pathways across multiple autoimmune diseases. Interestingly, some of these genetic variants associated with SSc are associated with susceptibility to developing SSc, while others increase risk only in specific autoantibody subsets seen in SSc [9,10].

Two genome wide association studies in systemic lupus erythematosus (SLE) have implicated the *C8orf13-BLK* (B lymphoid tyrosine kinase gene) region of chromosome 8p23.1 as a susceptibility locus for SLE [16,17]. These findings were further replicated in a Japanese population [18]. B lymphoid kinase (Blk), encoded by the *BLK* gene is a member of the Src family kinases (SFKs) which includes Blk, Lck, Fyn, Lyn, c-Src, c-Yes, Fgr, and Hck [19]. Blk is the only SFK that is exclusively expressed in B cells and thymocytes and not in mature T cells [20–22]. Blk transduces signals downstream of the B cell receptor (BCR) and plays a role in BCR signaling and B cell development [23,24]. B cell development is dependent upon the activation of the transcription factor nuclear factor κB (NFκB) by SFKs (Blk, Fyn, Lyn) [23]. The *C8orf13* gene is ubiquitously expressed but its exact function is currently unknown.

Given the importance of B-cells and autoantibodies in the pathogenesis of SSc as well as SLE and the emerging paradigm that autoimmune diseases may share common susceptibility genes, the current study sought to investigate the potential association of the *C8orf13-BLK* region variants with SSc in two large, independent, and well-described case–control series. Herein we demonstrate a significant association of both *C8orf13-BLK* region variants with susceptibility to SSc and more strongly with the anti-centromere antibody subset of SSc in both case–control series. We observe a dysregulation of the B-cell receptor and NFκB signaling based on the risk haplotype of these variants in peripheral blood gene expression arrays.

2. Patients and methods

2.1. SSc patients and controls

In this study, we combined 1050 patients of North Americans of European descent with SSc and 694 unrelated healthy controls of North Americans of European descent from the Scleroderma Family Registry and DNA Repository, the Genetics versus Environment in Scleroderma Outcomes Study (GENISOS), and from SSc patients evaluated in the University of Texas Rheumatology Division, dating from 1986 to present, as previously described [8,25]. For a replication cohort, an independent Spanish case–control series of 589 SSc patients and 722 healthy controls was included to confirm the genetic association. The control population was matched by age, sex and ethnicity with the SSc patients group, as previously described [15]. All SSc patients fulfilled American College of Rheumatology (ACR) preliminary criteria for disease classification or had at least 3 of the 5 CREST features (***C***alcinosis, ***R***aynaud's phenomenon, ***E***sophageal dysfunction, ***S***clerodactyly, and ***T***elangiectasias) [9,10,26,27]. Patients were classified as limited (cutaneous changes distal to the elbow) and diffuse (cutaneous changes proximal to the elbow) as defined by Leroy et al. [28]. Clinical and serologic characteristics of the SSc patients are presented in Table 1. All study subjects provided written informed consent and the study was approved by the institutional review board (IRB) of the University of Texas Health Science Center at Houston and the corresponding IRB in Spain.

2.2. Autoantibody analysis

All SSc patients were tested for anti-nuclear antibodies using indirect immunofluorescence (IIF) and HE*p*-2 cells as antigen substrate (Antibodies Inc., Davis, CA). Titers of 1/80 and higher were considered positive. Anti-centromere antibodies (ACA) were determined by their distinctive IIF pattern on HE*p*-2 cells. Autoantibodies to topoisomerase I (ATA) were determined by passive immunodiffusion against calf thymus extract using a commercially available kit (Inova Diagnostics, San Diego, CA) and read on an Electronic RID Plate Reader AD001.c (The Binding Site, San Diego, CA) and magnified visual inspection. Anti-RNA polymerase III antibodies (ARA) were determined using a commercially available enzyme-linked immunoassay (EIA) kit (MBL Co. Ltd, Nagoya, Japan). For ARA, the cutoff was defined as 2.5 standard deviations above the mean of 40 healthy controls.

2.3. SNP selection and genotyping

The two *C8orf13-BLK* gene region variants (rs13277113 and rs2736340) showing the strongest association with SLE in a genome wide study were selected for genotyping (13). Genomic DNA was extracted from peripheral blood according to the manufacturer's protocol with the PureGene genomic DNA isolation kit (Gentra Systems). The two variants were genotyped using a predesigned TaqMan SNP genotyping assay from Applied Biosystems (ABI, Foster City, CA). PCR amplification was performed and the genotypes were

Table 1
Clinical and serologic characteristics of patients with systemic sclerosis and gender distribution in controls.*

	North American	Spanish
Controls		
	($N = 694$)*	($N = 722$)*
Female	331 (50.7)	358 (49.9)
Male	322 (49.3)	360 (50.1)
Systemic Sclerosis		
	($N = 1050$)	($N = 589$)*
Female	928 (88.4)	516 (89.1)
Male	122 (11.6)	63 (10.9)
Skin involvement	($N = 995$)	($N = 475$)
Limited systemic sclerosis	605(60.8)	326 (68.7)
Diffuse systemic sclerosis	390 (39.2)	149 (31.3)
Antibodies	($N = 1050$)	($N = 589$)
Anti-centromere	297 (28.3)	213 (36.2)
Anti-topoisomerase I	173 (16.5)	103 (17.5)
Anti-RNA Polymerase III	191 (18.2)	–
Negative for above Antibodies	389 (63.0)	273 (53.7)

*Missing gender information.

automatically attributed by measuring the allele-specific fluorescence in the ABI 7900HT system (ABI). Automated allele calling was performed by allelic discrimination plots using SDS 2.3 software from ABI. Standard control DNA (CEPH 1347-02 from ABI) was added in replicates to minimize error and check genotyping quality.

2.4. Gene expression array of peripheral white blood cells

Blood samples were drawn directly into PAXgene™ tubes (PreAnalytix, Franklin Lakes, NJ). The protocol was modified by adding RNase inhibitor to the Paxgene tubes during thawing and total RNA was isolated according to manufacturer's protocol utilizing PAXgene RNA kit. The RNA quality and yield were assessed by Agilent 2100 Bioanalyzer (Agilent Technologies, Inc., Palo Alto, CA) and NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE). Two hundred nanograms of total RNA were amplified and purified using Illumina TotalPrep RNA Amplification Kit (Applied Biosystems/Ambion, Austin, TX). The amplified cRNA was hybridized on Illumina Human Ref-8 v2 arrays and the data were extracted with Illumina Beadstudio software suite (Illumina Inc, San Diego, CA).

2.5. Gene expression array data analysis

The initial analysis was performed with Illumina's Beadstudio. Raw data were also analyzed in BRB-ArrayTools version 3.7 (NCI) developed by Dr. Richard Simon and BRB-ArrayTools Development Team (National Cancer Institute, USA). Any gene whose signal detection *P*-value was not significantly different from the negative controls ($P < 0.05$) was removed from the analysis. Data were normalized using the median over the entire array. A gene was defined as differentially expressed when the significance level for the comparison was $p \leq 0.05$. BRB array tools were used for class comparison analysis and these sets of differentially expressed genes were modeled in Ingenuity Pathway Analysis (IPA) (Ingenuity Systems, www.ingenuity.com).

2.6. Statistical analysis

Statistical analyses were performed using SAS 9.1.3 (SAS Institute Inc., Cary, NC, USA) software. Allelic and genotypic associations were calculated using the standard Pearson's chi-square test or, when appropriate, Fisher's exact test. Odds ratios and 95% confidence intervals were calculated. The two series were also combined and tested for association using Mantel Haenszel test to increase the power to detect an association and reduce chances of false positive associations. Significance was calculated at alpha value ≤ 0.05. Logistic regression analysis was performed to identify significant independent risk factors and to evaluate the role of the genetic variability of *C8orf13-BLK* region polymorphisms in SSc and its autoantibody and phenotypic subsets. Measures of pairwise LD were determined using Haploview (Whitehead Institute for Biomedical Research, Cambridge, MA).

3. Results

3.1. SNP genotyping, linkage disequilibrium, and haplotype block structure

The *C8orf13-BLK* gene region variants (rs13277113–rs2736340) were in Hardy–Weinberg Equilibrium in both the North American and Spanish control samples. CEPH DNA samples were used as controls to improve the accuracy of allele calling on Taqman 5′ genotyping assay. The SNP call rate for both the SNPs in both the case–control series was greater than 97%. Pairwise LD was calculated by both D′ and r^2 between rs13277113 and rs2736340 variants in both the control series. The SNPs demonstrated strong LD in both the North American and Spanish healthy control series ($r^2 = 0.97$ and $r^2 = 0.86$, respectively).

3.2. C8orf13-BLK gene region variant – rs2736340 association analysis

Case–control comparisons of the rs2736340 variant for association with SSc and its subsets were performed in the North American series (Table 2). The "TT" genotype frequency was significantly increased in the SSc cases (9.8%) as compared to controls (7.4%) with a $P = 1.4 \times 10^{-2}$. The "TT" genotype was significantly increased in the limited skin subset and the ACA autoantibody subset of SSc. The "T" allele was increased in SSc cases (29.5%) as compared to controls (24.9%) with a $P = 3.0 \times 10^{-3}$, OR 1.26 [95% CI 1.1–1.5]. The "T" allele was significantly increased in both the limited and diffuse skin

Table 2
Genotypic and allelic distribution of the rs2736340 variant in North American and Spanish SSc patients and healthy controls.*

	North American				Spanish				MH combined	
	N (%)				*N* (%)					
	CC	CT	TT	*P*-value	CC	CT	TT	*P*-value	*P*-value	OR (95% CI)
Control subjects	398 (57.7)	241 (34.9)	51 (7.4)		436 (60.4)	248 (34.3)	38 (5.3)			
Patients with SSc	518 (50.8)	401 (39.4)	100 (9.8)	1.4×10^{-2}	315 (53.5)	231 (39.2)	43 (7.3)	3.0×10^{-2}	9.4×10^{-5}	1.33 (1.1–1.5)
Limited SSc	289 (49.5)	237 (40.6)	58 (9.9)	1.1×10^{-2}	169 (51.8)	130 (39.9)	27 (8.3)	1.8×10^{-2}	4.7×10^{-5}	1.41 (1.2–1.7)
Diffuse SSc	199 (52.4)	142 (37.4)	39 (10.3)	0.13	79 (53.4)	53 (35.8)	6 (4.1)	0.16	0.06	1.22 (1.0–1.5)
Antibodies										
ACA	128 (44.6)	127 (44.3)	32 (11.1)	7.1×10^{-4}	107 (50.2)	88 (41.3)	18 (8.5)	0.30	3.4×10^{-6}	1.61 (1.3–2)
ATA	98 (58.3)	56 (33.3)	14 (8.3)	0.92	53 (51.5)	45 (43.7)	5 (4.9)	0.18	0.38	1.12 (0.9–1.5)
ARA	99 (53.8)	62 (33.7)	23 (12.5)	0.09	–	–	–	–	–	–
	C	T	*P*-value	OR (95% CI)	C	T	*P*-value	OR (95% CI)	*P*-value	OR (95% CI)
Control subjects	1037 (75.1)	343 (24.9)			1120 (77.6)	324 (22.4)				
Patients with SSc	1437 (70.5)	601 (29.5)	3.0×10^{-3}	1.26 (1.1–1.5)	861 (73.1)	317 (26.9)	8.0×10^{-3}	1.27 (1.1–1.5)	6.8×10^{-5}	1.27 (1.1–1.4)
Limited SSc	815 (69.8)	353 (30.2)	2.5×10^{-3}	1.31 (1.1–1.6)	468 (71.8)	184 (28.2)	4.2×10^{-3}	1.36 (1.1–1.7)	3.3×10^{-5}	1.33 (1.2–1.5)
Diffuse SSc	540 (71.1)	220 (28.9)	4.0×10^{-2}	1.23 (1.0–1.5)	211 (71.3)	65 (22)	0.69	1.06 (0.8–1.5)	0.05	1.18 (1–1.4)
Antibodies										
ACA	383 (66.7)	191 (33.3)	1.4×10^{-4}	1.51 (1.2–1.9)	302 (70.9)	124 (29.1)	4.6×10^{-3}	1.42 (1.1–1.8)	2.2×10^{-6}	1.47 (1.2–1.7)
ATA	252 (75.0)	84 (25.0)	0.96	1.01 (0.8–1.3)	151 (73.3)	55 (26.7)	0.17	1.26 (0.9–1.8)	0.37	1.10 (0.9–1.4)
ARA	260 (70.7)	108 (29.3)	0.08	1.26 (0.96–1.6)	–	–	–	–	–	–

*The odds ratios and 95% confidence intervals (95% CI) are for the carriage of the minor genotype.
Control subjects are used as reference for all comparisons.
ACA = Anti-Centromere Antibody; ATA = Anti-Topoisomerase I Antibody; ARA = Anti-RNA Polymerase III Antibody.

subsets of SSc but had the highest increase in the ACA autoantibody subset.

The independent case–control series of Spanish samples also demonstrated an increase in the "TT" genotype in SSc cases (7.3%) as compared to controls (5.3%) with a $P = 3.0 \times 10^{-2}$. The "T" allele frequency was also increased in the Spanish SSc cases (26.9%) as compared to the controls ($P = 8.0 \times 10^{-3}$, OR 1.27. 95 %CI 1.1–1.5). and in both the limited skin subset and ACA autoantibody subset of SSc ($P = 4.2 \times 10^{-3}$, OR 1.36, 95% CI 1.1–1.7 and $P = 4.6 \times 10^{-3}$, OR 1.42, 95% CI 1.1–1.8, respectively). The combined Mantel–Haenszel analysis of the two case–control series confirmed the association with SSc. This analysis also emphasized the strongest association of this variant with the ACA-positive subset of SSc. Logistic regression analysis revealed an additive model as the best fit model. The "TT" and "CT" genotypes were independent risk factors for SSc after controlling for the confounding effects of gender and the two case–control series with an OR 1.71 [95% CI 1.2–2.4] and OR 1.31 [95% CI 1.1–1.5], respectively (Fig. 1 D–F). The ACA positive autoantibody subset of SSc demonstrated the highest risk by logisitic regression analysis with an OR = 2.27 [95% CI 1.5–3.5] for "TT" genotype and an OR = 1.6 [95% CI 1.2–2.0] for "CT" genotype. These data demonstrate a role of rs2736340 variant in the risk for SSc.

3.3. *C8orf13-BLK gene region variant – rs13277113 association analysis*

On genotypic comparisons of the rs13277113 variant with SSc and its subsets the "AA" genotype frequency was significantly increased in the SSc cases (10.3%) as compared to controls (6.6%) in the North American series with a $P = 2.3 \times 10^{-3}$ (Table 3). The "AA" genotype was significantly increased in the limited and diffuse skin subsets and also in the ACA and ARA autoantibody subsets of SSc. No significant association was observed with the ATA autoantibody subset of SSc with this variant. The "A" allele was present in 30.1% of SSc cases as compared to 24.6% in controls with a $P = 3.6 \times 10^{-4}$, OR 1.32 [95% CI 1.1–1.6]. The "A" allele was also significantly increased in the limited and diffuse skin subsets. Both genotypic and allelic comparisons revealed no statistically significant association with SSc or any of its subsets for the rs13277113 variant in the Spanish series. However, on combining the two series by Mantel–Haenszel analysis the association with SSc was reproduced (genotypic- $P = 2.5 \times 10^{-3}$, OR 1.25 [95% CI 1.1–1.4] and allelic- $P = 2.0 \times 10^{-3}$, OR 1.20 [95% CI 1.1–1.3]). This analysis also highlighted the strongest association of this variant with ACA-positive subset of SSc. Logistic regression analysis showed an additive model as the best fit model. The "AA" and "AG" genotypes were independent risk factors for SSc after controlling for the confounding effects of gender and the two case–control series with an OR 1.57 [95% CI 1.1–2.2] and OR 1.23 [95% CI 1.04–1.5], respectively (Fig. 1 A–C).

3.4. *Functional role of C8orf13-BLK gene region variants*

To further elucidate the role of *C8orf13-BLK* gene region variants in SSc, we analyzed the whole blood gene expression profile of SSc patients from our gene expression data repository. Gene expression and genotyping data were available on 15 SSc patients of North American European descent. Eight of these were heterozygous for the rs13277113–rs2736340 variants ("AG"/"CT") and the other seven carried the wildtype genotype ("GG"/"CC"). A comparison of SSc patients with heterozygous genotype versus wildtype genotype demonstrated 582 differentially expressed transcripts at $P < 0.05$. The top 20 genes differentially expressed are ranked according to *p*-value in Table 4A. The top ten upregulated transcripts by fold change included *FLT3* which is involved in B cell lymphopoiesis and the top ten downregulated transcripts included *SH2D2A*, which is involved in normal differentiation and activation of T cells (Table 4B). The *BLK* mRNA transcript level was −1.14 fold lower in the heterozygotes as compared to wildtypes but the difference was not significant at $P < 0.05$. Similarly, the *C8orf13* mRNA transcript was increased 1.13 fold in the heterozygotes as compared to wildtypes but the difference was not significant at $P < 0.05$ either. Modeling of these data in the IPA canonical pathway knowledge base revealed the B cell receptor signaling pathway as the second most significantly dysregulated pathway with a $P = 1.59 \times 10^{-14}$ (Figs. 2A, 3A) in heterozygous individuals. The NFκB signaling pathway was the most significantly dysregulated with a $P = 9.5 \times 10^{-8}$ in the IPA molecular toxicology knowledge base (Figs. 2B, 3B). These data suggest a role for BLK in regulating BCR signaling and downregulation of NFκB pathway based on the risk heterozygote haplotype of *C8orf13-BLK*.

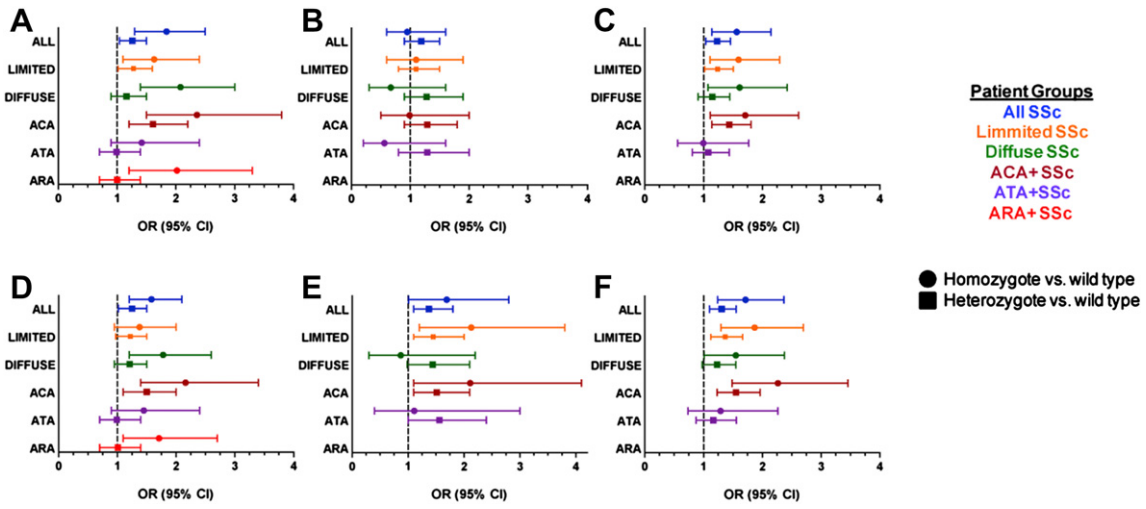

Fig. 1. Estimated risk of rs2736340 (A–C) and rs13277113 (D–F) variants in SSc patients versus controls, by logistic regression analysis. A, D – North American samples controlling for the confounding effects of gender. B, E – Spanish samples controlling for the confounding effects of gender. C, F – Combined samples controlling for the confounding effects of gender and two case–control series. SSc = systemic sclerosis; OR = odds ratio; 95% CI = 95% confidence interval. ACA = anti-centromere antibody; ATA = anti-topoisomerase I antibody; ARA = anti-RNA polymerase III antibody. Control subjects are used as reference for all comparisons. Both variants showed an additive mode of inheritance.

Table 3
Genotypic and allelic distribution of the rs13277113 variant in North American and Spanish SSc patients and healthy controls.*

	North American				Spanish				MH combined	
	N (%)				*N* (%)					
	AA	AG	GG	*P*-value	AA	AG	GG	*P*-value	*P*-value	OR (95% CI)
Control subjects	46 (6.6)	248 (35.8)	398 (57.5)		50 (7.1)	248 (35.3)	404 (57.5)			
Patients with SSc	107 (10.3)	413 (39.7)	521 (50)	2.3×10^{-3}	38 (6.5)	226 (38.6)	322 (54.9)	0.48	2.5×10^{-3}	1.25 (1.1–1.4)
Limited SSc	58 (9.7)	252 (41.9)	291 (48.4)	2.9×10^{-3}	23 (7.1)	120 (37.3)	179 (55.6)	0.83	3.4×10^{-3}	1.28 (1.1–1.5)
Diffuse SSc	44 (11.4)	141 (36.6)	200 (51.9)	1.7×10^{-2}	7 (4.7)	61 (40.9)	81 (54.4)	0.31	3.0×10^{-2}	1.25 (1.0–1.5)
Antibodies										
ACA	31 (10.4)	135 (45.5)	131 (44.1)	3.7×10^{-4}	13 (6.1)	87 (40.8)	113 (53.1)	0.33	6.5×10^{-4}	1.42 (1.2–1.7)
ATA	13 (7.6)	61 (35.7)	97 (56.7)	0.91	4 (3.9)	43 (41.7)	56 (54.4)	0.27	0.80	1.03 (0.8–1.3)
ARA	24 (12.7)	66 (34.9)	99 (52.4)	2.3×10^{-2}	–	–	–	–	–	–
	A	G	*P*-value	OR (95% CI)	A	G	*P*-value	OR (95% CI)	*P*-value	OR (95% CI)
Control subjects	340 (24.6)	1044 (75.4)			348 (24.8)	1056 (75.2)				
Patients with SSc	627 (30.1)	1455 (69.9)	3.6×10^{-4}	1.32 (1.1–1.6)	302 (25.8)	870 (74.2)	0.57	1.05 (0.9–1.2)	2.0×10^{-3}	1.20 (1.1–1.3)
Limited SSc	368 (30.6)	834 (69.4)	5.8×10^{-4}	1.35 (1.1–1.6)	166 (25.8)	478 (74.2)	0.63	1.05 (0.8–1.3)	2.9×10^{-3}	1.23 (1.1–1.4)
Diffuse SSc	229 (29.7)	541 (70.3)	9.1×10^{-3}	1.30 (1.1–1.6)	75 (25.2)	223 (74.8)	0.89	1.02 (0.8–1.4)	2.6×10^{-2}	1.20 (1.0–1.4)
Antibodies										
ACA	197 (33.2)	397 (66.8)	8.1×10^{-5}	1.52 (1.2–1.9)	113 (26.5)	313 (73.5)	0.47	1.10 (0.8–1.4)	5.5×10^{-4}	1.32 (1.1–1.5)
ATA	87 (25.4)	255 (74.6)	0.74	1.05 (0.8–1.4)	51 (24.8)	155 (75.2)	0.99	1.00 (0.7–1.4)	0.8	1.03 (0.8–1.3)
ARA	114 (30.2)	264 (69.8)	2.8×10^{-2}	1.33 (1.02–1.7)	–	–	–	–	–	–

*The odds ratios and 95% confidence intervals (95% CI) are for the carriage of the minor genotype.
Control subjects are used as reference for all comparisons.
ACA = Anti-Centromere Antibody; ATA = Anti-Topoisomerase I Antibody; ARA = Anti-RNA Polymerase III Antibody.

4. Discussion

Herein, we report the association of SSc with two variants (rs13277113 and rs2736340) in the *C8orf13-BLK* gene region in case–control subjects from North Americans of European descent and confirm this association in a Spanish case–control series. Studying a combined 1639 SSc patients and 1416 healthy controls we were able to demonstrate an additive effect of these variants for susceptibility to SSc. Furthermore, pathway analysis using the whole blood mRNA gene expression arrays highlighted dysregulation of BCR signaling and NFκB signaling based on the risk haplotype of the *C8orf13-BLK* gene region.

The rs13277113 and rs2736340 variants are located between *C8orf13* and *BLK* genes, which have reading frames in the opposite directions. Previous work in SLE has shown that these two variants are not in LD with any known coding-region variants on either *C8orf13* or *BLK* (13). In our combined analyses of two independent cohorts we find associations of both the "T" allele of the rs2736340 variant and the "A" allele of the rs13277113 variant. After controlling for gender and the case–control series an additive model was the best fit model and the ACA subset of SSc showed the strongest risk for susceptibility to SSc. These findings place *C8orf13* and *BLK* gene region in the category of common autoimmunity susceptibility genes due to their association with SLE, rheumatoid arthritis and now SSc. This group of common autoimmune disease genes also includes *PTPN22*, *IRF5*, and *STAT4* [15,17,29]. These genetic data combined with immune studies that show the presence of a type I interferon signature as well as the importance of B-cells and autoantibodies in both SSc and SLE support the hypothesis that there are also overlapping immune alterations between SSc and SLE [30–36].

BLK expression is highly specific to B cell lineage. It is involved in both BCR signaling and B cell development. In murine models, *blk* transcripts are expressed in pro-B, pre-B, mature-B cell line but are undetectable at the plasma cell stage [19]. The rs13277113 variant in the SLE study showed an increased expression of *C8orf13* and decreased expression of *BLK* mRNA based on the risk haplotype [16].

Table 4a
Top genes differentially expressed by SSc peripheral blood cells between the rs2736340–rs13277113 heterozygotes versus homozygotes ranked by *P*-value.

	p-value	Fold-change	Gene Symbol	Gene Name
1	4.26×10^{-4}	1.6	BCAR3	Breast cancer anti-estrogen resistance 3
2	4.30×10^{-4}	1.9	BIRC8	Baculoviral IAP repeat-containing 8
3	5.70×10^{-4}	1.5	SENP2	SUMO1/sentrin/SMT3 specific peptidase 2
4	6.86×10^{-4}	0.6	AHNAK	AHNAK nucleoprotein
5	9.96×10^{-4}	1.2	RPA1	Replication protein A1, 70 kDa
6	1.18×10^{-3}	0.7	RNF126	Ring finger protein 126
7	1.53×10^{-3}	0.6	PEX14	Peroxisomal biogenesis factor 14
8	1.54×10^{-3}	1.9	UPK3A	Uroplakin 3A
9	1.56×10^{-3}	0.7	MGC15523	Hypthetical protein MGC15523
10	1.67×10^{-3}	1.4	XKR3	Kell blood group complex subunit-related family, member 3
11	1.82×10^{-3}	0.8	C20orf116	Chromosome 20 open reading frame 116
12	2.02×10^{-3}	0.6	C3AR1	Complement component 3a receptor 1
13	2.09×10^{-3}	2.1	CMTM2	CKLF-like MARVEL transmembrane domain containing 2
14	2.30×10^{-3}	0.7	DPP3	Dipeptidyl-peptidase 3
15	2.47×10^{-3}	0.6	SLC35A2	Solute carrier family 35 (UDP-galactose trnasporter), member A2
16	2.67×10^{-3}	1.9	SPTLC1	Serine palmitoyltransferase, long chain base subunit 1
17	3.01×10^{-3}	1.6	CKLF	Chemokine-like factor
18	3.03×10^{-3}	1.7	MS4A2	Fc fragment of IgE, high affinity I receptor
19	3.09×10^{-3}	0.8	ALDOB	Aldose B, fructose-bisphosphonate
20	3.23×10^{-3}	0.5	INSIG1	Insulin induced gene 1

Table 4b
Top genes differentially expressed by SSc peripheral blood cells between the rs2736340-rs13277113 heterozygotes versus homozygotes ranked by fold change.

	Fold Change	Gene Symbol	Gene Name
1	2.96	CPA3	carboxypeptidase A3 (mast cell)
2	2.58	FLT3	fms-related tyrosine kinase 3
3	2.56	DNM3	dynamin 3
4	2.46	WNT8B	wingless-type MMTV integration site family, member 8B
5	2.45	SPARC	secreted protein, acidic, cysteine-rich (osteonectin)
6	2.44	GGT7	gamma-glutamyltransferase 7
7	2.43	SLC1A7	solute carrier family 1 (glutamate transporter), member 7
8	2.42	CHRM2	cholinergic receptor, muscarinic 2
9	2.42	CXCL5	chemokine (C-X-C motif) ligand 5
10	2.40	NIPA2	non imprinted in Prader–Willi/Angelman syndrome 2
1	−2.43	TNNT1	troponin T type 1
2	−2.18	ALG2	asparagine-linked glycosylation 2, alpha-1,3-mannosyltransferase homolog
3	−2.07	CDKN1C	cyclin-dependent kinase inhibitor 1C
4	−2.01	TNNC2	troponin C type 2
5	−1.95	GNLY	granulysin
6	−1.91	CDCA7	cell division cucle associated 7
7	−1.91	LTC4S	leukotriene C4 synthase
8	−1.91	SH2D2A	T cell specific adpater protein TSAd
9	−1.90	EOMES	eomesodermin homolog

In the current report, we were able to compare only the heterozygotes with the wildtypes and observed similar trends with an increased C8orf13 expression and a decreased BLK expression based on the risk haplotype, albeit not statistically significant. BLK is downstream of the BCR but its cellular functions are poorly defined. Deficiency in *blk*, *fyn* and *lyn* (SFK deficient) reduces the number of pre-B cells and a substantial reduction in peripheral B cell numbers. It is intriguing that we observed that the BCR signaling pathway was the second most significantly dysregulated pathway based on comparing heterozygotes with the wildtypes on the pathway analysis. Another report suggests that BLK is a growth inhibitor signaling molecule and prevents apoptosis [37]. Consistent with this, in the current report "Death receptor signaling" was the fourth most dysregulated pathway on comparing heterozygotes with the wildtypes. Finally, SFKs play an important role in pre-BCR mediated NFκB activation and B cell development [23]. In B cells, the NFκB family of transcription factors are important regulators of immune and inflammatory responses [38]. SFK deficient mice completely abrogated pre-BCR mediated NFκB activation [23]. Thus, this dysregulation in BCR and NFκB signaling may alter the delicate immune balance and tolerance and predispose to autoimmunity. Interestingly, NFκB signaling pathway was the most dysregulated toxicological pathway on comparing heterozygotes with the wildtypes in the current report. Furthermore, all the functional NFκB transcription factors (p50, p52, RelA, RelB, c-Rel) were downregulated in the heterozygotes. Together the whole genome microarray expression data support a role for the *C8orf13-BLK* region variants in the regulation of B-cell function.

In summary, our study is the first to report an association of the *C8orf13-BLK* region variants with SSc in two independent case–control

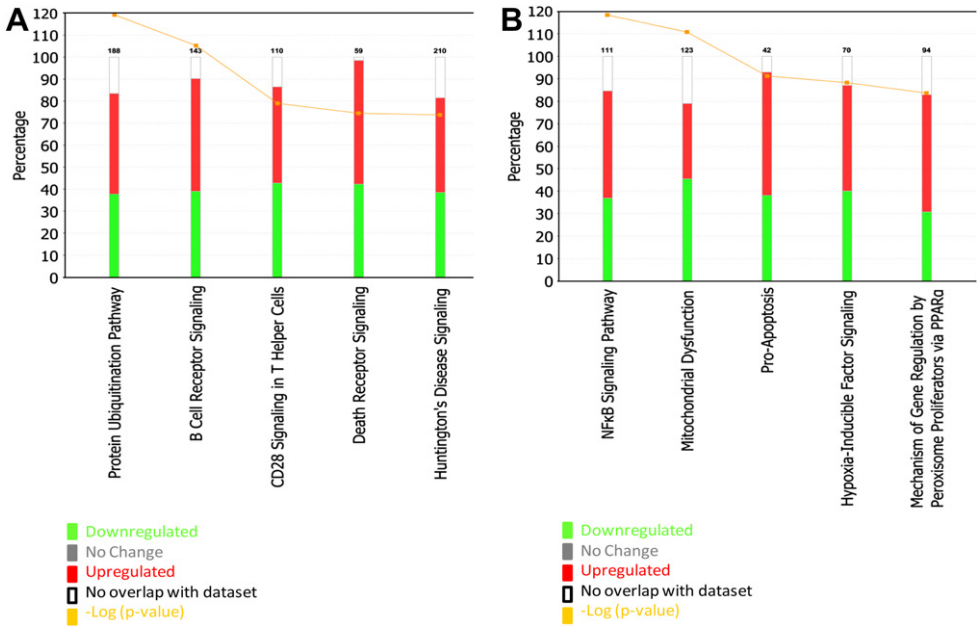

Fig. 2. Top pathways which discriminate SSc peripheral blood cells stratified by *C8orf13-BLK* variants modeling on the IPA canonical pathway knowledge base (A) and the IPA molecular toxicology knowledge base (B). Comparisons between the rs2736340-rs13277113 heterozygote versus homozygote in SSc were performed. Green represents downregulated, gray represents no change, red represents upregulated, and open represents no overlap with the data set. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

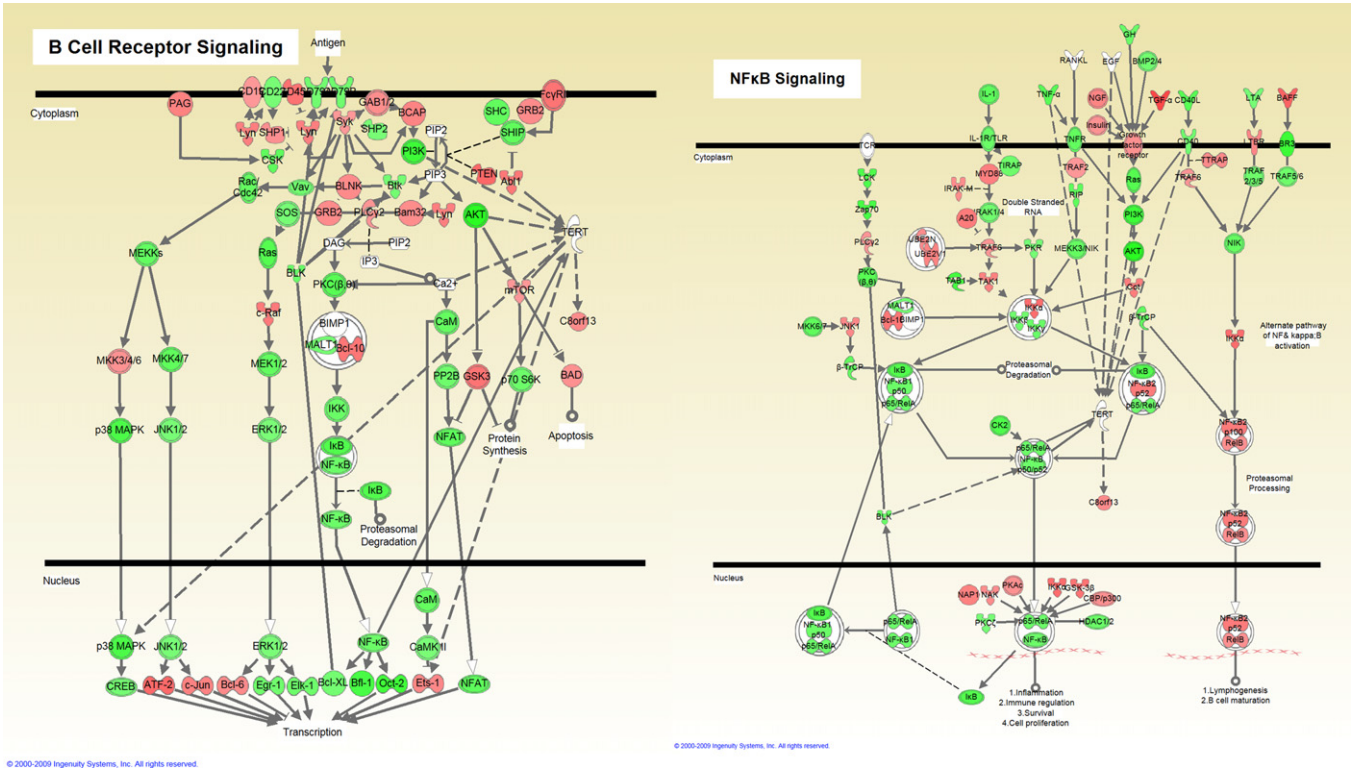

Fig. 3. Comparison of gene expression in peripheral blood cells stratified by the *C8orf13-BLK* variants rs2736340-rs13277113 heterozygote versus homozygote in SSc demonstrates alterations in B Cell Receptor Signaling Pathway (A) and NFκB Signaling Pathway (B). Red represents upregulated genes. Green represents downregulated genes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

series comprised of North Americans of European descent and Spanish SSc patients. These genetic variations may lead to the development of a more sophisticated molecular classification of SSc. Although additional genetic studies of SSc will be useful for replication of our data in other ethnic groups with SSc, resolving this issue will likely require comprehensive functional evaluation of these polymorphisms. Lastly, our data support the emerging paradigm that similar alterations in immune pathways may be involved in the pathogenesis of multiple autoimmune diseases.

Acknowledgments

This study was supported by the NIH/NIAMS Center of Research Translation (CORT) in Scleroderma (P50AR054144) (Dr. Arnett), the NIH/NIAMS Scleroderma Family Registry and DNA Repository (N01-AR-0-2251) (Dr Mayes), UTHSC-H Center for Clinical and Translational Sciences (Houston CTSA Program) (NIH/NCRR 3UL1RR024148) (Dr. Arnett). This study was also supported by grants SAF2006-00398, CTS-1180 and GENFER (Dr. Martin). B.R. is supported by the I3P CSIC program funded by the "Fondo Social Europeo". T.R. was sponsored by an unrestricted Grant from the "LandelijkKatholiekReuma Centrum (LKRC)", and VENI and VIDI laureate from the Dutch association of Research (NOW).

References

[1] Chizzolini C. T cells, B cells, and polarized immune response in the pathogenesis of fibrosis and systemic sclerosis. Curr Opin Rheumatol 2008;20:707–12.
[2] Hasegawa M, Fujimoto M, Takehara K, Sato S. Pathogenesis of systemic sclerosis: altered B cell function is the key linking systemic autoimmunity and tissue fibrosis. J Dermatol Sci 2005;39:1–7.
[3] Yoshizaki A, Iwata Y, Komura K, Ogawa F, Hara T, Muroi E, et al. CD19 regulates skin and lung fibrosis via Toll-like receptor signaling in a model of bleomycin-induced scleroderma. Am J Pathol 2008;172:1650–63.
[4] McHugh NJ, Whyte J, Artlett C, Briggs DC, Stephens CO, Olsen NJ, et al. Anti-centromere antibodies (ACA) in systemic sclerosis patients and their relatives: a serological and HLA study. Clin Exp Immunol 1994;96:267–74.
[5] Douvas AS, Achten M, Tan EM. Identification of a nuclear protein (Scl-70) as a unique target of human antinuclear antibodies in scleroderma. J Biol Chem 1979;254:10514–22.
[6] Okano Y, Steen VD, Medsger Jr TA. Autoantibody reactive with RNA polymerase III in systemic sclerosis. Ann Intern Med 1993;119:1005–13.
[7] Meyer OC, Fertig N, Lucas M, Somogyi N, Medsger Jr TA. Disease subsets, antinuclear antibody profile, and clinical features in 127 French and 247 US adult patients with systemic sclerosis. J Rheumatol 2007;34:104–9.
[8] Reveille JD, Fischbach M, McNearney T, Friedman AW, Aguilar MB, Lisse J, et al. Systemic sclerosis in 3 US ethnic groups: a comparison of clinical, socio-demographic, serologic, and immunogenetic determinants. Semin Arthritis Rheum 2001;30:332–46.
[9] Gourh P, Tan FK, Assassi S, Ahn CW, McNearney TA, Fischbach M, et al. Association of the PTPN22 R620W polymorphism with anti-topoisomerase I- and anticentromere antibody-positive systemic sclerosis. Arthritis Rheum 2006;54:3945–53.
[10] Alkassab F, Gourh P, Tan FK, McNearney T, Fischbach M, Ahn C, et al. An allograft inflammatory factor 1 (AIF1) single nucleotide polymorphism (SNP) is associated with anticentromere antibody positive systemic sclerosis. Rheumatology (Oxford) 2007;46:1248–51.
[11] Agarwal SK, Tan FK, Arnett FC. Genetics and genomic studies in scleroderma (systemic sclerosis). Rheum Dis Clin North Am 2008;34:17–40.
[12] Wu SP, Leng L, Feng Z, Liu N, Zhao H, McDonald C, et al. Macrophage migration inhibitory factor promoter polymorphisms and the clinical expression of scleroderma. Arthritis Rheum 2006;54:3661–9.
[13] Tan FK, Stivers DN, Arnett FC, Chakraborty R, Howard R, Reveille JD. HLA haplotypes and microsatellite polymorphisms in and around the major histocompatibility complex region in a Native American population with a high prevalence of scleroderma (systemic sclerosis). Tissue Antigens 1999; 53:74–80.
[14] Tan FK, Stivers DN, Foster MW, Chakraborty R, Howard RF, Milewicz DM, et al. Association of microsatellite markers near the fibrillin 1 gene on human chromosome 15q with scleroderma in a Native American population. Arthritis Rheum 1998;41:1729–37.
[15] Rueda B, Broen J, Simeon C, Hesselstrand R, Diaz B, Suarez H, et al. The STAT4 gene influences the genetic predisposition to systemic sclerosis phenotype. Hum Mol Genet 2009;18:2071–7.
[16] 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. N Engl J Med 2008;358:900–9.

[17] Harley JB, arcon-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.
[18] Ito I, Kawasaki A, Ito S, Hayashi T, Goto D, Matsumoto I, et al. Replication of the association between the C8orf13-BLK region and systemic lupus erythematosus in a Japanese population. Arthritis Rheum 2009;60:553–8.
[19] Dymecki SM, Zwollo P, Zeller K, Kuhajda FP, Desiderio SV. Structure and developmental regulation of the B-lymphoid tyrosine kinase gene blk. J Biol Chem 1992;267:4815–23.
[20] Dymecki SM, Niederhuber JE, Desiderio SV. Specific expression of a tyrosine kinase gene, blk, in B lymphoid cells. Science 1990;247:332–6.
[21] Drebin JA, Hartzell SW, Griffin C, Campbell MJ, Niederhuber JE. Molecular cloning and chromosomal localization of the human homologue of a B-lymphocyte specific protein tyrosine kinase (blk). Oncogene 1995;10:477–86.
[22] Islam KB, Rabbani H, Larsson C, Sanders R, Smith CI. Molecular cloning, characterization, and chromosomal localization of a human lymphoid tyrosine kinase related to murine Blk. J Immunol 1995;154:1265–72.
[23] Saijo K, Schmedt C, Su IH, Karasuyama H, Lowell CA, Reth M, et al. Essential role of Src-family protein tyrosine kinases in NF-kappaB activation during B cell development. Nat Immunol 2003;4:274–9.
[24] Tretter T, Ross AE, Dordai DI, Desiderio S. Mimicry of pre-B cell receptor signaling by activation of the tyrosine kinase Blk. J Exp Med 2003;198:1863–73.
[25] Arnett FC, Reveille JD, Goldstein R, Pollard KM, Leaird K, Smith EA, et al. Autoantibodies to fibrillarin in systemic sclerosis (scleroderma). An immunogenetic, serologic, and clinical analysis. Arthritis Rheum 1996;39:1151–60.
[26] Anonymous. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic criteria Committee. Arthritis Rheum 1980;23:581–90.
[27] Hummers LK, Hall A, Wigley FM, Simons M. Abnormalities in the regulators of angiogenesis in patients with scleroderma. J Rheumatol 2009;36:576–82.
[28] LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger Jr TA, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988;15:202–5.
[29] 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. N Engl J Med 2007;357:977–86.
[30] Arnett FC. Is scleroderma an autoantibody mediated disease? Curr Opin Rheumatol 2006;18:579–81.
[31] Tan FK, Zhou X, Mayes MD, Gourh P, Guo X, Marcum C, et al. Signatures of differentially regulated interferon gene expression and vasculotrophism in the peripheral blood cells of systemic sclerosis patients. Rheumatology (Oxford) 2006;45:694–702.
[32] York MR, Nagai T, Mangini AJ, Lemaire R, van Seventer JM, Lafyatis R. A macrophage marker, Siglec-1, is increased on circulating monocytes in patients with systemic sclerosis and induced by type I interferons and toll-like receptor agonists. Arthritis Rheum 2007;56:1010–20.
[33] Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proc Natl Acad Sci U S A 2003;100:2610–5.
[34] Graham RR, Kyogoku C, Sigurdsson S, Vlasova IA, Davies LR, Baechler EC, et al. Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. Proc Natl Acad Sci U S A 2007;104:6758–63.
[35] Peng SL. Altered T and B lymphocyte signaling pathways in lupus. Autoimmun Rev 2009;8:179–83.
[36] Cancro MP, D'Cruz DP, Khamashta MA. The role of B lymphocyte stimulator (BLyS) in systemic lupus erythematosus. J Clin Invest 2009;119:1066–73.
[37] Yao XR, Scott DW. Antisense oligodeoxynucleotides to the blk tyrosine kinase prevent anti-mu-chain-mediated growth inhibition and apoptosis in a B-cell lymphoma. Proc Natl Acad Sci U S A 1993;90:7946–50.
[38] Zwollo P, Rao S, Wallin JJ, Gackstetter ER, Koshland ME. The transcription factor NF-kappaB/p50 interacts with the blk gene during B cell activation. J Biol Chem 1998;273:18647–55.