REVIEW

Genetics of systemic sclerosis

Lara Bossini-Castillo[1] · Elena López-Isac[1] · Maureen D. Mayes[2] · Javier Martín[1]

© Springer-Verlag Berlin Heidelberg 2015

Abstract Systemic sclerosis (SSc) is connective tissue disorder in which fibrosis of the skin and internal organs is the main hallmark. Despite the difficulties of studying a complex disease, significant advances have been achieved in the SSc genetics field. In this review, we will describe the firmest genetic susceptibility markers known to date. We will analyze the most recent findings in the HLA region and in non-HLA genes. Furthermore, we will propose functional connections of these *loci* with the mechanisms involved in SSc pathogenesis. However, only non-HLA genetic regions that have been associated with SSc at the genome-wide significance level or that have been reported to be associated with the disease in at least two different independent studies will be considered. In spite of the increasing number of SSc genetic susceptibility factors identified, further studies with larger sample sizes, deeper phenotype characterization of the patients and innovative analyses will be needed to translate SSc genetics into clinical practice and patient care in the future.

This article is a contribution to the Special Issue on Immunopathology of systemic sclerosis – Guest Editors: Jacob M. van Laar and John Varga

✉ Javier Martín
martin@ipb.csic.es

[1] Consejo Superior de Investigaciones Científicas (IPBLN-CSIC), Instituto de Parasitología y Biomedicina López-Neyra, PTS Granda, Granada, Spain

[2] The University of Texas Health Science Center at Houston, Houston, USA

Introduction

Complex diseases, such as systemic sclerosis (SSc), are defined as those involving the interaction between genetic predisposition and environmental triggers. However, the environmental factors that influence the disease onset or prognosis of SSc are unknown. Moreover, from the genetics point of view, the number of relevant genes or the extent of their involvement in the disease pathogenesis have not been established yet.

The heritability of SSc was considered controversial in the largest published SSc twin study, which in general suggested a modest genetic contribution to the phenotype [1]. Nevertheless, this study included only 42 sets of twins, and it should be considered that, in a family study of 703 cases, an affected first-degree relative increased the risk of SSc 13 times compared to the general population [1, 2]. Moreover, having an affected sibling increased SSc risk by 15 times [2], and there was a remarkable concordance of autoantibodies between SSc twins [1]. Additionally, recent analyses have shown that the standardized incidence ratio of SSc seemed to be less than those observed in autoimmune diseases (ADs) such as rheumatoid arthritis or ankylosing spondylitis, but similar to those observed for Hashimoto thyroiditis or psoriasis [3]. In addition, SSc prevalence, clinical outcomes, and autoantibody profiles have been reported to vary depending on patient ancestry [4]. Therefore, the role of genetic factors in SSc susceptibility can now be considered solidly established.

Case-controls studies, in particular those based on single-nucleotide polymorphisms (SNPs), have provided a continuously growing panel of genetic players in the SSc pathogenic process. Genome-wide association studies (GWASs), which include hundreds of thousands of SNPs located throughout the genome, have been very effective in identifying a huge number of genetic *loci* associated with complex traits, including

autoimmune diseases (ADs) [5]. GWASs offered for the first time a hypothesis-blind approach to the analysis of complex traits. Moreover, analyses covering a great proportion of the variability of the human genome, especially in Caucasian cohorts, including large cohorts, became feasible for the first time. Unfortunately, SSc was not included in the first GWAS in ADs in 2007 [6], and until 2010, no SSc-related non-HLA *locus* had been established at the genome-wide significance level (*p* value<5×10^{-8}) [7] (Table 1). Although the initial genetic association reports in SSc were hardly reproducible, the coming of GWASs [7–11] and the recruitment of large patient cohorts have resulted in a growing number of firmly established SSc genetic susceptibility *loci*.

In this review, we will provide an updated overview of the known SSc-related genetic factors and we will address their possible functional implication in the pathogenic events that are characteristic of this chronic and disabling condition.

HLA

The human leukocyte antigen (HLA) was the first genetic association with SSc to be discovered and, similar to other ADs, it remains as a major autoimmunity genetic marker [12]. HLA-class II is expressed exclusively on antigen presenting cells and presents extracellular antigens. Considering that (1) the genes encoding the three major types of HLA class II molecules (HLA-DP, HLA-DQ, and HLA-DR) have been associated with SSc, (2) the associations with SSc in the HLA region are closely related to the autoantibodies in the patient sera, and (3) the reported remarkable variability between populations, we will address a comprehensive analysis of this region in the present section.

Classical HLA alleles

HLA associations have been studied in different ethnic populations, as discussed below, and have been shown to vary according to the presence/absence of the three most common SSc-related (and mutually exclusive) autoantibodies. The HLA-DQB1*0301 allele has been associated with the entire disease group of SSc patients in whites, blacks, and Hispanics [4], while HLA-DRB1*01, DRB1*04, and DQB1*0501 alleles have been related to anticentromere positive SSc patients (ACA+) and HLA-DRB1*11 and HLA-DPB1*1301 have been associated with antitopoisomerase positive SSc patients (ATA+), independently from their ethnic origin [4, 13, 14].

In white European ancestry populations, it has been described that the HLA-DRB1*1104/HLA-DQA1*0501/HLA-DQB1*0301 haplotype and the HLA-DQB1 alleles that encode a non-leucine residue at position 26 (DQB1 26 epi) predispose to develop SSc. On the other hand, the HLA-DRB1*0701-HLA-DQA1*0201-HLA-DQB1*0202

haplotype and the HLA-DRB1*1501 haplotype have a protective effect [14]. These HLA effects are also seen in Hispanic populations [14]. Furthermore, in whites, the HLA-DRB1*04 or HLA-DRB1*08 alleles and the HLA-DQB1*05 and HLA-DQB1*26 epi alleles cause an increased risk for ACA+ SSc [14, 15]. However, ATA+ SSc is associated with an over-representation of the HLA-DRB1*11 allele, the HLA-DRB1*1104/HLA-DQA1*0501/HLA-DQB1*0301 haplotype, and the HLA-DQB1*03 and HLA-DPB1*1301 alleles, or a reduced frequency of the HLA-DQB1*0501 allele [13–15]. Moreover, the HLA-DRB1*01 alleles have been described to predispose to ACA+ SSc and to protect against ATA+ SSc [15]. The anti-RNA polymerase III antibody positive SSc patient subset (ARA+) is defined by DRB1*0404, DRB1*11, and DQB1*03 in both white and Hispanic populations [14].

Regarding black populations, the HLA-DRB1*0804, HLA-DQA1*0501, and HLA-DQB1*0301 alleles are associated with the overall disease and HLA-DRB1*08 with ARA+ patients [14].

SNP-based analyses

As it was expected from candidate gene studies, GWASs confirmed the great contribution of the HLA region to the SSc genetic component. In the three GWASs carried out to date in SSc, the most prominent genetic association signal corresponded to the HLA class II region in chromosome 6 [7–11]. In the interrogated European descent populations, the highest peaks corresponded to the *HLA-DRB1* gene (rs6457617) and the *HLA-DQB1 locus* (haplotype block defined by rs9275224, rs6457617, and rs9275245), respectively [7, 11].

In addition, the Immunochip, a custom platform designed to provide a dense-mapping of different autoimmune-related chromosomal regions, has allowed the link between SNPs, classical HLA-alleles, and polymorphic amino acid positions in the HLA molecules [16, 17]. This platform, with complete coverage of the known variants in the extended HLA region, offers a suitable basis for novel imputation strategies that allow the deduction of the amino acids that better tag the associations pointed out by SNP-based analyses and that define the classical HLA alleles [16]. In the case of SSc, the differential associations of the classical HLA alleles with the main SSc serological subsets (ACA+ or ATA+) were further confirmed, since it was necessary to define a different HLA model for each serological subset to explain all the HLA associations observed in the whole disease analysis [18]. The ACA+ associations were best explained by certain residues in the 13th polymorphic position in *HLA-DRB1* and the 69th position in *HLA-DQA1*. On the other hand, the ATA+ subgroup associations were dependent on the 67th and 86th positions in *HLA-DRB1* and the positions 76th and 96th in *HLA-DPB1* [18]. Moreover, both the subtype and the whole disease groups

Table 1 Locus with or without significance level

Locus	Locus coordinates Chr:bp	Locus function	Associated genetic polymorphism/s	Risk/protection	References
Type I IFN pathway					
IRF5[a]	7:128,937,940-128,950,035	Virus-mediated activation of IFN	rs2004640, rs10954213, rs2280714, rs10488631, rs12537284, rs4728142, rs3757385	Risk	[7, 11, 18, 27, 28, 29]
IRF8[a]	16:85,899,168-85,922,606	Differentiation into monocyte precursor cells	rs11642873	Protection	[33]
IRF7	11:612,555-615,999	Regulates IFNα genes	rs4963128	Protection	[35]
IL-12 pathway					
IL12A[a]	3:159,988,836-159,996,019	Differentiation of Thl	rs77583790	Risk	[18]
IL12RB2[a]	1:67,307,364-67,396,900	IL-12 receptor β chain	rs3790567	Risk	[39]
IL12RB1[a]	19:18,058,995-18,098,986	IL-12 receptor β chain	rs2305743	Protection	[40]
STAT4[a]	2:191,029,576-191,172,671	T cell maturation and signaling	rs7574865, rs11889341, rs8179673, rs10181656, rs6752770, rs3821236	Risk	[7, 11, 18, 43, 44, 45, 46]
Debris clearance, autophagy, and detoxification					
DNASE1L3[a]	3:58,192,626-58,211,003	DNA breakdown during apoptosis	rs35677470	Risk	[18]
ATG5[a]	6:106,184,476-106,325,820	Autophagosome elongation	rs9373839	Risk	[18]
PPARG[a]	3:12,287,850-12,434,356	Regulator of adipocyte differentiation	rs310746	Risk	[56]
T cell-associated loci					
CD247[a]	1:167,430,640-167,518,610	Forms the TCR-CD3 complex	rs2056626	Protection	[7, 11, 59]
CSK[a]	15:74,782,084-74,803,198	Tyrosine kinase of the Src family	rs1378942	Risk	[62]
PTPN22	1:113,813,811-113,871,753	TCR signaling pathway phosphatase	rs2476601	Risk	[8, 9, 65]
B cell-associated loci					
BANK1	4:101,790,607-102,074,812	B cell receptor-induced calcium mobilization	rs10516487, rs3733197, rs17266594	Risk	[67, 68]
BLK	8:11,494,012-11,564,599	B cell receptor signaling and B cell development	rs2736340, rs13277113	Risk	[69, 70, 71]
TNF pathway and family					
TNFAIP3	6:137,867,188-137,883,314	Inhibits NF-kappa B activation and TNF-mediated apoptosis	rs5029939	Risk	[18, 54, 76]
TNIP1[a]	5:151,029,943-151,087,660	Regulation of nuclear factor kappa-B activation	rs2233287, rs4958881, rs3792783	Risk	[11, 78]
TNFSF4	1:173,183,731-173,207,332	T cell-APC interactions	rs1234314, rs2205960, rs844644, rs12039904, rs844665, rs844648	Risk	[81, 82]

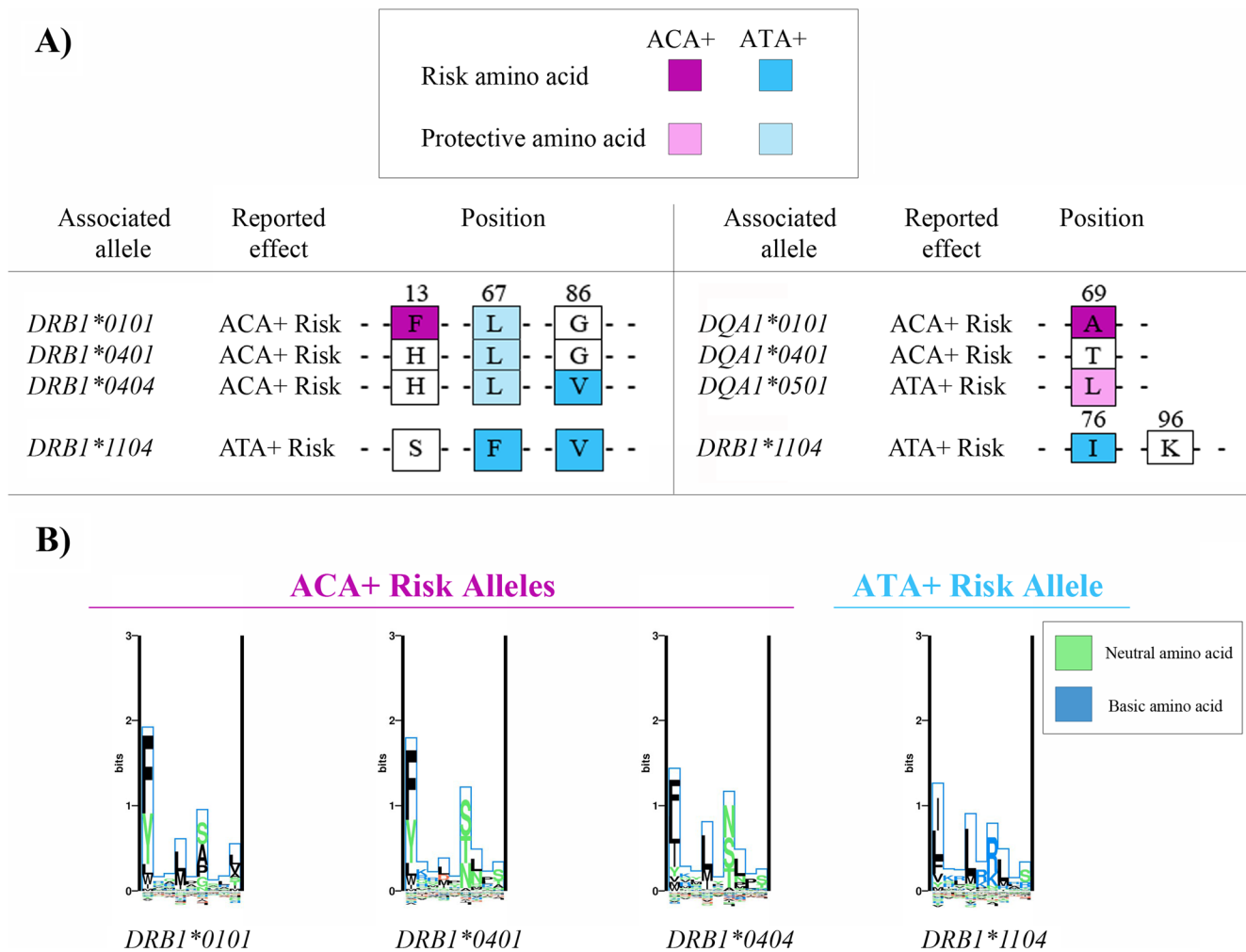[a] *Locus* that reached the genome-wide significance level

A)

			ACA+	ATA+
	Risk amino acid		■ (magenta)	■ (blue)
	Protective amino acid		■ (pink)	■ (light blue)

Associated allele	Reported effect	Position			Associated allele	Reported effect	Position
		13	67	86			69
DRB1*0101	ACA+ Risk	F	L	G	DQA1*0101	ACA+ Risk	A
DRB1*0401	ACA+ Risk	H	L	G	DQA1*0401	ACA+ Risk	T
DRB1*0404	ACA+ Risk	H	L	V	DQA1*0501	ATA+ Risk	L
							76 96
DRB1*1104	ATA+ Risk	S	F	V	DRB1*1104	ATA+ Risk	I K

B)

Fig. 1 **a** Previously reported classical four-digit alleles associated with ACA+ or ATA+ SSc patients and their relation with the amino acid model in Mayes et al. 2014. **b** HLA-DRB1 sequence logos for SSc-associated classical alleles

required the addition of seven SNPs to the model to condition all the HLA signals [18].

Mayes et al. underlined the relevance of the identified amino acids either in peptide binding groove or in the structure of the HLA proteins [18]. Indeed, as illustrated in Fig. 1, the classical HLA alleles that included the most highly associated risk amino acids for each subtype seemed to cause different amino acid binding preferences in the corresponding peptide binding grooves [19, 20]. In addition, the implication that the majority of the reported SNPs served as expression quantitative trait *loci* (eQTLs), that is, that their role was to modulate the expression of multiple HLA class I and class II genes, suggested novel functional implications for these selected variants [18, 21].

NON-HLA

In the following sections, we focus on the non-HLA genetic regions that have been associated with SSc at the genome-wide significance level and those that did not reach this

threshold but that have been associated with the disease in at least two different independent studies.

Type I interferon pathway

Type I interferon (IFN) responses induce antiviral immunity. These responses are initiated by a variety of pattern recognition receptors, such as Toll-like receptors (TLRs). IFNs modulate the amount of antigen presented to T cells, induce APC maturation, and increase natural killer (NK) cell activation [22]. Moreover, the balance and timing between IFNs and other stimulation pathways on T cells control the repression or promotion of T cell antiviral responses [22]. IFNβ is produced in any kind of cell infected by a virus, while APCs, especially plasmacytoid dendritic cells (pDCs), are the key producers of IFNα [22]. In the case of SSc, pDCs have also been shown to secrete CXCL4, which leads to endothelial-cell activation and increased responses of TLRs [23]. Furthermore, SSc has been classified as an "IFN signature" AD based on multiple lines of evidence that support a deregulation of the

type I IFN pathway [24]. Increased expression of IFN-regulated genes in peripheral whole blood cells, peripheral blood mononuclear cells, and sera of SSc patients have been described [24]. Regarding genetic findings, several IFN-related genes have been found to be associated with SSc.

Interferon regulatory factor 5 (*IRF5*) is expressed by macrophages, dendritic cells (conventional dendritic cells or cDCs) and pDCs as well as B cells. This protein promotes IFNα expression, and it is essential in defining an inflammatory macrophage lineage [25, 26]. Furthermore, IRF5 is involved in cell cycle and apoptosis, in microbial infection, and in inflammation [25, 26]. *IRF5* is a common autoimmunity susceptibility factor and also one of the major non-HLA gene associations with SSc [24]. Interestingly, the risk association of this *locus* with SSc has been reported in different ethnicities, and it reached the genome-wide significance level [7, 18, 27, 28]. While initial reports pointed towards a subtype specific association of this *locus* with patients with dcSSc or lung involvement [28, 29], novel reports show an association with the whole disease [7, 18]. In this line, Carmona et al. observed an additive effect of three distinct functional haplotype blocks (the haplotype tagged by rs10488631 is involved in protein stability, rs2004640 tags a splicing-related haplotype, and rs4728142 tags a haplotype in the 3′UTR that affects the expression of the gene) [30]. The *IRF5* haplotype structure was similar to the effect observed in SLE patients [30]. This genetic resemblance in the *IRF5 locus* between SSc, SLE, and other ADs has been confirmed recently using both the frequentist and the Bayesian approaches [31]. There is evidence that this *locus* affects the survival of SSc patients, which suggests that it may influence not just the susceptibility but also the severity of some SSc clinical complications such as lung involvement [32]. Remarkably, the minor allele of *IRF5* rs4728142, which is the most significant risk *IRF5* variant associated with SSc identified in the GWAS by Radstake et al., is associated with longer survival in SSc patients [7, 32].

IRF8 is another member of the IFN regulatory factor family that has been associated with SSc. This gene was identified as a genome-wide level SSc susceptibility factor in a phenotype-directed GWAS analysis by Gorlova et al. [33]. In this study, the minor allele of *IRF8* rs11642873 was less frequent in the lcSSc subset of patients than in the control group [33]. The encoded protein has a relevant role in the innate immune response carried out by myeloid cells and DCs and is also involved in B cell biology [34].

There is evidence of the association of an additional *locus* of this family, *IRF7*, with a protective effect for ACA+ SSc [35]. It is noteworthy that IRF7 is considered the master regulator of type I IFN antiviral immunity [36]. However, this association requires replication by independent groups.

Interleukin-12 pathway

Interleukin-12 is produced by phagocytes, APCs, and B cells after infection [37]. The main consequence of IL-12 production is the secretion of IFNγ, with the consequent proinflammatory effects. IL-12 controls T cell expansion towards a Th1 phenotype to the detriment of the Th2 compartment [37]. Moreover, it has been implicated in the development of auto-reactive Th1 cells in disease [37]. A variety of IL-12 pathway genes have been associated with ADs [38]. It has been suggested that ADs can be divided into two clusters based on their association with the IL-23 receptor (*IL23R*, which belongs to the same cytokine family but involved in Th17 amplification) or to *IL12A* (which encodes the IL-12 specific subunit, IL-12p35) [38]. This second group characterized by a key role of IL-12 and/or IL-35 would encompass SSc, which has several IL-12-related *loci* among its major genetic associations.

In fact, the rs77583790 rare variant, located upstream the *IL12A locus*, showed a genome-wide level risk association with SSc, and especially with the lcSSc subset in the recent Immunochip study by Mayes et al. [18]. Moreover, the coding genes for both chains in the IL-12 receptor (IL-12R), *IL12RB1* and *IL12RB2*, have been associated with SSc [39, 40]. Although the functional relevance of the *IL12RB2*-associated variants (rs3790567 was found to be the lead SNP) is unclear, the study by Bossini-Castillo et al. concluded that the observed signal was independent from the nearby *IL23R* gene [39]. This evidence supports the above-mentioned hypothesis of the relevance of IL-12 in SSc. In the case of *IL12RB1*, the results in López-Isac et al. point towards the rs436857 promoter SNP as the most plausible causal variant for the region [40]. The minor allele of rs436857 was a protective variant, and *in silico* analysis correlated this allele with a lower expression of IL12RB1, concordant with a decrease in SSc risk due to a lower response to IL-12 driven inflammation [40].

The *STAT4 locus* is another clear example of the IL-12 predominance in SSc. *STAT4* is a well-known autoimmunity susceptibility genetic factor, which encodes a transcription factor that plays a central role in IL-12 triggered inflammation [41]. Additionally, type I IFN can directly activate *STAT4* which induces the production of IFNγ [42]. Therefore, this molecule can act as link between the innate and the adaptive immune responses [42]. In SSc, *STAT4* was soon established as an SSc risk factor by Rueda et al. and then confirmed in different independent cohorts of European and Asian origin [43–47]. It should be highlighted that *STAT4* reached the genome-wide significance level in the two SSc GWAS comprising white European individuals [7, 11]. Moreover, the Immunochip study helped to narrow down the associated region, but the functional basis for this association is still unknown. STAT4 knock-out mice seem to be protected from inflammation-driven fibrosis in the SSc bleomycin-induced

SSc model [48]. In addition, Dieudé et al. reported an effect of the *STAT4* risk variants on pulmonary fibrosis in SSc [44]. Therefore, this *locus* is considered as a promising therapeutic target both for SSc and other ADs [41, 49].

Debris clearance, autophagy, and detoxification

One of the most relevant advantages of carrying out well-powered GWASs or immune-focused studies, in which several *loci* are tested without a very stringent pre-selection criteria, is the generation of new hypothesis that propose mechanisms that may have been previously overlooked, but that contribute to the onset or progression of ADs. In this context, the recently published Immunochip study identified two new SSc genetic risk factors, *DNASE1L3* and *ATG5*, which included for the first time debris clearance and autophagy, respectively, as pathogenic mechanisms for SSc.

DNASE1L3, homologue to DNase I, is a single- and double-stranded DNA endonuclease expressed by liver and spleen macrophages [50, 51]. Mutations in the *DNASE1L3* gene have been associated with familial forms of SLE, presumably due to an impaired elimination of the detritus of apoptotic macrophages, which lead to the production of auto-antibodies and immune imbalance [51]. This hypothesis is in line with the association of *DNASE1L3* rs35677470 SNP with ACA+ SSc patients [18]. The signal in this variant reached a *p* value=4.25×10^{-31}, and it is the most significant non-HLA association with SSc described to date [18]. It should be noted that the reported variant encodes a non-synonymous change from arginine to cysteine in the 206th position of the protein, which abolishes the activity of the protein, probably due to an alteration in its tertiary structure [52]. Moreover, the same association was replicated in an independent Immunochip study in an Australian SSc cohort, confirming the relevance of this *locus* in SSc [53]. Furthermore, the described association in the *DNASE1L3* seemed to explain the observed association in the nearby *PXK locus*, identified by Martín et al. [54].

In the case of *ATG5*, involved in the elongation of autophagosomes, the observed association corresponds to an intronic region of unknown function, rs9373839 [18]. Autophagy has emerged as an important piece of the immune response process [55]. Autophagy-related molecules interact with the immune cells at different levels such as the following: T and B cell development and function, phagocytosis, antigen presentation, cytokine secretion, etc [55]. Consequently, the association of *ATG5*, with SSc introduces autophagy as a new area of research and drug target exploration for this condition.

In addition, one of the suggestive *loci* in Allanore et al., the *PPARG* (peroxisome proliferator-activated receptor gamma) gene, was confirmed in a later SSc meta-GWAS and replication by López-Isac and collaborators [11, 56]. The risk variant identified in this gene, the rs310746 SNP, did not reach the genome-wide significance level in the overall analysis, but it was confirmed in the replication phase [56]. Thus, it is possible that this gene, involved in the peroxisome detoxification system with clear implications in fatty acid metabolism, has a role in SSc. Remarkably, as pointed out by López-Isac et al., this molecule also has an antifibrotic effect and has been found to affect dermal fibrosis in the SSc bleomycin mouse model [56, 57].

T cell-associated loci

According to the evidence, the T cell compartment is a key component of SSc pathogenesis. These lymphocytes appear in fibrotic zones and show altered phenotype and numbers [58]. Therefore, it is not surprising that several genes expressed by T cells belong to the SSc genetic network.

In fact, *CD247*, the encoding gene for the ζ-chain of the T cell receptor (TCR), was identified as a novel SSc susceptibility factor in the first GWAS in European descent SSc patients [7]. The protective association of rs2056626 observed by Radstake et al. was independently replicated in European cohorts, which confirms the implication of the TCR modulation in the disease [7, 59]. On the other hand, the same signal was not observed in Chinese individuals, underlining the high influence of ancestry and the heterogeneity between patients in SSc [60].

The tyrosine-protein kinase CSK (or C-Src kinase or C-terminal Src kinase) is an AD genetic marker that is involved in the inactivation of the Src-family kinases, which participate in signaling cascades such as the TCR pathway, B cell signaling, and skin fibrosis [61, 62]. Interestingly, the rs1378942 SNP, which maps in a *CSK* intron, was identified in a GWAS follow-up study by Martin et al. as a variant associated with increased SSc risk [62].

Furthermore, the lymphoid tyrosine phosphatase (LYP) encoded by *PTPN22* can exert its negative regulation over the TCR activation only when it is separated from CSK [63]. Interestingly, a non-synonymous variant in *PTPN22*, known as C1858T, R620W, or rs2476601, has been associated with multiple ADs and prevents this protein-protein interaction [64]. Of note, a meta-analysis of several SSc cohorts showed that this variant, but not another AD-associated functional SNP (rs33996649), has an impact on SSc susceptibility [65].

B cell-associated loci

SSc is characterized by immune imbalance in which also B cells react in a pathogenic manner [66]. B cells are responsible not only for auto-antibody production, but for cytokine release that activates the immune response. *BANK1* (B cell scaffold protein with ankyrin repeats 1) was the first B cell marker gene to be associated with SSc. The association of this *locus*, despite being modest, has been proven to be consistent in different studies [18, 67, 68]. What is more, *BANK1* has been

reported to have additive effects with *STAT4*, *IRF5*, and with an additional B cell marker, *BLK* (B lymphocyte kinase) [67, 69]. The association in this *locus* is again modest, but consistent in different European and Asian cohorts [18, 69–72].

TNF pathway and family

Abnormal levels of TNFα in SSc patient sera, leucocytes, bronchoalveolar lavage fluid, and skin have been long reported [73]. Moreover, TNFα inhibitors have been suggested as possible treatment for SSc patients [74]. Thus, it is not surprising that several TNFα pathway genes have been reported to be associated with SSc.

TNFAIP3, also known as A20, inhibits the proinflammatory NF-κβ signaling after TNFα activation [75]. Moreover, TNFAIP3 is also involved in apoptosis, IRF activation in response to pathogens, and even in autophagy [75]. Several studies have addressed the association of this *locus* with SSc. Initially, the *TNFAIP3 locus* was found to be related to SSc and especially to its severe phenotypes (dcSSc and pulmonary involvement) [76]. Then, a peak of association in this region was replicated in the Immunochip study by Mayes et al. [18]. Furthermore, this gene has been found to be associated with polyautoimmunity in SSc patients, and it has been identified as a shared *locus* between SLE and SSc in pan-meta-GWAS reports including both conditions [54, 77].

TNIP1 encodes the TNFAIP3-interacting protein 1, which regulates TNFAIP3, the previously described TNF-induced NF-κβ pathway inhibitor. Allanore et al. reported for the first time the association of *TNIP1* with SSc in a GWAS [11]. The association of three highly linked SNPs in *TNIP1* (rs2233287, rs4958881, and rs3792783) as SSc risk factors was independently replicated by Bossini-Castillo et al., and this *locus* was reinforced as a genome-wide level genetic factor in the meta-analysis [78].

TNFSF4 is a costimulatory molecule of the TNF family, also known as OX40L. The binding of this ligand, expressed on activated APCs and endothelial cells, to its receptor (CD134 or OX40) promotes T cell and B cell proliferation and survival [79]. Polymorphisms in the *TNFSF4* gene have been found to be associated with ADs [80], including SSc. Several SNPs located in the *TNFSF4* promoter have been reported to be associated with SSc [81, 82]. However, controversial phenotype specific associations were reported in both studies. A later meta-analysis including a new cohort, confirmed an especially strong association of rs2205960 with the ACA+ subset of patients [83].

Conclusion

International collaboration has allowed the analysis of the genetic basis of SSc in powerful and reproducible studies.

Furthermore, the implementation of new genotyping platforms and innovative biocomputational and statistical methods have provided the scientific community with increasing numbers of identified *loci* and new insights into the relationships that connect them. Genetic evidence supports a key role of the immune system in SSc predisposition, particularly with the IFN type I and IL-12 pathways and the deregulation of several immune cell compartments. Nevertheless, future approaches including even larger cohorts, deep clinical characterization and longitudinal measures of the individuals, or integrative analyses of genomic, epigenomic, transcriptomic, and proteomic data (systems biology) will help to establish the pathogenic mechanisms that result in the onset and progression of SSc. These advances would be especially valuable in understanding the most severe clinical outcomes, such as lung involvement, which remain widely unexplained. Eventually, this knowledge would lead to the validation of SSc biomarkers, the selection of drug targets, and the development of precision medicine.

References

1. Feghali-Bostwick C, Medsger TA Jr, Wright TM (2003) Analysis of systemic sclerosis in twins reveals low concordance for disease and high concordance for the presence of antinuclear antibodies. Arthritis Rheum 48:1956–1963

2. Arnett FC, Cho M, Chatterjee S, Aguilar MB, Reveille JD, Mayes MD (2001) Familial occurrence frequencies and relative risks for systemic sclerosis (scleroderma) in three United States cohorts. Arthritis Rheum 44:1359–1362

3. Hemminki K, Li X, Sundquist J, Sundquist K (2009) Familial associations of rheumatoid arthritis with autoimmune diseases and related conditions. Arthritis Rheum 60:661–668

4. Reveille JD, Fischbach M, McNearney T, Friedman AW, Aguilar MB, Lisse J et al (2001) Systemic sclerosis in 3 US ethnic groups: a comparison of clinical, sociodemographic, serologic, and immunogenetic determinants. Semin Arthritis Rheum 30:332–346

5. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H et al (2014) The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 42:D1001–1006

6. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661-678

7. Radstake TR, Gorlova O, Rueda B, Martin JE, Alizadeh BZ, Palomino-Morales R et al (2010) Genome-wide association study

of systemic sclerosis identifies CD247 as a new susceptibility locus. Nat Genet 42:426–429

8. Gourh P et al (2006) Association of the PTPN22 R620 W polymorphism with anti-topoisomerase I- and anticentromere antibodypositive systemic sclerosis. Arthritis Rheum 54:3945–3953

9. Dieude P et al (2008) The PTPN22 620 W allele confers susceptibility to sysremic sclerosis: Wndings of a large case-control study of european Causckasians and a meta-analysis. Arthritis Rheum 58:2183–2188

10. Zhou X, Lee JE, Arnett FC, Xiong M, Park MY, Yoo YK et al (2009) HLA-DPB1 and DPB2 are genetic loci for systemic sclerosis: a genome-wide association study in Koreans with replication in North Americans. Arthritis Rheum 60:3807–3814

11. Allanore Y, Saad M, Dieude P, Avouac J, Distler JH, Amouyel P et al (2011) Genome-wide scan identifies TNIP1, PSORS1C1, and RHOB as novel risk loci for systemic sclerosis. PLoS Genet 7:e1002091

12. Fernando MM, Stevens CR, Walsh EC, De Jager PL, Goyette P, Plenge RM et al (2008) Defining the role of the MHC in autoimmunity: a review and pooled analysis. PLoS Genet 4:e1000024

13. Gilchrist FC, Bunn C, Foley PJ, Lympany PA, Black CM, Welsh KI et al (2001) Class II HLA associations with autoantibodies in scleroderma: a highly significant role for HLA-DP. Genes Immun 2:76–81

14. Arnett FC, Gourh P, Shete S, Ahn CW, Honey RE, Agarwal SK et al (2010) Major histocompatibility complex (MHC) class II alleles, haplotypes and epitopes which confer susceptibility or protection in systemic sclerosis: analyses in 1300 Caucasian, African-American and Hispanic cases and 1000 controls. Ann Rheum Dis 69:822–827

15. Beretta L, Rueda B, Marchini M, Santaniello A, Simeon CP, Fonollosa V et al (2012) Analysis of Class II human leucocyte antigens in Italian and Spanish systemic sclerosis. Rheumatology (Oxford) 51:52–59

16. Parkes M, Cortes A, van Heel DA, Brown MA (2013) Genetic insights into common pathways and complex relationships among immune-mediated diseases. Nat Rev Genet 14:661–673

17. Raychaudhuri S, Sandor C, Stahl EA, Freudenberg J, Lee HS, Jia X et al (2012) Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. Nat Genet 44:291–296

18. Mayes MD, Bossini-Castillo L, Gorlova O, Martin JE, Zhou X, Chen WV et al (2014) Immunochip analysis identifies multiple susceptibility loci for systemic sclerosis. Am J Hum Genet 94:47–61

19. Rapin N, Hoof I, Lund O, Nielsen M (2008) MHC motif viewer. Immunogenetics 60:759–765

20. Nielsen M, Lundegaard C, Blicher T, Peters B, Sette A, Justesen S et al (2008) Quantitative predictions of peptide binding to any HLA-DR molecule of known sequence: NetMHCIIpan. PLoS Comput Biol 4:e1000107

21. Albert FW, Kruglyak L (2015) The role of regulatory variation in complex traits and disease. Nat Rev Genet 16:197–212

22. Crouse J, Kalinke U, Oxenius A (2015) Regulation of antiviral T cell responses by type I interferons. Nat Rev Immunol 15:231–242

23. van Bon L, Affandi AJ, Broen J, Christmann RB, Marijnissen RJ, Stawski L et al (2014) Proteome-wide analysis and CXCL4 as a biomarker in systemic sclerosis. N Engl J Med 370:433–443

24. Wu M, Assassi S (2013) The role of type 1 interferon in systemic sclerosis. Front Immunol 4:266

25. Barnes BJ, Richards J, Mancl M, Hanash S, Beretta L, Pitha PM (2004) Global and distinct targets of IRF-5 and IRF-7 during innate response to viral infection. J Biol Chem 279:45194–45207

26. Ryzhakov G, Eames HL, Udalova IA (2015) Activation and function of interferon regulatory factor 5. J Interf Cytokine Res: Off J Int Soc Interf Cytokine Res 35:71–78

27. Ito I, Kawaguchi Y, Kawasaki A, Hasegawa M, Ohashi J, Hikami K et al (2009) Association of a functional polymorphism in the IRF5 region with systemic sclerosis in a Japanese population. Arthritis Rheum 60:1845–1850

28. Dieude P, Guedj M, Wipff J, Avouac J, Fajardy I, Diot E et al (2009) Association between the IRF5 rs2004640 functional polymorphism and systemic sclerosis: a new perspective for pulmonary fibrosis. Arthritis Rheum 60:225–233

29. Dieude P, Dawidowicz K, Guedj M, Legrain Y, Wipff J, Hachulla E et al (2010) Phenotype-haplotype correlation of IRF5 in systemic sclerosis: role of 2 haplotypes in disease severity. J Rheumatol 37:987–992

30. Carmona FD, Martin JE, Beretta L, Simeon CP, Carreira PE, Callejas JL et al (2013) The systemic lupus erythematosus IRF5 risk haplotype is associated with systemic sclerosis. PLoS One 8:e54419

31. Kottyan LC, Zoller EE, Bene J, Lu X, Kelly JA, Rupert AM et al (2015) The IRF5-TNPO3 association with systemic lupus erythematosus has two components that other autoimmune disorders variably share. Hum Mol Genet 24:582–596

32. Sharif R, Mayes MD, Tan FK, Gorlova OY, Hummers LK, Shah AA et al (2012) IRF5 polymorphism predicts prognosis in patients with systemic sclerosis. Ann Rheum Dis 71:1197–1202

33. Gorlova O, Martin JE, Rueda B, Koeleman BP, Ying J, Teruel M et al (2011) Identification of novel genetic markers associated with clinical phenotypes of systemic sclerosis through a genome-wide association strategy. PLoS Genet 7:e1002178

34. Wang H, Morse HC 3rd (2009) IRF8 regulates myeloid and B lymphoid lineage diversification. Immunol Res 43:109–117

35. Carmona FD, Gutala R, Simeon CP, Carreira P, Ortego-Centeno N, Vicente-Rabaneda E et al (2012) Novel identification of the IRF7 region as an anticentromere autoantibody propensity locus in systemic sclerosis. Ann Rheum Dis 71:114–119

36. Honda K, Yanai H, Negishi H, Asagiri M, Sato M, Mizutani T et al (2005) IRF-7 is the master regulator of type-I interferon-dependent immune responses. Nature 434:772–777

37. Trinchieri G (1997) Function and clinical use of interleukin-12. Curr Opin Hematol 4:59–66

38. van Wanrooij RL, Zwiers A, Kraal G, Bouma G (2012) Genetic variations in interleukin-12 related genes in immune-mediated diseases. J Autoimmun 39:359–368

39. Bossini-Castillo L, Martin JE, Broen J, Gorlova O, Simeon CP, Beretta L et al (2012) A GWAS follow-up study reveals the association of the IL12RB2 gene with systemic sclerosis in Caucasian populations. Hum Mol Genet 21:926–933

40. Lopez-Isac E, Bossini-Castillo L, Guerra SG, Denton C, Fonseca C, Assassi S et al (2014) Identification of IL12RB1 as a novel systemic sclerosis susceptibility locus. Arthritis Rheum 66:3521–3523

41. Liang Y, Pan HF, Ye DQ (2014) Therapeutic potential of STAT4 in autoimmunity. Expert Opin Ther Targets 18:945–960

42. Nguyen KB, Watford WT, Salomon R, Hofmann SR, Pien GC, Morinobu A et al (2002) Critical role for STAT4 activation by type 1 interferons in the interferon-gamma response to viral infection. Science 297:2063–2066

43. Rueda B, Broen J, Simeon C, Hesselstrand R, Diaz B, Suarez H et al (2009) The STAT4 gene influences the genetic predisposition to systemic sclerosis phenotype. Hum Mol Genet 18:2071–2077

44. Dieude P, Guedj M, Wipff J, Ruiz B, Hachulla E, Diot E et al (2009) STAT4 is a genetic risk factor for systemic sclerosis having additive effects with IRF5 on disease susceptibility and related pulmonary fibrosis. Arthritis Rheum 60:2472–2479

45. Gourh P, Agarwal SK, Divecha D, Assassi S, Paz G, Arora-Singh RK et al (2009) Polymorphisms in TBX21 and STAT4 increase the risk of systemic sclerosis: evidence of possible gene-gene interaction and alterations in Th1/Th2 cytokines. Arthritis Rheum 60:3794–3806

46. Tsuchiya N, Kawasaki A, Hasegawa M, Fujimoto M, Takehara K, Kawaguchi Y et al (2009) Association of STAT4 polymorphism

with systemic sclerosis in a Japanese population. Ann Rheum Dis 68:1375–1376

47. Yi L, Wang JC, Guo XJ, Gu YH, Tu WZ, Guo G et al (2013) STAT4 is a genetic risk factor for systemic sclerosis in a Chinese population. Int J Immunopathol Pharmacol 26:473–478

48. Avouac J, Furnrohr BG, Tomcik M, Palumbo K, Zerr P, Horn A et al (2011) Inactivation of the transcription factor STAT-4 prevents inflammation-driven fibrosis in animal models of systemic sclerosis. Arthritis Rheum 63:800–809

49. Barnes J, Agarwal SK (2011) Targeting STAT4 in systemic sclerosis: a promising new direction. Expert Rev Clin Immunol 7:445–448

50. Rodriguez AM, Rodin D, Nomura H, Morton CC, Weremowicz S, Schneider MC (1997) Identification, localization, and expression of two novel human genes similar to deoxyribonuclease I. Genomics 42:507–513

51. Al-Mayouf SM, Sunker A, Abdwani R, Abrawi SA, Almurshedi F, Alhashmi N et al (2011) Loss-of-function variant in DNASE1L3 causes a familial form of systemic lupus erythematosus. Nat Genet 43:1186–1188

52. Ueki M, Kimura-Kataoka K, Takeshita H, Fujihara J, Iida R, Sano R et al (2014) Evaluation of all non-synonymous single nucleotide polymorphisms (SNPs) in the genes encoding human deoxyribonuclease I and I-like 3 as a functional SNP potentially implicated in autoimmunity. FEBS J 281:376–390

53. Zochling J, Newell F, Charlesworth JC, Leo P, Stankovich J, Cortes A et al (2014) An Immunochip-based interrogation of scleroderma susceptibility variants identifies a novel association at DNASE1L3. Arthritis Res Ther 16:438

54. Martin JE, Assassi S, Diaz-Gallo LM, Broen JC, Simeon CP, Castellvi I et al (2013) A systemic sclerosis and systemic lupus erythematosus pan-meta-GWAS reveals new shared susceptibility loci. Hum Mol Genet 22:4021–4029

55. Bhattacharya A, Eissa NT (2013) Autophagy and autoimmunity crosstalks. Front Immunol 4:88

56. Lopez-Isac E, Bossini-Castillo L, Simeon CP, Egurbide MV, Alegre-Sancho JJ, Callejas JL et al (2014) A genome-wide association study follow-up suggests a possible role for PPARG in systemic sclerosis susceptibility. Arthritis Res Ther 16:R6

57. Wei J, Zhu H, Komura K, Lord G, Tomcik M, Wang W et al (2014) A synthetic PPAR-gamma agonist triterpenoid ameliorates experimental fibrosis: PPAR-gamma-independent suppression of fibrotic responses. Ann Rheum Dis 73:446–454

58. O'Reilly S, Hugle T, van Laar JM (2012) T cells in systemic sclerosis: a reappraisal. Rheumatology (Oxford) 51:1540–1549

59. Dieude P, Boileau C, Guedj M, Avouac J, Ruiz B, Hachulla E et al (2011) Independent replication establishes the CD247 gene as a genetic systemic sclerosis susceptibility factor. Ann Rheum Dis 70:1695–1696

60. Wang J, Yi L, Guo X, He D, Li H, Guo G et al (2014) Lack of Association of the CD247 SNP rs2056626 with systemic sclerosis in Han Chinese. Open Rheum J 8:43–45

61. Okada M (2012) Regulation of the SRC family kinases by Csk. Int J Biol Sci 8:1385–1397

62. Martin JE, Broen JC, Carmona FD, Teruel M, Simeon CP, Vonk MC et al (2012) Identification of CSK as a systemic sclerosis genetic risk factor through genome wide association study follow-up. Hum Mol Genet 21:2825–2835

63. Vang T, Liu WH, Delacroix L, Wu S, Vasile S, Dahl R et al (2012) LYP inhibits T-cell activation when dissociated from CSK. Nat Chem Biol 8:437–446

64. Fiorillo E, Orru V, Stanford SM, Liu Y, Salek M, Rapini N et al (2010) Autoimmune-associated PTPN22 R620W variation reduces phosphorylation of lymphoid phosphatase on an inhibitory tyrosine residue. J Biol Chem 285:26506–26518

65. Diaz-Gallo LM, Gourh P, Broen J, Simeon C, Fonollosa V, Ortego-Centeno N et al (2011) Analysis of the influence of PTPN22 gene polymorphisms in systemic sclerosis. Ann Rheum Dis 70:454–462

66. Yoshizaki A, Sato S (2015) Abnormal B lymphocyte activation and function in systemic sclerosis. Ann Dermatol 27:1–9

67. Dieude P, Wipff J, Guedj M, Ruiz B, Melchers I, Hachulla E et al (2009) BANK1 is a genetic risk factor for diffuse cutaneous systemic sclerosis and has additive effects with IRF5 and STAT4. Arthritis Rheum 60:3447–3454

68. Rueda B, Gourh P, Broen J, Agarwal SK, Simeon C, Ortego-Centeno N et al (2010) BANK1 functional variants are associated with susceptibility to diffuse systemic sclerosis in Caucasians. Ann Rheum Dis 69:700–705

69. Coustet B, Dieude P, Guedj M, Bouaziz M, Avouac J, Ruiz B et al (2011) C8orf13-BLK is a genetic risk locus for systemic sclerosis and has additive effects with BANK1: results from a large french cohort and meta-analysis. Arthritis Rheum 63:2091–2096

70. Ito I, Kawaguchi Y, Kawasaki A, Hasegawa M, Ohashi J, Kawamoto M et al (2010) Association of the FAM167A-BLK region with systemic sclerosis. Arthritis Rheum 62:890–895

71. Gourh P, Agarwal SK, Martin E, Divecha D, Rueda B, Bunting H et al (2010) Association of the C8orf13-BLK region with systemic sclerosis in North-American and European populations. J Autoimmun 34:155–162

72. Shu C, Du W, Mao X, Li Y, Zhu Q, Wang W et al (2014) Possible single-nucleotide polymorphism loci associated with systemic sclerosis susceptibility: a genetic association study in a Chinese Han population. PLoS One 9:e113197

73. Hugle T, O'Reilly S, Simpson R, Kraaij MD, Bigley V, Collin M et al (2013) Tumor necrosis factor-costimulated T lymphocytes from patients with systemic sclerosis trigger collagen production in fibroblasts. Arthritis Rheum 65:481–491

74. Murdaca G, Spano F, Contatore M, Guastalla A, Puppo F (2014) Potential use of TNF-alpha inhibitors in systemic sclerosis. Immunotherapy 6:283–289

75. Catrysse L, Vereecke L, Beyaert R, van Loo G (2014) A20 in inflammation and autoimmunity. Trends Immunol 35:22–31

76. Dieude P, Guedj M, Wipff J, Ruiz B, Riemekasten G, Matucci-Cerinic M et al (2010) Association of the TNFAIP3 rs5029939 variant with systemic sclerosis in the European Caucasian population. Ann Rheum Dis 69:1958–1964

77. Koumakis E, Giraud M, Dieude P, Cohignac V, Cuomo G, Airo P et al (2012) Brief report: candidate gene study in systemic sclerosis identifies a rare and functional variant of the TNFAIP3 locus as a risk factor for polyautoimmunity. Arthritis Rheum 64:2746–2752

78. Bossini-Castillo L, Martin JE, Broen J, Simeon CP, Beretta L, Gorlova OY et al (2013) Confirmation of TNIP1 but not RHOB and PSORS1C1 as systemic sclerosis risk factors in a large independent replication study. Ann Rheum Dis 72:602–607

79. Gough MJ, Weinberg AD (2009) OX40 (CD134) and OX40L. Adv Exp Med Biol 647:94–107

80. Ishii N, Takahashi T, Soroosh P, Sugamura K (2010) OX40-OX40 ligand interaction in T-cell-mediated immunity and immunopathology. Adv Immunol 105:63–98

81. Gourh P, Arnett FC, Tan FK, Assassi S, Divecha D, Paz G et al (2010) Association of TNFSF4 (OX40L) polymorphisms with susceptibility to systemic sclerosis. Ann Rheum Dis 69:550–555

82. Bossini-Castillo L, Broen JC, Simeon CP, Beretta L, Vonk MC, Ortego-Centeno N et al (2011) A replication study confirms the association of TNFSF4 (OX40L) polymorphisms with systemic sclerosis in a large European cohort. Ann Rheum Dis 70:638–641

83. Coustet B, Bouaziz M, Dieude P, Guedj M, Bossini-Castillo L, Agarwal S et al (2012) Independent replication and meta analysis of association studies establish TNFSF4 as a susceptibility gene preferentially associated with the subset of anticentromere-positive patients with systemic sclerosis. J Rheum 39:997–1003