Inflammation and Regeneration

REVIEW Open Access

Updates on genetics in systemic sclerosis

Yuko Ota and Masataka Kuwana*

Abstract

Systemic sclerosis (SSc) is a complex disease, in which an interaction of genetic and environmental factors plays an important role in its development and pathogenesis. A number of genetic studies, including candidate gene analysis and genome-wide association study, have found that the associated genetic variants are mainly localized in noncoding regions in the expression quantitative trait locus and influence corresponding gene expression. The gene variants identified as a risk for SSc susceptibility include those associated with innate immunity, adaptive immune response, and cell death, while there are only few SSc-associated genes involved in the fibrotic process or vascular homeostasis. Human leukocyte antigen class II genes are associated with SSc-related autoantibodies rather than SSc itself. Since the pathways between the associated genotype and phenotype are still poorly understood, further investigations using multi-omics technologies are necessary to characterize the complex molecular architecture of SSc, identify biomarkers useful to predict future outcomes and treatment responses, and discover effective drug targets.

Introduction

Systemic sclerosis (SSc) is a complex autoimmune disease with heterogeneous clinical manifestations. The pathogenesis of SSc includes microvasculopathy, chronic inflammation and autoimmunity, and excessive fibrosis in the skin and internal organs, such as the lungs, heart, and gastrointestinal tract [1]. One of the autoimmune features is production of autoantibodies to various nuclear proteins, including centromere/kinetochore, topoisomerase I (topo I), and RNA polymerase III [2]. The primary event in the pathogenesis of SSc is thought to be endothelial injury, followed by aberrant vascular and immune dysregulation, leading to excessive tissue fibrosis [3]. The etiology of SSc is largely unknown, but accumulating evidence has shown that the combination of environmental and genetic factors contributes to the development and heterogeneous expression of the disease. Several environmental factors have been shown to correlate with increased SSc susceptibility, such as the exposure to certain chemical compounds, e.g., silica, organic solvents, dry cleaning detergents, vinyl chloride, and epoxy resin [4–6], and microorganisms such as cytomegalovirus [7]. Nevertheless, during the last decade, a number of genetic markers have been reported to be associated with SSc susceptibility and/or certain SSc subsets. However, the pathways between the associated genotype and phenotype as well as the interplay between the genetic risk and environmental triggers are still poorly understood. This review features updated knowledge of roles of genetic factors in susceptibility and disease expression of SSc.

Family association studies

Roles of the genetic background in susceptibility of SSc were first examined in familial association studies. Analysis of combined American cohorts involving 703 families found that SSc occurred significantly more frequently in families with SSc (1.6%) than in the general population (0.026%) [8]. In a follow-up study, affected first-degree relatives within multicase SSc families were concordant for SSc-related autoantibodies and human leukocyte antigen (HLA) class II haplotypes than expected by chance [9]. The heritability of the disease is often assessed by the disease concordance in monozygotic twins, and this strategy successfully demonstrated contribution of genetic backgrounds to susceptibility of systemic lupus erythematosus (SLE) [10] and

* Correspondence: kuwanam@nms.ac.jp
Department of Allergy and Rheumatology, Nippon Medical School Graduate School of Medicine, 1-1-5 Sendagi, Bunkyo-ku, Tokyo 113-8603-8582, Japan

rheumatoid arthritis (RA) [11]. The largest SSc twin study included 42 twin pairs, including 24 monozygotic and 18 dizygotic twins, and found that overall concordance of SSc was as low as 4.7%, and was similar in monozygotic and dizygotic twins [10]. This concordance rate was much lower than those in other autoimmune diseases, ~ 25% in SLE or RA. Consistency for the presence of anti-nuclear antibodies (ANAs) was significantly higher in monozygotic twins compared to dizygotic twins (90% versus 40%), suggesting that genetic background contributes mainly to autoantibody responses, but the underlying genetic backgrounds themselves are not sufficient for development of the disease [12]. In addition, families of SSc patients have an increased risk to develop other autoimmune diseases, such as autoimmune thyroid diseases and SLE, or some of SSc manifestations including Raynaud's phenomenon and interstitial lung disease (ILD), implicating shared genetic components between SSc and other autoimmune diseases [13].

Genetic association studies

The types of genetic variation include single-nucleotide polymorphism (SNP), variable number of tandem repeat (VNTR) or microsatellite, and copy-number variation (CNV). Of these, the most frequent genetic variation in human is the SNP, which potentially influences the protein function due to alteration in the amino acid sequence or modifies the gene expression. Two basic approaches used for genetic association studies include the candidate gene approach (CGA) and the genome-wide association study (GWAS). These approaches identify genetic variations and determine the likelihood that the variant occurs more or less frequently in the cases than in the controls. The associations are first tested in a discovery cohort and then are verified in a non-overlapping group of cases and controls for replication. The CGA is hypothesis-driven and is able to analyze all types of genetic variations, primarily based on associations with the potential pathogenic process and associations previously reported in other autoimmune diseases, although it is principally impossible to identify the gene loci that have not been assumed to be associated with the disease. Several genes of interest identified by the CGA in SSc patients are listed in Table 1 [14–41, 43]. The majority of genetic variations identified were SNPs, but there were some microsatellite or CNV associations. The genes analyzed were mainly those involved in innate and acquired immune responses or fibrosis.

Over the past 10 years, GWAS that scans the entire genome for SNPs provides powerful approach to analyze the genetic components of the polygenic diseases in hypothesis-free setting [44]. This strategy enables us to allow identification of new disease-related gene loci and

Table 1 SSc susceptibility genes identified by the CGA

Gene	Genetic polymorphism	Molecular function	References
PTPN22	SNP	T cell receptor (TCR) signaling	[14–16]
BANK1	SNP	B cell signaling	[17, 18]
CTGF	SNP	Fibroblast proliferation and production of extracellular matrix	[19]
FAM167A-BLK	SNP	B cell receptor signaling and B cell development	[20]
IRF5	SNP	TLR-dependent type I interferon production	[21, 22]
TNFAIP3	SNP	Negative feedback regulation of the NF-κB pathway	[26, 27]
STAT4	SNP	Induction of T helper 1 cells	[28]
FAS	SNP	Apoptosis of a wide variety of cell types	[31]
TBX21	SNP	T helper 1 cell differentiation	[32]
TNFSF4	SNP	Immune regulation	[15, 33]
TNIP1	SNP	Regulation of the NF-κB signaling pathway	[34]
IRAK1	SNP	Innate immune signaling	[35]
KCNA5	SNP	Potassium voltage-gated channel	[36]
TNFa13	VNTR	Modulator of inflammation	[37]
COL1A2	VNTR	Component of type I collagen	[38]
NOS2	VNTR	Production of a reactive free radical	[39]
CD19	VNTR	Regulation of B cell function	[40]
IRAK1	SNP	Mediator of innate immune response	[41, 42]
IL-21	SNP	Mediator of acquired immune response	[43]
IL-2RA	SNP	T cell activation	[43]

pathways in an unbiased manner. However, GWAS approach often miss unusual or rare variants since most GWAS covers up to 80% of common polymorphisms in the human genome. In addition, SNPs assessed by the GWAS are selected as landmarks of the surrounding SNPs in strong linkage disequilibrium. Therefore, additional analysis including whole genome sequencing, functional assays, and expression analysis in the affected and un-affected tissue is always required to identify the "true" SNPs responsible for disease pathogenesis. In early GWAS conducted in SSc patients, the strongest association identified was found in the HLA class II region on chromosome 6 [45], but it was difficult to identify the responsible gene loci due to the considerable variability in allele distribution among ethnic groups and the complex genetic structure of the HLA system. The GWAS approaches followed by replication studies and functional assay have led to identification of several non-HLA loci as an SSc susceptibility genes (Table 2) [42, 46–66]. It is of note that CGA is less reliable under the viewpoint of statistical significance, while identification of the primary associated variant in the locus is often difficult due to linkage disequilibrium. The majority of robustly replicated SSc susceptibility loci are involved in innate or adaptive immune system, and some were associated with cell death pathways. Additional immune-related genes responsible for SSc susceptibility were identified using the ImmunoChip array, another genotyping platform for SNP genotyping with high-density mapping of 196,524 variants across 186 known risk loci for autoimmune and inflammatory diseases in European Caucasian population [44]. These high-throughput genotyping studies found that most of the SSc-associated immune-related genes were shared among other autoimmune diseases, such as SLE and RA [67]. Interestingly, genes directly involved in the fibrotic process and/or vascular homeostasis were scarcely detected in the GWAS, although subsequent integration of multi-ethnic data and meta-analysis with increased sample size have revealed some candidate genes associated with fibrosis as the SSc-associated genes [52]. The hypothesis-free results from the GWAS support the hypothesis that the genetic background shared by many autoimmune diseases primarily contributes to dysregulated autoimmune responses in patients with SSc but suggest additional indispensable roles of environmental factors and epigenetic influences in the development of SSc.

Most of the genetic risk factors for SSc are located in intronic regions, rather than coding regions, and act as regulatory variants modulating the expression of nearby genes, i.e., transcription factor binding sites in expression quantitative trait locus (eQTL) [68]. A recent eQTL analysis in combination with GWAS data on SSc-associated genes successfully identified differentially regulated genes in the affected tissues in SSc patients and candidate genes potentially targeted by approved medications for immune-mediated diseases [69].

SSc susceptibility genes outside the HLA region

Genes involved in innate immunity

Type 1 interferon (IFN) is an important mediator of innate immunity often triggered by microbial infection. Over the past several years, there has been increasing evidence of dysregulation of the type 1 IFN pathway in autoimmune diseases, including SLE, dermatomyositis, and SSc [70–72]. Increased expression and activation of type I IFN-inducible genes termed "type I IFN signature" has been observed in peripheral blood and affected skin of SSc patients [70–72]. The GWAS identified transcription factors involved in regulation of type I IFN signaling, such as interferon regulatory factor (*IRF*) 4 [42], *IRF5* [42, 46–51], *IRF7* [52, 53], and *IRF8* [47, 50, 52–57]. Interestingly, these genes were associated with susceptibility of SLE and other autoimmune diseases [73–77]. Since all single markers within the *IRF5* loci failed to detect association signals, disease susceptibility could be regulated by the haplotype within the *IRF5* locus [78]. *IRF5* mediates induction of proinflammatory cytokines such as interleukin (IL)-6, IL-12, IL-23 and tumor-necrosis factor (TNF)-α and defines the phenotype of macrophages. In fact, macrophages carrying the *IRF5* risk allele haplotype have an increased expression of IRF5 protein and pattern recognition receptor-induced Akt2 activation, leading to proinflammatory cytokine production and M1 macrophage polarization [78]. The association between the *IRF5* genotype and SSc patients, especially those with anti-topo I-positive diffuse cutaneous SSc (dcSSc) with ILD, was first reported in French population by CGA [21] and was later replicated in independent studies [46]. One of the *IRF5* SNPs was shown to be useful in predicting a longer survival and preserved lung function [23]. A non-synonymous SNP located in the *IRF7* was associated with SSc with anticentromere antibody (ACA) in the USA and European cohorts [53] and was replicated in a meta-GWAS [52].

Another gene identified by the GWAS in SSc patients includes TNF-α-induced protein 3 (*TNFAIP3*), also known as A20 protein, which negatively regulates the TNF-induced nuclear factor (NF)-κB signaling pathway [26]. Three intronic risk variants that were linked to a decreased expression of A20 protein and one exonic variant were associated with SSc. The risk non-synonymous variant with an amino acid substitution was associated with reduction of activity of A20 protein. In this regard, the decrease of A20 expression by siRNA in foreskin fibroblasts resulted in an enhanced stimulation of collagen and α-smooth muscle

Table 2 Non-HLA SSc susceptibility genes identified by the GWAS

Gene	SNP	Study type	References
Innate immunity			
IRF4	rs9328192	GWAS	[42]
IRF5	rs4728142, rs10488631, rs10488631, rs3757385, rs109542313, rs2004640, rs12537284, rs2280714	GWAS	[42, 47–50, 46, 51]
IRF5-TNPO3	rs36073657, rs12155080	meta-GWAS	[52]
IRF7	rs1131665, rs4963128, rs702966	CGA, meta-GWAS	[52, 53]
IRF8	rs11642873, rs2280381, rs11117432, rs11644034, rs12711490, rs7202472, rs11117420	GWAS, meta-GWAS,	[47, 50, 53–56, 52, 57]
TNFAIP3	rs5029929, rs2230926, rs6932056	GWAS	[47, 50, 56, 58]
TNIP1	rs4958881, rs2233287, rs3792783	GWAS, meta-GWAS,	[47, 49, 52, 59]
TAP2	rs12538892, rs17500468	ImmunoChip	[51]
NFKB1	rs230534	meta-GWAS	[52]
Adaptive immune response			
TNFSF4	rs4916334, rs10798269, rs12039904	GWAS	[47, 50, 60]
TNFSF4-LOC100506023-PRDX6	rs2022449, rs1857066	meta-GWAS	[52]
CD247	rs2056626	GWAS, meta-GWAS	[48, 52, 54, 61]
CSK	rs1378942	GWAS, GWAS follow-up	[47, 52, 60]
PTPN22	rs2476601	GWAS	[42]
STAT4	rs7574865, rs3821236, rs4853458, rs10168266, rs3821236	GWAS, ImmunoChip, meta-GWAS	[46, 49, 48, 50, 52, 54, 55]
BLK	rs13277113, rs2736340	GWAS	[47]
IL-12 Signaling Pathway and cytokines			
IL-12A	rs7758790, rs589446	GWAS, ImmunoChip	[51, 52]
TYK2	rs2304256, rs34536443, rs12720356, rs35018800	ImmunoChip follow-up	[62]
IL-12RB1	rs436857, rs2305743, rs8109496, rs11668601	meta-GWAS	[52, 63]
IL-12RB2	rs3790566, rs924080, rs3790567	GWAS, meta-GWAS	[52, 64]
Apoptosis, Autophagy Pathway			
DNASEIL3	rs35677470	ImmunoChip	[51]
FLNB-DNASE1L3-PXK	rs7355798, rs4076852	meta-GWAS	[52]
ATG5	rs9373839, rs633724	GWAS, ImmunoChip, meta-GWAS	[42, 47, 51, 52]
PRDM1	rs4134466	GWAS	[55]
GSDMA	rs3894194	GWAS	[65]
GSDMB	rs883770	meta-GWAS	[52]
NOTCH4	rs443198	GWAS	[53]
Vascular homeostasis and fibrosis			
PPARG	rs310746	GWAS follow-up	[66]
Other			
NAB1	rs16832798	meta-GWAS	[52]
DDX6	rs11217020	meta-GWAS	[52]
DGKQ	rs11724804	meta-GWAS	[52]
POGLUT1-TIMMDC1-CD80-ARHGAP31	rs9884090	meta-GWAS	[52]
RAB2A-CHD7	rs6598008	meta-GWAS	[52]
TSPAN32, CD81-AS1	rs2651804	meta-GWAS	[52]
NUP85-GRB2	rs1005714	meta-GWAS	[52]

actin (α-SMA) gene expression after transforming growth factor-β (TGF-β) stimulation [79], suggesting that impaired A20 activity contributes to increased collagen production mediated by TGF-β. TNFAIP3-interacting protein 1 (*TNIP1*), which regulates TNFA IP3 activity, was also identified as the SSc-associated gene by GWAS in European population [49] and was replicated in a meta-GWAS [52].

Genes involved in adaptive immune response

TNF ligand superfamily member 4 (*TNFSF4*) encoding the T cell co-stimulatory molecule OX40 ligand was identified as the SSc-associated gene by the GWAS [47, 50, 60], and was replicated in CGA of a large European cohort. On the other hand, CD247 or zeta chain of the T cell receptor (TCR)/CD3 complex was identified as a susceptibility gene for SSc by the GWAS in European population [48], but this association was not replicated by CGA in Chinese cohort [80] and by a trans-ethnic meta-GWAS analysis [55]. Protein tyrosine phosphatase, non-receptor type 22 (*PTPN22*) was identified as the susceptibility gene for SSc by the GWAS [42] as well as for a wide range of autoimmune diseases [81]. This gene encodes the lymphoid tyrosine phosphatase (LYP), which directly interacts with c-Src kinase (CSK) and negatively regulates the TCR/CD3 complex signaling. The association of missense *PTPN22* allele detected in SSc patients disrupts the interaction of LYP with CSK and leads to an increased LYP activity [14–16]. *CSK* was also identified as the susceptibility gene associated with SSc by the GWAS [47, 52, 60]. Interestingly, CSK is known to function not only as a regulator of T cell activation, but also as a regulator of myofibroblast differentiation by modulating the function of Src kinase [82].

Signal transducer activator transcriptional factor 4 (*STAT4*) is one of susceptible genes for many autoimmune diseases [81] and is a transcription factor activated by a variety of cytokines, including type 1 IFN, IL-2, IL-12, IL-23, IL-27, and IL-35. STAT4 was first discovered to be crucial for promoting cellular-mediated immune responses via the differentiation of T helper 1 (Th1) cells through IFN-γ production [83], but is involved in a variety of inflammatory and immune processes. There were some contradictory findings in terms of the *STAT4* SNP associations with SSc susceptibility: association with limited cutaneous SSc (lcSSc), but not with dcSSc in European Caucasian cohorts [28, 46, 48, 49], while association with dcSSc and anti-topo I antibody in Chinese cohort [29]. Nevertheless, the recent GWAS meta-analysis and meta-GWAS confirmed the association of the responsible allele with overall SSc [46, 52, 55]. In addition, a large European study showed the additive effect of the *STAT4* and *IRF5* polymorphisms on susceptibility to SSc and SSc-related ILD [30]. The SSc-susceptible SNPs within the intron of *STAT4* locus are known to be eQTL. The STAT4 protein expression might control susceptibility of tissue fibrosis, since STAT4 knock-out mice were protective against bleomycin-induced dermal fibrosis [84]. The genes involved in B cell differentiation were also identified as SSc susceptible genes. B cell-specific scaffold protein with ankyrin 1 (*BANK1*) [15, 16] encodes a signaling molecule involved in B cell mobilization, and *BLK* encodes a tyrosine kinase crucial for B cell development and signaling. The association of *BANK1* SNPs in SSc susceptibility was revealed by two independent CGA studies [17, 18], and later confirmed by whole-exon sequencing [85]. The association of B lymphocyte kinase (*BLK*) was established in both European and Japanese populations [20, 22] and was confirmed by meta-analysis [86]. The trans-ethnic meta-analysis of GWAS identified B lymphocyte-induced maturation protein 1 (*PRDM1*) [55] as the SSc susceptibility gene.

Three genes in the IL-12 signaling pathway, including the intergenic region of *IL12A* [51, 52], IL-12 receptor B (*IL-12RB*)-1 [52, 63], and *IL-12RB2* [52, 64], were reported to be associated with SSc by GWAS, implying an important role of IL-12-mediated Th1 response in SSc pathogenesis. TYK2 encoding a tyrosine kinase member of the Janus kinase-STAT family, and mediates signaling of IL-12 family cytokines, such us IL-12 and IL-23, and is a common genetic risk factor for several autoimmune diseases, such as RA and SLE [87]. Meta-analysis of ImmunoChip analysis revealed that a common TYK2 missense variant was associated with SSc susceptibility [62]. On the other hand, SNPs within the *IL-21* gene, which plays a critical role in follicular helper T cell differentiation and germinal center formation, was shown to be associated with SSc in European/US Caucasian population by the CGA [43].

Genes involved in cell death

Deoxyribonuclease 1-like 3 (*DNASE1L3*) plays an important role in DNA fragmentation during apoptosis. The ImmunoChip analysis revealed that the non-synonymous *DNASE1L3* SNP, resulting in diminished DNase activity, was associated with SSc [51]. The GWAS follow-up study identified growth factor receptor-bound protein 10 (*GRB10*) as SSc susceptible gene [54]. GRB10 is an adaptor protein known to interact with a number of tyrosine kinase receptors and signaling molecules and has a potential role in apoptosis regulation. On the other hand, genes associated with autophagy were also detected by GWAS and ImmunoChip study. These included autophagy-related 5 (*ATG5*), which plays a role in assisting in autophagosome elongation and regulating lymphocyte maturation via autophagy [51], and Ras-related protein Rab-2A (*RAB2A*), which is involved in

autophagosome clearance [52]. These variants could impair proper functioning of autophagy, leading to endothelial cell stress pathways activation. A meta-GWAS analysis identified the association of SSc susceptibility with *GSDMA/B* encoding gasdermin A/B, which have a potential role in pyroptosis, a highly inflammatory cellular death [52, 55]. The gene encoding neurogenic locus notch homolog protein 4 (NOTCH4), which is involved in cell proliferation, differentiation, and apoptosis, was identified as a SSc susceptibility gene by the GWAS [53]. A recent CGA of large Caucasian and Chinese cohorts found associations of multiple NOTCH4 exonic variants with SSc and/or SSc subtypes [88].

Genes involved in vascular homeostasis and fibrosis

Only few SSc-associated genes involved in the vascular homeostasis and fibrotic process have been identified the SSc-susceptible gene. The GWAS follow-up analysis identified a SNP located upstream of the gene for the peroxisome proliferator-activated receptor gamma (PPARG) as one of SSc susceptible genes [66]. PPARG was initially identified in adipose tissue and was shown to be an anti-fibrotic effector through suppression of collagen synthesis, myofibroblast differentiation, and other TGF-β-induced fibrotic responses [89]. The genes whose molecular and cellular function has not been investigated in detail in mammalians, such as DDX6, NAB1, and DGKQ, have been identified by meta-GWAS study [52], but additional studies are necessary to clarify roles of these genes in the pathogenic process of SSc. *CAV1* was shown to be associated with SSc susceptibility by the CGA [90]. This gene encodes caveolin 1, which is an inhibitor of tissue fibrosis by suppressing TGF-β signaling. However, this association has not been replicated in independent studies. Dense microsatellite analysis of the HLA region in Japanese SSc patients identified a relationship between a rare variant of retinoid X receptor-beta (*RXRB*) and SSc patients with anti-topo I antibody on the risk haplotype harboring *HLA-DPB1*13:01* [91]. *RXRB* plays roles in anti-fibrotic activity through formation of a heterodimer with peroxisome proliferator-activated receptors and 9-cis retinoic acid ligands [92].

Roles of HLA gene polymorphisms in SSc susceptibility

Since immune dysregulation is one of characteristic pathogenic features of SSc [1], HLA has been examined extensively as one of potential genetic factors. Despite this, HLA associations with susceptibility to SSc were generally weak and inconsistent among studies, owing to diverse distribution of gene polymorphisms among ethnic groups [93]. Twelve different gene loci, including *HLA-B*, *C*, *DRA*, *DRB1*, *DRB5*, *DQA1*, *DQB1*, *DMB*, *DOA*, *DPA1*, *DPB1*, and *DPB2* were reported to be associated with SSc. Of these, *DRB1*, *DQA1*, *DQB1* and *DPB1* loci were extensively analyzed, but it was difficult to identify the responsible gene loci due to strong linkage disequilibrium. One of the most extensive studies enrolling 1300 SSc patients and 1000 controls with Caucasian, African, and Hispanic American backgrounds found that the associated HLA class II alleles were different among ethnic groups, and all associations were not robust [94]. The associations between susceptibility of SSc and the third hypervariable region (HVR) sequences of the *DRB1* gene were also investigated but were again borderline [95]. On the other hand, the HLA region (6p21.3), especially the *HLA-DPB1* and *DPB2*, was consistently identified as the gene region most strongly associated with SSc by GWAS, and this association was most prominent in SSc patients with anti-topo I antibody [45], but it remains controversial if the primarily associated genes were located within HLA or non-HLA genes.

A number of studies have found that HLA class II genes are associated with SSc-related autoantibodies rather than SSc itself, while the associated HLA class II alleles are different among ethnic groups [96]. The association of HLA class II alleles with SSc-related autoantibodies were analyzed most intensively for the anti-topo I antibody. It was reported that anti-topo I was associated with *DQB1*03*, which was thought to play the primary role in Caucasians [97]. On the other hand, we found that the *DRB1*15;02*, *DRB5*01;02*, *DQB1*06;01* haplotype was associated with anti-topo I-positive SSc in Japanese population [98]. Since the *DRB1/B5* alleles associated with anti-topo I in various ethnic groups, *DRB1*11;04* in Caucasians, *DRB1*08;04/*11;01* in American Africans, *DRB1*11;04/*08;02* in Hispanics, and *DRB5*01;02* in Japanese, have the common amino acid sequence FLEDR at amino acid positions at 67–71 in the hypervariable β1 domain of the *DRB* gene, we have proposed that the *DRB* is the primary gene associated with anti-topo I antibody in SSc patients (Fig. 1A). Full-blooded Choctaw Native Americans living in southeastern Oklahoma have the highest prevalence of SSc [99]. Anti-topo I antibody is the predominant autoantibody in this patent population and is associated with the unique Amerindian HLA haplotype containing *DRB1*1602*, which has the ILEDR sequence at amino acid positions at 67–71 [93]. Since antigen-presenting cells present an antigenic peptide to $CD4^+$ T cells in the context of HLA class II molecules, it is likely that the amino acid sequence FLEDR at positions 67–71 of the *DRB* gene located at the bottom of the antigen-binding groove controls antigen-specific $CD4^+$ T cell responses. To test this hypothesis, we examined in vitro T cell proliferative response to recombinant topo I fragments and found that the HLA-DR-restricted T cell response was found

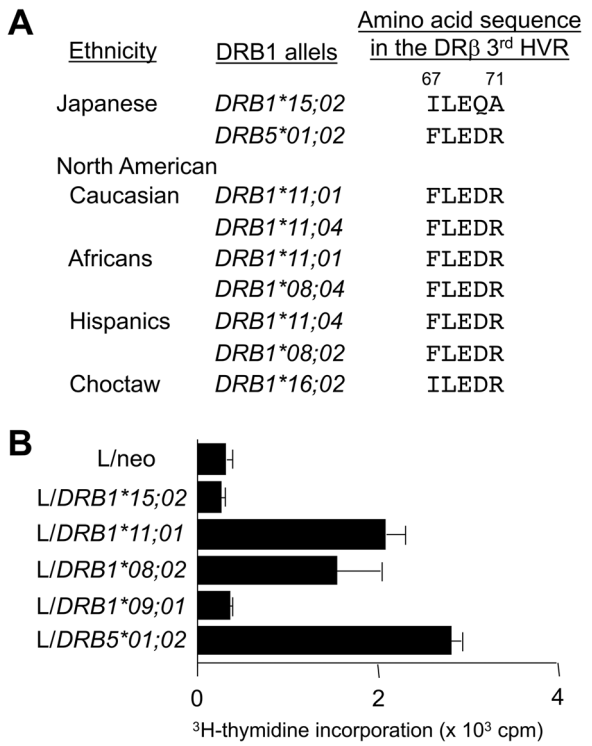

Fig. 1 **A** *DRB1* or *DRB5* alleles associated with the presence of anti-topo I antibody in various ethnicities and their amino acid sequence at positions 67–71 in the third hypervariable region (HVR). **B** Proliferative responses of a representative topo I-reactive $CD4^+$ T cell clone derived from an anti-topo I-positive patient with SSc in the presence of antigenic peptide-pulsed L cell transfectants expressing a series of human HLA-DR molecules harboring the *DRB1* or *DRB5* alleles associated with the presence of anti-topo I antibody. The peptide-induced T cell proliferation was measured by ^{3}H-thymidine incorporation

in SSc patients with anti-topo I as well as in heathy controls who had the *DRB* alleles with the FLEDR sequence [93]. We further established topo I-reactive $CD4^+$ T cell clones and examined their HLA class II restriction using a series of mouse L cell transfectants pulsed with the antigenic topo I peptide [100]. The antigen-induced response of the T cell clones was observed upon a co-culture with L cell transfectants expressing the DRA molecule in combination with DRB molecules harboring *DRB1*11;01*, **08;02*, or *DRB5*01;02* (Fig. 1B), indicating that the FLEDR sequence was critical to present the antigenic topo I peptide and resultant $CD4^+$ T cell activation, potentially leading to anti-topo I autoantibody production. A recent 3D structure models for prediction of HLA α/β heterodimers using the associated amino acid residues in the peptide-binding groove successfully identified immunodominant peptides of topo I [101]. On the other hand, previous studies have proposed that the *DPB1* locus is the primary susceptibility gene for anti-topo I-positive SSc within the HLA region and *DPB1*13:01* is strongly associated with SSc and anti-topo I antibody in various ethnic groups [97, 101]. The *DPB1*13:01* is linked with a variety of *DRB1* alleles with low linkage disequilibrium value, i.e., *DRB1*01;01, 15;01, 04;06, 11;01*, and *12;01* in Japanese, suggesting that association of *DPB1*13:01* with anti-topo I antibody-positive SSc is independent of the *DRB1/B5* allele association with production of anti-topo I antibody. Interestingly, the recent high-density microsatellite analysis of the HLA region identified *RXRB* as the responsible gene on the risk haplotype harboring *HLA-DPB1*13:01* in SSc patients with anti-topo I antibody in Japanese population [91].

Many lines of evidence have shown that individual SSc-related autoantibodies have associations with different HLA class II alleles and haplotypes, including ACA with the *DRB1*01;01, DQB1*05;01* haplotype, anti-RNA polymerase III with *DRB1*04;01/*04;04*, anti-PM-Scl with the *DRB1*03;01, DQB1*02;01* haplotype, and anti-U1RNP with the *DRB1*04;01, DQB1*03;02* haplotype [94, 98, 102]. These strong links could explain difference in prevalence of individual SSc-related autoantibodies among ethnic groups: anti-RNA polymerase III was more prevalent in cohorts from the UK, Northeast USA, and Australia, compared with other European countries and Japan [103], while anti-PM-Scl is almost exclusively found in Caucasian patients with SSc [104]. Conditional analysis in the autoantibody subsets of SSc revealed several associated amino acid residues, mostly in the peptide-binding groove of the HLA class II molecules [97]. It is interesting to note that bioinformatically predicted immunodominant peptides of topo I, fibrillarin, and centromere protein A are homologous to viral protein sequences, suggesting a possible link between HLA alleles, autoantibodies, and environmental triggers in the pathogenesis of SSc.

Gene expression profiling of the affected organ systems

The technologies to analyze gene expression profiling have enabled us to evaluate expression levels of comprehensive genes in peripheral blood, skin, and other affected tissues in patients with SSc. The gene expression analysis using microarray on skin biopsies from patients with SSc or morphea and healthy controls found four unique expression patterns in SSc patients: “fibro-proliferative”, “inflammatory”, “limited”, and “normal-like” [105]. The fibro-proliferative pattern comprised of patients with dcSSc, with the gene set associated with the biological processes of cell cycle. The inflammatory pattern was characterized by increased expression of a series of immune response genes. The limited pattern was predominantly found in lcSSc patients with a high expression of a distinct signature found heterogeneously across the samples. Lastly, the normal-like pattern had

increased expression of genes associated with fatty acid metabolism and lacked any expression associated with inflammation or proliferation. The gene expression profiling in the affected skin has called attention because this information is useful as biomarkers for predicting progression of the disease and treatment responses. For example, Hinchcliff et al. found that patients who responded to mycophenolate mofetil predominantly had inflammatory pattern, whereas all patients with the fibro-proliferative pattern were non-responders [106]. In a phase IIb, placebo-controlled, randomized clinical trial of abatacept in patients with early dcSSc, there was no statistically significant difference in changes of modified Rodnan total skin thickness score (mRSS) at 52 weeks in patients treated with abatacept compared with those treated with placebo [107]. Interestingly, in a subgroup analysis, abatacept significantly improved mRSS compared with placebo in patients with inflammatory pattern of gene expression in the skin, but not in those with other gene expression patterns.

Conclusions and future perspectives

Recent advances of technology of the genetic study such as GWAS have successfully identified a number of associations between genetic polymorphisms and SSc. A series of studies have suggested that dysregulated innate and adaptive immunity linked to genetic predisposition is involved in the pathogenic process of SSc but is insufficient to fully elicit microvasculopathy and excessive fibrosis, which are characteristics to SSc. We now know that estimated heritability of SSc is lower than other autoimmune diseases, such as RA and SLE, and contribution of environmental factors and epigenetic influences is more important in the development of SSc [108]. Acquired alteration in processes involved in DNA methylation and histone modification, and dysregulated miRNA network plays a critical role in the development of SSc, although the pathways linked between genetic factors and environmental triggers are still not fully understood. Therefore, the multi-omics analyses, including transcriptome, proteome, and metabolome will open up new avenues for improving understanding of the complex molecular architecture of SSc, predicting outcomes and treatment responses, and discovering new drug targets.

Abbreviations

SSc: Systemic sclerosis; topo I: Topoisomerase I; HLA: Human leukocyte antigen; RA: Rheumatoid arthritis; SLE: Systemic lupus erythematosus; ANA: Anti-nuclear antibody; ILD: Interstitial lung disease; SNP: Single-nucleotide polymorphism; VNTR: Variable number of tandem repeat; CNV: Copy-number variation; CGA: Candidate gene approach; GWAS: Genome-wide association study; eQTL: Expression quantitative trait loci; IFN: Interferon; IRF: Interferon regulatory factor; IL: Interleukin; TNF: Tumor-necrosis factor; dcSSc: Diffuse cutaneous systemic sclerosis; ACA: Anticentromere antibody; TNFAIP3: Tumor-necrosis factor-α-induced protein 3; α-SMA: α-Smooth muscle actin; TGF-β: Transforming growth factor-β; TNIP1: Tumor-necrosis factor-α-induced protein 3-interacting protein 1; TNFSF4: Tumor-necrosis factor ligand superfamily member 4; TCR: T cell receptor; PTPN22: Protein tyrosine phosphatase, non-receptor type 22; LYP: Lymphoid tyrosine phosphatase; CSK: c-Src kinase; STAT4: Signal transducer activator transcriptional factor 4; Th1: T helper 1; lcSSc: Limited cutaneous systemic sclerosis; IL-12RB: Interleukin-12 receptor B; IL-2RA: Interleukin 2 receptor subunit alpha; BANK1: B cell-specific scaffold protein with ankyrin 1; BLK: B lymphocyte kinase; PRDM1: B lymphocyte-induced maturation protein 1; DNASE1L3: Deoxyribonuclease 1-like 3; GRB10: Growth factor receptor-bound protein 10; ATG5: Autophagy-related 5; RAB2A: Ras-related protein Rab-2A; GSDMA: Gasdermin A; NOTCH4: Neurogenic locus notch homolog protein 4; PPARG: Peroxisome proliferator-activated receptor γ; CAV1: Caveolin 1; RXRB: Rare variant of retinoid X receptor-beta; mRSS: Modified Rodnan total skin thickness score; HVR: Hypervariable region

Acknowledgements

Not applicable

Authors' contributions

Y.O. and M.K. wrote the manuscripts. The authors read and approved the final manuscript.

Funding

Not applicable

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Received: 16 March 2021 Accepted: 24 May 2021
Published online: 15 June 2021

References

1. Varga J, Trojanowska M, Kuwana M. Pathogenesis of systemic sclerosis: recent insights of molecular and cellular mechanisms and therapeutic opportunities. J Scleroderma Relat Disord. 2017;2(3):137–52. https://doi.org/10.5301/jsrd.5000249.
2. Kuwana M. Circulating anti-nuclear antibodies in systemic sclerosis: utility in diagnosis and disease subsetting. J Nippon Med Sch. 2017;84(2):56–63. https://doi.org/10.1272/jnms.84.56.
3. Ota Y, Kuwana M. Endothelial cells and endothelial progenitor cells in the pathogenesis of systemic sclerosis. Eur J Rheumatol. 2020;7(Suppl 3):139–46. https://doi.org/10.5152/eurjrheum.2019.19158.
4. Makol A, Reilly MJ, Rosenman KD. Prevalence of connective tissue disease in silicosis (1985-2006)-a report from the state of Michigan surveillance system for silicosis. Am J Ind Med. 2011;54(4):255–62. https://doi.org/10.1002/ajim.20917.
5. Kettaneh A, Al Moufti O, Tiev KP, Chayet C, Tolédano C, Fabre B, et al. Occupational exposure to solvents and gender-related risk of systemic sclerosis: a metaanalysis of case-control studies. J Rheumatol. 2007;34(1):97–103.
6. Nicholson WJ, Henneberger PK, Seidman H. Occupational hazards in the VC-PVC industry. Prog Clin Biol Res. 1984;141:155–75.
7. Mora GF. Systemic sclerosis: environmental factors. J Rheumatol. 2009;36(11):2383–96. https://doi.org/10.3899/jrheum.090207.
8. Arnett FC, Cho M, Chatterjee S, Aguilar MB, Reveille JD, Mayes MD. Familial occurrence frequencies and relative risks for systemic sclerosis (scleroderma) in three United States cohorts. Arthritis Rheum. 2001;44(6):1359–62. https://doi.org/10.1002/1529-0131(200106)44:6<1359::AID-ART228>3.0.CO;2-S.
9. Assassi S, Arnett FC, Reveille JD, Gourh P, Mayes MD. Clinical, immunologic, and genetic features of familial systemic sclerosis. Arthritis Rheum. 2007;56(6):2031–7. https://doi.org/10.1002/art.22647.
10. Ulff-Møller CJ, Svendsen AJ, Viemose LN, Jacobsen S. Concordance of autoimmune disease in a nationwide Danish systemic lupus erythematosus

twin cohort. Semin Arthritis Rheum. 2018;47(4):538–44. https://doi.org/10.1016/j.semarthrit.2017.06.007.
11. Silman AJ, MacGregor AJ, Thomson W, Holligan S, Carthy D, Farhan A, et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. Br J Rheumatol. 1993;32(10):903–7. https://doi.org/10.1093/rheumatology/32.10.903.
12. Feghali-Bostwick C, Medsger TA Jr, Wright TM. Analysis of systemic sclerosis in twins reveals low concordance for disease and high concordance for the presence of antinuclear antibodies. Arthritis Rheum. 2003;48(7):1956–63. https://doi.org/10.1002/art.11173.
13. Arora-Singh RK, Assassi S, del Junco DJ, Arnett FC, Perry M, Irfan U, et al. Autoimmune diseases and autoantibodies in the first degree relatives of patients with systemic sclerosis. J Autoimmun. 2010;35(1):52–7. https://doi.org/10.1016/j.jaut.2010.02.001.
14. 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(12):3945–53. https://doi.org/10.1002/art.22196.
15. Dieudé P, Guedj M, Wipff J, Avouac J, Hachulla E, Diot E, et al. The PTPN22 620W allele confers susceptibility to systemic sclerosis: findings of a large case-control study of European Caucasians and a meta-analysis. Arthritis Rheum. 2008;58(7):2183–8. https://doi.org/10.1002/art.23601.
16. Diaz-Gallo LM, Gourh P, Broen J, Simeon C, Fonollosa V, Ortego-Centeno N, et al. Analysis of the influence of PTPN22 gene polymorphisms in systemic sclerosis. Ann Rheum Dis. 2011;70(3):454–62. https://doi.org/10.1136/ard.2010.130138.
17. Rueda B, Gourh P, Broen J, Agarwal SK, Simeon C, Ortego-Centeno N, et al. BANK1 functional variants are associated with susceptibility to diffuse systemic sclerosis in Caucasians. Ann Rheum Dis. 2010;69(4):700–5. https://doi.org/10.1136/ard.2009.118174.
18. Dieudé P, Wipff J, Guedj M, Ruiz B, Melchers I, Hachulla E, et al. BANK1 is a genetic risk factor for diffuse cutaneous systemic sclerosis and has additive effects with IRF5 and STAT4. Arthritis Rheum. 2009;60(11):3447–54. https://doi.org/10.1002/art.24885.
19. Fonseca C, Lindahl GE, Ponticos M, Sestini P, Renzoni EA, Holmes AM, et al. A polymorphism in the CTGF promoter region associated with systemic sclerosis. N Engl J Med. 2007;357(12):1210–20. https://doi.org/10.1056/NEJMoa067655.
20. Gourh P, Agarwal SK, Martin E, Divecha D, Rueda B, Bunting H, et al. Association of the C8orf13-BLK region with systemic sclerosis in North-American and European populations. J Autoimmun. 2010;34(2):155–62. https://doi.org/10.1016/j.jaut.2009.08.014.
21. Dieudé P, Guedj M, Wipff J, Avouac J, Fajardy I, Diot E, et al. Association between the IRF5 rs2004640 functional polymorphism and systemic sclerosis: a new perspective for pulmonary fibrosis. Arthritis Rheum. 2009;60(1):225–33. https://doi.org/10.1002/art.24183.
22. Ito I, Kawaguchi Y, Kawasaki A, Hasegawa M, Ohashi J, Kawamoto M, et al. Association of the FAM167A-BLK region with systemic sclerosis. Arthritis Rheum. 2010;62(3):890–5. https://doi.org/10.1002/art.27303.
23. Sharif R, Mayes MD, Tan FK, Gorlova OY, Hummers LK, Shah AA, et al. IRF5 polymorphism predicts prognosis in patients with systemic sclerosis. Ann Rheum Dis. 2012;71(7):1197–202. https://doi.org/10.1136/annrheumdis-2011-200901.
24. Carmona FD, Martin JE, Beretta L, Simeón CP, Carreira PE, Callejas JL, et al. The systemic lupus erythematosus IRF5 risk haplotype is associated with systemic sclerosis. PLoS One. 2013;8(1):e54419. https://doi.org/10.1371/journal.pone.0054419.
25. Dieude P, Dawidowicz K, Guedj M, Legrain Y, Wipff J, Hachulla E, et al. Phenotype-haplotype correlation of IRF5 in systemic sclerosis: role of 2 haplotypes in disease severity. J Rheumatol. 2010;37(5):987–92. https://doi.org/10.3899/jrheum.091163.
26. Dieudé P, Guedj M, Wipff J, Ruiz B, Riemekasten G, Matucci-Cerinic M, et al. Association of the TNFAIP3 rs5029939 variant with systemic sclerosis in the European Caucasian population. Ann Rheum Dis. 2010;69(11):1958–64. https://doi.org/10.1136/ard.2009.127928.
27. Koumakis E, Giraud M, Dieudé P, Cohignac V, Cuomo G, Airò P, et al. 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. 2012;64(8):2746–52. https://doi.org/10.1002/art.34490.
28. Rueda B, Broen J, Simeon C, Hesselstrand R, Diaz B, Suárez H, et al. The STAT4 gene influences the genetic predisposition to systemic sclerosis phenotype. Hum Mol Genet. 2009;18(11):2071–7. https://doi.org/10.1093/hmg/ddp119.
29. Yi L, Wang JC, Guo XJ, Gu YH, Tu WZ, Guo G, et al. STAT4 is a genetic risk factor for systemic sclerosis in a Chinese population. Int J Immunopathol Pharmacol. 2013;26(2):473–8. https://doi.org/10.1177/039463201302600220.
30. Dieudé P, Guedj M, Wipff J, Ruiz B, Hachulla E, Diot E, et al. STAT4 is a genetic risk factor for systemic sclerosis having additive effects with IRF5 on disease susceptibility and related pulmonary fibrosis. Arthritis Rheum. 2009;60(8):2472–9. https://doi.org/10.1002/art.24688.
31. Liakouli V, Manetti M, Pacini A, Tolusso B, Fatini C, Toscano A, et al. The -670G>A polymorphism in the FAS gene promoter region influences the susceptibility to systemic sclerosis. Ann Rheum Dis. 2009;68(4):584–90. https://doi.org/10.1136/ard.2008.088989.
32. Gourh P, Agarwal SK, Divecha D, Assassi S, Paz G, Arora-Singh RK, et al. 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. 2009;60(12):3794–806. https://doi.org/10.1002/art.24958.
33. Gourh P, Arnett FC, Tan FK, Assassi S, Divecha D, Paz G, et al. Association of TNFSF4 (OX40L) polymorphisms with susceptibility to systemic sclerosis. Ann Rheum Dis. 2010;69(3):550–5. https://doi.org/10.1136/ard.2009.116434.
34. Coustet B, Bouaziz M, Dieudé P, Guedj M, Bossini-Castillo L, Agarwal S, et al. 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 Rheumatol. 2012;39(5):997–1003. https://doi.org/10.3899/jrheum.111270.
35. Dieudé P, Bouaziz M, Guedj M, Riemekasten G, Airò P, Müller M, et al. Evidence of the contribution of the X chromosome to systemic sclerosis susceptibility: association with the functional IRAK1 196Phe/532Ser haplotype. Arthritis Rheum. 2011;63(12):3979–87. https://doi.org/10.1002/art.30640.
36. Wipff J, Dieudé P, Guedj M, Ruiz B, Riemekasten G, Cracowski JL, et al. Association of a KCNA5 gene polymorphism with systemic sclerosis-associated pulmonary arterial hypertension in the European Caucasian population. Arthritis Rheum. 2010;62(10):3093–100. https://doi.org/10.1002/art.27607.
37. Takeuchi F, Nabeta H, Füssel M, Conrad K, Frank KH. Association of the TNFa13 microsatellite with systemic sclerosis in Japanese patients. Ann Rheum Dis. 2000;59(4):293–6. https://doi.org/10.1136/ard.59.4.293.
38. Hata R, Akai J, Kimura A, Ishikawa O, Kuwana M, Shinkai H. Association of functional microsatellites in the human type I collagen alpha2 chain (COL1A2) gene with systemic sclerosis. Biochem Biophys Res Commun. 2000;272(1):36–40. https://doi.org/10.1006/bbrc.2000.2731.
39. Kawaguchi Y, Tochimoto A, Hara M, Kawamoto M, Sugiura T, Katsumata Y, et al. NOS2 polymorphisms associated with the susceptibility to pulmonary arterial hypertension with systemic sclerosis: contribution to the transcriptional activity. Arthritis research & therapy. 2006;8(4):R104. https://doi.org/10.1186/ar1984.
40. Tsuchiya N, Kuroki K, Fujimoto M, Murakami Y, Tedder TF, Tokunaga K, et al. Association of a functional CD19 polymorphism with susceptibility to systemic sclerosis. Arthritis Rheum. 2004;50(12):4002–7. https://doi.org/10.1002/art.20674.
41. Carmona FD, Cénit MC, Diaz-Gallo LM, Broen JC, Simeón CP, Carreira PE, et al. New insight on the Xq28 association with systemic sclerosis. Ann Rheum Dis. 2013;72(12):2032–8. https://doi.org/10.1136/annrheumdis-2012-202742.
42. Lopez-Isac E, Martin JE, Assassi S, Simeon CP, Carreira P, Ortego-Centeno N, et al. IRF4 newly identified as a common susceptibility locus for systemic sclerosis and rheumatoid arthritis in a cross-disease meta-analysis of genome-wide association studies. Arthritis Rheumatol. 2016;68(9):2338–44. https://doi.org/10.1002/art.39730.
43. Diaz-Gallo LM, Simeon CP, Broen JC, Ortego-Centeno N, Beretta L, Vonk MC, et al. Implication of IL-2/IL-21 region in systemic sclerosis genetic susceptibility. Ann Rheum Dis. 2013;72(7):1233–8. https://doi.org/10.1136/annrheumdis-2012-202357.
44. Parkes M, Cortes A, van Heel DA, Brown MA. Genetic insights into common pathways and complex relationships among immune-mediated diseases. Nat Rev Genet. 2013;14(9):661–73. https://doi.org/10.1038/nrg3502.
45. Zhou X, Lee JE, Arnett FC, Xiong M, Park MY, Yoo YK, et al. HLA-DPB1 and DPB2 are genetic loci for systemic sclerosis: a genome-wide association study in Koreans with replication in North Americans. Arthritis Rheum. 2009;60(12):3807–14. https://doi.org/10.1002/art.24982.
46. Xu Y, Wang W, Tian Y, Liu J, Yang R. Polymorphisms in STAT4 and IRF5 increase the risk of systemic sclerosis: a meta-analysis. Int J Dermatol. 2016;55(4):408–16. https://doi.org/10.1111/ijd.12839.
47. Martin JE, Assassi S, Diaz-Gallo LM, Broen JC, Simeon CP, Castellvi I, et al. A systemic sclerosis and systemic lupus erythematosus pan-meta-GWAS

reveals new shared susceptibility loci. Hum Mol Genet. 2013;22(19):4021–9. https://doi.org/10.1093/hmg/ddt248.
48. Radstake TR, Gorlova O, Rueda B, Martin JE, Alizadeh BZ, Palomino-Morales R, et al. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. Nat Genet. 2010;42(5):426–9. https://doi.org/10.1038/ng.565.
49. Allanore Y, Saad M, Dieudé P, Avouac J, Distler JH, Amouyel P, et al. Genome-wide scan identifies TNIP1, PSORS1C1, and RHOB as novel risk loci for systemic sclerosis. PLoS Genet. 2011;7(7):e1002091. https://doi.org/10.1371/journal.pgen.1002091.
50. Martin JE, Broen JC, Carmona FD, Teruel M, Simeon CP, Vonk MC, et al. Identification of CSK as a systemic sclerosis genetic risk factor through Genome Wide Association Study follow-up. Hum Mol Genet. 2012;21(12):2825–35. https://doi.org/10.1093/hmg/dds099.
51. Mayes MD, Bossini-Castillo L, Gorlova O, Martin JE, Zhou X, Chen WV, et al. Immunochip analysis identifies multiple susceptibility loci for systemic sclerosis. Am J Hum Genet. 2014;94(1):47–61. https://doi.org/10.1016/j.ajhg.2013.12.002.
52. López-Isac E, Acosta-Herrera M, Kerick M, Assassi S, Satpathy AT, Granja J, et al. GWAS for systemic sclerosis identifies multiple risk loci and highlights fibrotic and vasculopathy pathways. Nat Commun. 2019;10(1):4955. https://doi.org/10.1038/s41467-019-12760-y.
53. Carmona FD, Gutala R, Simeón CP, Carreira P, Ortego-Centeno N, Vicente-Rabaneda E, et al. Novel identification of the IRF7 region as an anticentromere autoantibody propensity locus in systemic sclerosis. Ann Rheum Dis. 2012;71(1):114–9. https://doi.org/10.1136/annrheumdis-2011-200275.
54. Gorlova O, Martin JE, Rueda B, Koeleman BP, Ying J, Teruel M, et al. Identification of novel genetic markers associated with clinical phenotypes of systemic sclerosis through a genome-wide association strategy. PLoS Genet. 2011;7(7):e1002178. https://doi.org/10.1371/journal.pgen.1002178.
55. Terao C, Kawaguchi T, Dieude P, Varga J, Kuwana M, Hudson M, et al. Transethnic meta-analysis identifies GSDMA and PRDM1 as susceptibility genes to systemic sclerosis. Ann Rheum Dis. 2017;76(6):1150–8. https://doi.org/10.1136/annrheumdis-2016-210645.
56. Terao C, Ohmura K, Kawaguchi Y, Nishimoto T, Kawasaki A, Takehara K, et al. PLD4 as a novel susceptibility gene for systemic sclerosis in a Japanese population. Arthritis Rheum. 2013;65(2):472–80. https://doi.org/10.1002/art.37777.
57. Arismendi M, Giraud M, Ruzehaji N, Dieudé P, Koumakis E, Ruiz B, et al. Identification of NF-κB and PLCL2 as new susceptibility genes and highlights on a potential role of IRF8 through interferon signature modulation in systemic sclerosis. Arthritis Res Ther. 2015;17(1):71. https://doi.org/10.1186/s13075-015-0572-y.
58. Gorlova OY, Li Y, Gorlov I, Ying J, Chen WV, Assassi S, et al. Gene-level association analysis of systemic sclerosis: a comparison of African-Americans and White populations. PLoS One. 2018;13(1):e0189498. https://doi.org/10.1371/journal.pone.0189498.
59. Bossini-Castillo L, Martin JE, Broen J, Simeon CP, Beretta L, Gorlova OY, et al. Confirmation of TNIP1 but not RHOB and PSORS1C1 as systemic sclerosis risk factors in a large independent replication study. Ann Rheum Dis. 2013;72(4):602–7. https://doi.org/10.1136/annrheumdis-2012-201888.
60. Bossini-Castillo L, Broen JC, Simeon CP, Beretta L, Vonk MC, Ortego-Centeno N, et al. A replication study confirms the association of TNFSF4 (OX40L) polymorphisms with systemic sclerosis in a large European cohort. Ann Rheum Dis. 2011;70(4):638–41. https://doi.org/10.1136/ard.2010.141838.
61. Dieudé P, Boileau C, Guedj M, Avouac J, Ruiz B, Hachulla E, et al. Independent replication establishes the CD247 gene as a genetic systemic sclerosis susceptibility factor. Ann Rheum Dis. 2011;70(9):1695–6. https://doi.org/10.1136/ard.2010.147009.
62. López-Isac E, Campillo-Davo D, Bossini-Castillo L, Guerra SG, Assassi S, Simeón CP, et al. Influence of TYK2 in systemic sclerosis susceptibility: a new locus in the IL-12 pathway. Ann Rheum Dis. 2016;75(8):1521–6. https://doi.org/10.1136/annrheumdis-2015-208154.
63. López-Isac E, Bossini-Castillo L, Guerra SG, Denton C, Fonseca C, Assassi S, et al. Identification of IL12RB1 as a novel systemic sclerosis susceptibility locus. Arthritis Rheumatol. 2014;66(12):3521–3. https://doi.org/10.1002/art.38870.
64. Bossini-Castillo L, Martin JE, Broen J, Gorlova O, Simeón CP, Beretta L, et al. A GWAS follow-up study reveals the association of the IL12RB2 gene with systemic sclerosis in Caucasian populations. Hum Mol Genet. 2012;21(4):926–33. https://doi.org/10.1093/hmg/ddr522.
65. Moreno-Moral A, Bagnati M, Koturan S, Ko JH, Fonseca C, Harmston N, et al. Changes in macrophage transcriptome associate with systemic sclerosis and mediate GSDMA contribution to disease risk. Ann Rheum Dis. 2018;77(4):596–601. https://doi.org/10.1136/annrheumdis-2017-212454.
66. López-Isac E, Bossini-Castillo L, Simeon CP, Egurbide MV, Alegre-Sancho JJ, Callejas JL, et al. A genome-wide association study follow-up suggests a possible role for PPARG in systemic sclerosis susceptibility. Arthritis Res Ther. 2014;16(1):R6. https://doi.org/10.1186/ar4432.
67. Bossini-Castillo L, López-Isac E, Martín J. Immunogenetics of systemic sclerosis: defining heritability, functional variants and shared-autoimmunity pathways. J Autoimmun. 2015;64:53–65. https://doi.org/10.1016/j.jaut.2015.07.005.
68. Rockman MV, Kruglyak L. Genetics of global gene expression. Nat Rev Genet. 2006;7(11):862–72. https://doi.org/10.1038/nrg1964.
69. Kerick M, González-Serna D, Carnero-Montoro E, Teruel M, Acosta-Herrera M, Makowska Z, et al. eQTL analysis in systemic sclerosis identifies new candidate genes associated with multiple aspects of disease pathology. Arthritis Rheumatol. 2021. https://doi.org/10.1002/art.41657.
70. 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(6):694–702. https://doi.org/10.1093/rheumatology/kei244.
71. 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(3):1010–20. https://doi.org/10.1002/art.22382.
72. Higgs BW, Liu Z, White B, Zhu W, White WI, Morehouse C, et al. Patients with systemic lupus erythematosus, myositis, rheumatoid arthritis and scleroderma share activation of a common type I interferon pathway. Ann Rheum Dis. 2011;70(11):2029–36. https://doi.org/10.1136/ard.2011.150326.
73. Graham RR, Kozyrev SV, Baechler EC, Reddy MV, Plenge RM, Bauer JW, et al. A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. Nat Genet. 2006;38(5):550–5. https://doi.org/10.1038/ng1782.
74. Dieguez-Gonzalez R, Calaza M, Perez-Pampin E, de la Serna AR, Fernandez-Gutierrez B, Castañeda S, et al. Association of interferon regulatory factor 5 haplotypes, similar to that found in systemic lupus erythematosus, in a large subgroup of patients with rheumatoid arthritis. Arthritis Rheum. 2008;58(5):1264–74. https://doi.org/10.1002/art.23426.
75. Fu Q, Zhao J, Qian X, Wong JL, Kaufman KM, Yu CY, et al. Association of a functional IRF7 variant with systemic lupus erythematosus. Arthritis Rheum. 2011;63(3):749–54. https://doi.org/10.1002/art.30193.
76. Cunninghame Graham DS, Morris DL, Bhangale TR, Criswell LA, Syvänen AC, Rönnblom L, et al. Association of NCF2, IKZF1, IRF8, IFIH1, and TYK2 with systemic lupus erythematosus. PLoS Genet. 2011;7(10):e1002341. https://doi.org/10.1371/journal.pgen.1002341.
77. Eyre S, Bowes J, Diogo D, Lee A, Barton A, Martin P, et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat Genet. 2012;44(12):1336–40. https://doi.org/10.1038/ng.2462.
78. Kristjansdottir G, Sandling JK, Bonetti A, Roos IM, Milani L, Wang C, et al. Interferon regulatory factor 5 (IRF5) gene variants are associated with multiple sclerosis in three distinct populations. J Med Genet. 2008;45(6):362–9. https://doi.org/10.1136/jmg.2007.055012.
79. Bhattacharyya S, Wang W, Graham LV, Varga J. A20 suppresses canonical Smad-dependent fibroblast activation: novel function for an endogenous inflammatory modulator. Arthritis Res Ther. 2016;18(1):216. https://doi.org/10.1186/s13075-016-1118-7.
80. Wang J, Yi L, Guo X, He D, Li H, Guo G, et al. Lack of association of the CD247 SNP rs2056626 with systemic sclerosis in Han Chinese. Open Rheumatol J. 2014;8(1):43–5. https://doi.org/10.2174/1874312901408010043.
81. Kunz M, Ibrahim SM. Non-major histocompatibility complex rheumatoid arthritis susceptibility genes. Crit Rev Immunol. 2011;31(2):99–114. https://doi.org/10.1615/CritRevImmunol.v31.i2.20.
82. Skhirtladze C, Distler O, Dees C, Akhmetshina A, Busch N, Venalis P, et al. Src kinases in systemic sclerosis: central roles in fibroblast activation and in skin fibrosis. Arthritis Rheum. 2008;58(5):1475–84. https://doi.org/10.1002/art.23436.
83. Barnes J, Agarwal SK. Targeting STAT4 in systemic sclerosis: a promising new direction. Expert Rev Clin Immunol. 2011;7(4):445–8. https://doi.org/10.1586/eci.11.31.

84. Avouac J, Fürnrohr BG, Tomcik M, Palumbo K, Zerr P, Horn A, et al. Inactivation of the transcription factor STAT-4 prevents inflammation-driven fibrosis in animal models of systemic sclerosis. Arthritis Rheum. 2011;63(3):800–9. https://doi.org/10.1002/art.30171.
85. Mak AC, Tang PL, Cleveland C, Smith MH, Kari Connolly M, Katsumoto TR, et al. Whole-exome sequencing for identification of potential causal variants for diffuse cutaneous systemic sclerosis. Arthritis Rheumatol. 2016;68(9):2257–62. https://doi.org/10.1002/art.39721.
86. Coustet B, Dieudé P, Guedj M, Bouaziz M, Avouac J, Ruiz B, et al. 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. 2011;63(7):2091–6. https://doi.org/10.1002/art.30379.
87. Diogo D, Bastarache L, Liao KP, Graham RR, Fulton RS, Greenberg JD, et al. TYK2 protein-coding variants protect against rheumatoid arthritis and autoimmunity, with no evidence of major pleiotropic effects on non-autoimmune complex traits. PLoS ONE. 2015;10(4):e0122271. https://doi.org/10.1371/journal.pone.0122271.
88. Zhou X, Li H, Guo S, Wang J, Shi C, Espitia M, et al. Associations of multiple NOTCH4 exonic variants with systemic sclerosis. J Rheumatol. 2019;46(2):184–9. https://doi.org/10.3899/jrheum.180094.
89. Wei J, Zhu H, Komura K, Lord G, Tomcik M, Wang W, et al. A synthetic PPAR-γ agonist triterpenoid ameliorates experimental fibrosis: PPAR-γ-independent suppression of fibrotic responses. Ann Rheum Dis. 2014;73(2):446–54. https://doi.org/10.1136/annrheumdis-2012-202716.
90. Manetti M, Allanore Y, Saad M, Fatini C, Cohignac V, Guiducci S, et al. Evidence for caveolin-1 as a new susceptibility gene regulating tissue fibrosis in systemic sclerosis. Ann Rheum Dis. 2012;71(6):1034–41. https://doi.org/10.1136/annrheumdis-2011-200986.
91. Oka A, Asano Y, Hasegawa M, Fujimoto M, Ishikawa O, Kuwana M, et al. RXRB is an MHC-encoded susceptibility gene associated with anti-topoisomerase I antibody-positive systemic sclerosis. J Invest Dermatol. 2017;137(9):1878–86. https://doi.org/10.1016/j.jid.2017.04.028.
92. Dantas AT, Pereira MC, de Melo Rego MJ, da Rocha LF, Jr., Pitta Ida R, Marques CD, et al. The role of PPAR gamma in systemic sclerosis. PPAR Res. 2015;2015:124624.
93. Kuwana M, Kaburaki J, Arnett FC, Howard RF, Medsger TA Jr, Wright TM. Influence of ethnic background on clinical and serologic features in patients with systemic sclerosis and anti-DNA topoisomerase I antibody. Arthritis Rheum. 1999;42(3):465–74. https://doi.org/10.1002/1529-0131(199904)42:3<465::AID-ANR11>3.0.CO;2-Y.
94. Arnett FC, Gourh P, Shete S, Ahn CW, Honey RE, Agarwal SK, et al. 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. 2010;69(5):822–7. https://doi.org/10.1136/ard.2009.111906.
95. Gentil CA, Gammill HS, Luu CT, Mayes MD, Furst DE, Nelson JL. Characterization of the HLA-DRβ1 third hypervariable region amino acid sequence according to charge and parental inheritance in systemic sclerosis. Arthritis Res Ther. 2017;19(1):46. https://doi.org/10.1186/s13075-017-1253-9.
96. Reveille JD, Durban E, MacLeod-St Clair MJ, Goldstein R, Moreda R, Altman RD, et al. Association of amino acid sequences in the HLA-DQB1 first domain with antitopoisomerase I autoantibody response in scleroderma (progressive systemic sclerosis). J Clin Invest. 1992;90(3):973–80. https://doi.org/10.1172/JCI115974.
97. Kuwana M, Inoko H, Kameda H, Nojima T, Sato S, Nakamura K, et al. Association of human leukocyte antigen class II genes with autoantibody profiles, but not with disease susceptibility in Japanese patients with systemic sclerosis. Intern Med. 1999;38(4):336–44. https://doi.org/10.2169/internalmedicine.38.336.
98. Kuwana M, Kaburaki J, Okano Y, Inoko H, Tsuji K. The HLA-DR and DQ genes control the autoimmune response to DNA topoisomerase I in systemic sclerosis (scleroderma). J Clin Invest. 1993;92(3):1296–301. https://doi.org/10.1172/JCI116703.
99. Arnett FC, Howard RF, Tan F, Moulds JM, Bias WB, Durban E, et al. Increased prevalence of systemic sclerosis in a Native American tribe in Oklahoma. Association with an Amerindian HLA haplotype. Arthritis Rheum. 1996;39(8):1362–70. https://doi.org/10.1002/art.1780390814.
100. Kuwana M, Medsger TA Jr, Wright TM. Highly restricted TCR-alpha beta usage by autoreactive human T cell clones specific for DNA topoisomerase I: recognition of an immunodominant epitope. J Immunol. 1997;158(1):485–91.
101. Gourh P, Safran SA, Alexander T, Boyden SE, Morgan ND, Shah AA, et al. HLA and autoantibodies define scleroderma subtypes and risk in African and European Americans and suggest a role for molecular mimicry. Proc Natl Acad Sci U S A. 2020;117(1):552–62. https://doi.org/10.1073/pnas.1906593116.
102. Kuwana M, Pandey JP, Silver RM, Kawakami Y, Kaburaki J. HLA class II alleles in systemic sclerosis patients with anti-RNA polymerase I/III antibody: associations with subunit reactivities. J Rheumatol. 2003;30(11):2392–7.
103. Nikpour M, Hissaria P, Byron J, Sahhar J, Micallef M, Paspaliaris W, et al. Prevalence, correlates and clinical usefulness of antibodies to RNA polymerase III in systemic sclerosis: a cross-sectional analysis of data from an Australian cohort. Arthritis Res Ther. 2011;13(6):R211. https://doi.org/10.1186/ar3544.
104. Kuwana M, Okano Y, Kaburaki J, Tojo T, Medsger TA Jr. Racial differences in the distribution of systemic sclerosis-related serum antinuclear antibodies. Arthritis Rheum. 1994;37(6):902–6. https://doi.org/10.1002/art.1780370619.
105. Milano A, Pendergrass SA, Sargent JL, George LK, McCalmont TH, Connolly MK, et al. Molecular subsets in the gene expression signatures of scleroderma skin. PLoS One. 2008;3(7):e2696. https://doi.org/10.1371/journal.pone.0002696.
106. Hinchcliff M, Huang CC, Wood TA, Matthew Mahoney J, Martyanov V, Bhattacharyya S, et al. Molecular signatures in skin associated with clinical improvement during mycophenolate treatment in systemic sclerosis. J Invest Dermatol. 2013;133(8):1979–89. https://doi.org/10.1038/jid.2013.130.
107. Chakravarty EF, Martyanov V, Fiorentino D, Wood TA, Haddon DJ, Jarrell JA, et al. Gene expression changes reflect clinical response in a placebo-controlled randomized trial of abatacept in patients with diffuse cutaneous systemic sclerosis. Arthritis Res Ther. 2015;17(1):159. https://doi.org/10.1186/s13075-015-0669-3.
108. Tsou PS, Sawalha AH. Unfolding the pathogenesis of scleroderma through genomics and epigenomics. J Autoimmun. 2017;83(9):73–94. https://doi.org/10.1016/j.jaut.2017.05.004.

Publisher's Note