The role of genetics and epigenetics in the pathogenesis of systemic sclerosis

Jasper C. A. Broen, Timothy R. D. J. Radstake and Marzia Rossato

Abstract | Systemic sclerosis (SSc) is a complex autoimmune disease of unclear aetiology. A multitude of genetic studies, ranging from candidate-gene studies to genome-wide association studies, have identified a large number of genetic susceptibility factors for SSc and its clinical phenotypes, but the contribution of these factors to disease susceptibility is only modest. However, in an endeavour to explore how the environment might affect genetic susceptibility, epigenetic research into SSc is rapidly expanding. Orchestrated by environmental factors, epigenetic modifications can drive genetically predisposed individuals to develop autoimmunity, and are thought to represent the crossroads between the environment and genetics in SSc. Therefore, in addition to providing a comprehensive description of the current understanding of genetic susceptibility underlying SSc, this Review describes the involvement of epigenetic phenomena, including DNA methylation patterns, histone modifications and microRNAs, in SSc.

Broen, J. C. A. *et al. Nat. Rev. Rheumatol.* advance online publication 19 August 2014; doi:10.1038/nrrheum.2014.128

Introduction

The aetiology of systemic sclerosis (SSc) is poorly understood. The main manifestations of this severe autoimmune disease are vasculopathy, immune activation and extensive fibrosis of the skin and internal organs.[1] Two subsets of patients with SSc exist on the basis of the extent of skin fibrosis: limited cutaneous SSc (lcSSc) and diffuse cutaneous SSc (dcSSc). This classification also partially reflects the severity of the disease and the association with specific autoantibodies: lcSSc is more commonly associated with isolated pulmonary hypertension, late smouldering progression of the disease, and the presence of anti-centromere autoantibodies (ACA); dcSSc is more prominently linked to the development of interstitial lung disease and musculoskeletal manifestations with early, dramatic onset, and the presence of anti-topoisomerase autoantibodies (ATA).[2] In addition, the presence of anti-RNA polymerase III autoantibodies has been linked to susceptibility to renal crisis in patients with SSc.[3]

Intense research into SSc over the past 10 years has resulted in many genetic and environmental factors being implicated in disease susceptibility, and SSc has also been linked to the development of cancer.[4–6] In less than a decade, genetic analysis techniques have advanced from small cohorts of candidate-gene studies to complex genome-wide analyses in large multinational cohorts, expediting the discovery of various genetic factors that contribute to SSc susceptibility. Findings from these studies have given us novel insights into the pathways and cell types that are likely to be involved in SSc pathogenesis, but these data still face further investigation through functional research and for clinical relevance.

Competing interests
The authors declare no competing interests.

Department of Rheumatology & Clinical Immunology, Laboratory of Translational Immunology, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, Netherlands (**J.C.A.B.**, **T.R.D.J.R.**, **M.R.**).

Correspondence to: T.R.D.J.R. t.r.d.j.radstake@umcutrecht.nl

We aim, in this Review, to provide an extensive overview of these genetic risk factors, and to be critical of what this knowledge has revealed in terms of disease pathogenesis. Given that the pathogenesis of SSc cannot be attributed solely to genetic causes and the fact that, of all autoimmune diseases, this disease has the lowest concordance rate among monozygotic twins, research over the past 5 years has shifted focus to epigenetics. Multiple environmental factors have been suggested to be involved in SSc susceptibility—occupational exposure to silica (silicon dioxide) is most often reported[6]—inducing epigenetic modifications that can promote the development of autoimmunity in genetically predisposed individuals. A discussion of such epigenetic modifications is the focus of the latter part of this article.

The role of genetics in SSc

Family and twin studies

As with many rheumatic and autoimmune diseases, SSc is not inherited in a Mendelian fashion; however, the disease can occur in families in which a member has already been diagnosed with SSc or another autoimmune disease. Having a sibling with SSc is one of the highest risk factors for developing SSc (a 15–19-fold increase over the general population) and other first-degree relatives have a 13–15-fold increased risk of developing SSc compared with the general population. Although this finding seems to be a sharp increase in risk rates, the absolute risk of developing SSc is only 1.6% in these relatives, compared with 0.026% in the general population.[7–10] These findings were strengthened by a study of the Utah Population Database, a genealogical resource that contains data on the family structure and medical records of more than 7 million inhabitants of

Key points

- Systemic sclerosis (SSc) is a complex autoimmune disease, the pathogenesis of which is influenced by genetic, epigenetic and environmental factors
- Candidate-gene studies and genome-wide association studies have identified a large number of genetic susceptibility factors for SSc and its clinical phenotypes
- The low concordance rate for monozygotic twins demonstrates that a genetic basis cannot account exclusively for SSc pathogenesis and that epigenetic factors influenced by the environment are important
- Several DNA methylation patterns, histone modifications and microRNAs (miRNAs) are altered in different cell types from patients with SSc; altered circulating miRNAs are potentially useful as disease diagnostic and prognostic biomarkers

Box 1 | SNPs and GWAS

- SNPs are changes in single base pairs that influence gene function differently depending on where they occur in a gene. Whereas their presence in an exon could cause an amino-acid substitution or premature termination, SNPs that occur within a splice region could promote alternative splicing and those occurring within a promoter or in the 3' untranslated region could influence gene expression. Furthermore, on the basis of linkage disequilibrium, SNPs can also be used to indicate the presence of certain genomic regions in close proximity to their own locus. Consequently, the SNP itself might not be responsible for conferring disease susceptibility but might be physically linked to a genomic region that contains a real causative variant. So, whereas some candidate-gene studies use variants that confer altered protein function, others will be based on linkage disequilibrium.
- GWAS take advantage of the fact that SNPs provide information about their surrounding genomic region. Screening large numbers of SNPs enables differences within the whole genome between cases and controls to be identified. The precision of a genome-wide scan depends heavily on the number of markers used, and this approach is not always able to detect rare genomic variants. The next step in genomic analyses will, therefore, be full-genome sequencing, which covers all nucleotides in the genome.

Abbreviations: GWAS, genome-wide association studies; SNP, single-nucleotide polymorphisms.

Utah, USA.[11] Analysis of 1,037 unique patients with SSc from this database revealed that 50 families affected by SSc had significant elevated familial standardized incidence ratios (2.07–17.60) and that an increased relative risk existed for other autoimmune disease in the first and second-degree relatives.

A study in 2003 of 42 twin pairs (24 monozygotic and 18 dizygotic), of which at least one sibling had SSc, revealed that only two pairs—both female–female, one monozygotic and one dizygotic pair—were concordant for SSc. No statistically significant difference in the concordance rates between monozygotic twins (4.2%) and dizygotic twins (5.6%) was observed.[12] However, a number of the twins in the study were below the peak-of-onset age for SSc; therefore, some of the healthy twins might be prone to develop SSc in later life.

Candidate-gene studies

Multiple candidate-gene studies of SSc have been carried out. Most often, the candidate genes and single-nucleotide polymorphisms (SNPs; Box 1) were selected on the basis of their involvement in the susceptibility to other autoimmune diseases; a comprehensive list of associated non-HLA SNPs is provided in Supplementary Table 1 online. These SNPs, and multiple HLA associations, are discussed extensively elsewhere.[4,13] Here, we will try to consider the findings within a larger framework by showing that the candidate genes reside in overlapping pathways that function separate to antigen presentation, as suggested by HLA associations.

T-cell signalling

The most frequently described SNP involved in SSc susceptibility is located in the gene encoding signal transducer and activator of transcription 4 (STAT4).[14–18] STAT4 is required for the development of type 1 T helper (T_H1) cells from naive $CD4^+$ cells and mediates the production of IFN-γ, processes that are strongly implicated in SSc pathogenesis.[19] Consistent with a role for STAT4 in SSc, in a mouse model of bleomycin-induced skin fibrosis, deletion of *Stat4* resulted in impaired T-cell activation and reduced fibrosis.[18]

SNPs in *TNFSF4*, which encodes TNF ligand superfamily member 4 (also known as OX40 ligand [OX40L]) have also been found to exacerbate the risk of developing SSc. These polymorphisms highlight the involvement of T-cell signalling in SSc susceptibility, as TNFSF4 is the natural ligand of TNF receptor superfamily member 4, which is expressed on the surface of T cells and influences T-cell survival and proliferation.[20] However, in contrast to the aforementioned *STAT4* susceptibility regions, *TNFSF4* variants were found to be associated with disease phenotype rather than overall SSc susceptibility.[21,22]

As mice lacking *Tbx21*, which encodes the T_H1-cell transcription factor T-bet, develop extensive fibrosis after injection with bleomycin, researchers investigated a potential role for this gene in SSc as well.[22–24] A study of white American patients and healthy individuals controls found that a variant of *TBX21* was involved in SSc susceptibility, regardless of the clinical phenotype. Interestingly, homozygous carriers of the rs11650354 TT variant had higher expression of T_H2 cytokines than people with the CC genotype, and the CC genotype was associated with a type I interferon signature.[16]

On the basis that T cells can dampen the inflammasome by inhibiting nucleotide-binding oligomerization domain, leucine-rich repeat and pyrin domain-containing protein 1 (NLRP1) responses, and that *NLRP1* variants conferred susceptibility to autoimmune disorders,[25] a potential role for this gene in SSc was investigated. In a study comprising an identification cohort of 870 patients with SSc and 962 healthy controls and a replication set consisting of 1,060 patients with SSc and 625 healthy controls, a variant of *NLRP1* was found to confer susceptibility to ATA-positive SSc phenotypes and fibrosing alveolitis.[25,26]

B-cell signalling

Associations between defects in B-cell signalling components and SSc pathogenesis have also been described, highlighting a role for the immune system, rather than intrinsic fibroblast defects, in disease susceptibility. Two polymorphisms—rs13277113 and rs2736340—within the FAM167A-BLK region are known to be associated with lcSSc and dcSSc. The FAM167A-BLK region has been investigated in multiple autoimmune diseases;[27]

BLK is a kinase that is expressed in thymocytes and is involved in signalling downstream of the B-cell receptor.[29,30] Both variants also seem to be more prevalent in ACA-positive lcSSc.[22] The association between rs13277113 and SSc was replicated in a study of Japanese patients.[28] When stratifying patients based on their genotype, RNA expression profiling revealed a difference in activation of BCR-related pathways between genotypes.[31]

A second B-cell-specific candidate gene encodes B-cell scaffold protein with ankyrin repeats (BANK1), an adaptor protein involved in signalling downstream of the B-cell antigen receptor. Research into *BANK1* polymorphisms in SSc has mostly been based on findings from studies of patients with systemic lupus erythematosus (SLE) or rheumatoid arthritis (RA).[32,33] A large multicentre study of 2,380 white patients with SSc and 3,270 ethnically matched healthy controls revealed a significant association of rs10516487 and rs17266594 with dcSSc. An independent replication study confirmed these results.[34,35]

Also of particular interest are two studies that implicate the *IRAK1* gene in the development of pulmonary fibrosis in SSc.[36,37] As *IRAK1* is located on the X chromosome, these findings might have a role in the high prevalence of SSc in women. The exact contribution, however, needs to be further investigated by X-chromosomal gene expression and *in vitro* experiments.

Taken together, the candidate-gene studies show that polymorphisms tend to cluster in pathways involved in B-cell and T-cell receptor signalling and development. These findings lead to the hypothesis that SSc is a disease originating from the defective interaction of B cells and T cells.

Genome-wide association studies

Genome-wide association studies (GWAS) have expedited the discovery of multiple novel polymorphisms that increase susceptibility to SSc or to one of its clinical phenotypes. Definitions for SNPs and GWAS are given in Box 1.

GWAS and SSc susceptibility

The first SSc GWAS, carried out in 137 Korean patients with SSc and 564 healthy controls, analysed 500,568 SNP markers and discovered associations of SNPs within the HLA-DPB1 region with SSc susceptibility, ATA positivity and ACA positivity.[38] In 2010, the largest GWAS to date—involving 2,296 patients with SSc and 5,171 healthy controls—not only confirmed a strong association with HLA and replicated the association of SSc with the genes encoding interferon regulatory factor (IRF) 5 and STAT4 (Supplementary Table 1 online), but also identified five non-HLA loci, of which *CD247* reached genome-wide significance after replication;[39] the *CD247* association has been robustly replicated in a white French population.[40] *CD247* encodes the T-cell receptor ζ subunit, a component of the T-cell receptor complex, and the results, therefore, underscore the relevance of the antigen-presenting cell–T-cell interaction in the pathogenesis of SSc.

GWAS and SSc phenotype

A GWAS was conducted in a French cohort consisting of 564 SSc cases and 1,776 healthy controls followed by replication in 1,682 SSc cases and 3,926 healthy controls. Psoriasis susceptibility 1 candidate gene 1 protein (*PSORS1C1*), TNF alpha-induced protein 3 (TNFAIP3)-interacting protein 1 (*TNIP1*) and ras homolog gene family member B (*RHOB*) loci were identified as novel gene regions associated with SSc susceptibility.[41] Furthermore, the involvement of *STAT4*, *IRF5* and *CD247* was confirmed.[41] A large follow-up study corroborated the association of *TNIP1*, but did not find any association with *PSORS1C1* or *RHOB*.[42] Data from a follow-up of the French GWAS suggested that *PPARG*, which encodes peroxisome proliferator-activated receptor-γ, was associated with SSc, although genome-wide significance was not reached.[43] Another follow-up study found that *IL12RB2*, which encodes the β2 subunit of the IL-12 receptor that is functionally linked to STAT4, CD247 and IFN-γ, was associated with SSc.[44]

A genome-wide meta-analysis of SLE and SSc in a cohort with a sample size of 21,109 (6,835 cases and 14,274 controls) identified *KIAA0319L* as a novel susceptibility locus for SSc and SLE, and found the previously described SLE susceptibility loci *PXK* and *JAZF1* to be shared with SSc.[45] Intriguingly, the researchers found that *KIAA0319L* was overexpressed by the peripheral blood cells of patients with SSc or SLE compared with healthy controls. However, the role of *KIAA0319L* has not yet been fully elucidated.

Finally, an Immunochip study with high-resolution coverage of immune-associated loci of 1,833 systemic sclerosis (SSc) cases and 3,466 healthy controls identified *DNASE1L3* (involved in DNA degradation in apoptosis), *SCHIP1-IL12A* (encodes a component of the IFN-γ pathway), *ATG5* (involved in IFN-α signalling and autophagy), and *TREH-DDX6* (has a role in RNA degradation and microRNA [miRNA]-induced gene silencing) as susceptibility loci for SSc; the data were later validated using a replication cohort comprising 4,017 SSc cases and 5,935 controls.[46]

Several follow-up GWAS have focused on identifying the genetic components that contribute to clinical phenotype rather than SSc susceptibility. So far, these efforts have led to the discovery that *IRF8* is specifically associated with lcSSc. The same study also demonstrated that HLA-DQB1 is associated with ACA positivity, whereas HLA-DPA1/B1 was associated with ATA positivity. Less discriminating was *NOTCH4*, which was associated with the presence of both ACA and ATA.[47] Altogether, GWAS show that T-cell signalling and interferon signalling pathways are involved in SSc susceptibility, and reveal roles in apoptosis, DNA or RNA degradation and autophagy.[45,46]

Exploring the field of epigenetics

The aforementioned studies show that genetic factors are unable to fully account for the risk of SSc development—particularly when considering the extremely low concordance rates of SSc in monozygotic twins (4.2%).[12]

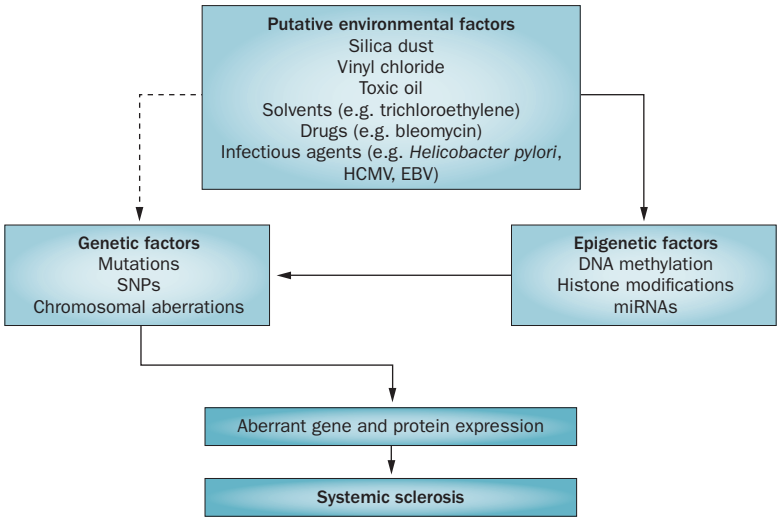

Figure 1 | The pathogenesis of SSc is influenced by genetic, epigenetic and environmental factors.[48,101] Several environmental factors have been proposed to contribute to the development of SSc; however, only exposure to silica has been proved to be associated with an increased SSc risk.[50] Some environmental agents can induce genetic mutations or chromosomal rearrangements,[4] but in most cases they probably induce epigenetic modifications to promote autoimmunity. By affecting the expression of individual genes on a genetic background that is already predisposed by the presence of SSc-susceptibility SNPs, aberrant epigenetic markers can promote the development of SSc. Abbreviations: EBV, Epstein–Barr virus; HCMV, human cytomegalovirus; SNP, single-nucleotide polymorphism; SSc, systemic sclerosis.

Box 2 | Epigenetic mechanisms

- **Epigenetics**: heritable changes that influence gene expression without altering the DNA sequence; these changes include DNA methylation, post-translational modification of histones and miRNAs.
- **DNA methylation**: describes the addition of a methyl group to CpG dinucleotides, mainly concentrated in CpG islands located within promoter regions. This addition promotes a more condensed DNA configuration, blocking accessibility to transcriptional activators and thereby inhibiting gene transcription; transcriptional activators can no longer bind because of tight chromatic configuration. Patterns of DNA methylation are established and maintained by DNA methyltransferases, whereas methyl-CpG-binding domain proteins associate with methyl-cytosine and recruit silencing complexes and histone deacetylases to consolidate the heterochromatic state (as in, maintain the tighter, closed DNA configuration of heterochromatin).
- **Histone modifications**: includes acetylation, phosphorylation and methylation of specific histone residues that influence the accessibility of chromatin to transcription factors at gene promoters and enhancers. Histone acetylation, which promotes a more open chromatin structure, is catalysed by histone acetyltransferase enzymes. Histone deacetylases remove acetyl groups causing DNA to wrap more tightly around the nucleosome, thereby repressing gene expression. Histones can be also methylated by histone methyltransferases, or demethylated by histone demethylases. Histone H3 trimethylation at lysine 4 (H3K4me3) is associated with active chromatin and gene expression. By contrast, methylation at histone H3 on lysine 9 (H3K9me) or histone H3 trimethylation at lysine 27 (H3K27me3) represses transcription by inducing a condensed heterochromatin status.
- **miRNAs**: short, noncoding RNAs of 18–23 nucleotides that function as endogenous inhibitors of gene expression by binding the 3' untranslated region of target mRNAs, and thereby inhibiting their translation or promoting mRNA destabilization and degradation. miRNAs are expressed in a tissue-specific and cell type-specific manner, but can also stably circulate in biological fluids by associating with carrier proteins or being incorporated in microvesicles such as exosomes and apoptotic bodies.

Abbreviations: CpG, cytosine–guanine linear dinucleotide; miRNA, microRNA.

These findings suggest that the environment can affect the breakage of immune tolerance and the development of fibrosis and SSc in certain genetic backgrounds. This concept is supported by numerous studies indicating that different environmental factors, including occupational exposure to silica dust, vinyl chloride or drugs such as bleomycin, can induce SSc-like symptoms[48] (Figure 1). Interestingly, no clear association between smoking and the risk of developing SSc has been established.[49]

The individual role of environmental factors in SSc risk is difficult to evaluate, owing to the rarity and heterogeneity of the disease.[50] The overall effect of single agents is, however, estimated to be modest, considering the size of the 'exposed' population compared with the low prevalence of SSc. Accordingly, it is thought that multiple factors, each of them having a mild effect, can contribute to the development of SSc in individuals with a genetic predisposition (Figure 1). Potential mechanisms for environmentally induced systemic autoimmunity include interference with immune tolerance, activation of the immune system, induction of genetic alterations and dysregulation of epigenetic patterns.[4,50] The observation that fibroblasts taken from patients with SSc keep their altered behaviour for multiple passages when cultured outside the pathological context is one of the strongest indicators for the fundamental role of epigenetic regulation in SSc to date. Consistent with this theory, epigenetic modifications such as DNA methylation and histone modification (Box 2) have been shown to be important in determining gene activity in other autoimmune conditions, including SLE and RA.[51,52] Consequently, the study of epigenetic phenomena in SSc has flourished over the past few years, providing new insights into disease pathogenesis and novel potential therapeutic targets.

Herein, we provide a comprehensive overview of epigenetic markers and modifiers that have been shown to be altered in specific cell types or tissues taken directly from patients with SSc (Tables 1 and 2; Figure 2). We also outline how common epigenetic factors might be appealing targets for future therapeutic intervention for SSc.

DNA methylation

Altered DNA methylation in lymphocytes

Several studies have shown that SLE and RA are generally associated with global DNA hypomethylation that causes the overexpression of autoimmune-related genes.[51] Similarly, $CD4^+$ T cells from patients with SSc show a general reduction in the levels of DNA methylation associated with a concurrent decreased expression of methylation-regulating genes, such as *DNMT1*, *MBD3* and *MBD4*.[53] The promoter of *CD70* (also known as *TNFSF7* or *CD27L*), which encodes the B-cell co-stimulatory molecule CD70, is substantially hypomethylated and probably contributes to the overexpression of CD70 by $CD4^+$ T cells of patients with SSc.[54] In addition, $CD4^+$ T cells from patients with SSc have high expression of *CD40LG* (which encodes CD40L; also known as CD154), which is located on the X chromosome

Table 1 | Epigenetic alterations in SSc at the level of DNA methylation and histone marks

Epigenetic alteration	Finding in SSc	Tissue	Gene affected	Gene product	Impact on gene expression	References
DNA methylation						
Pattern	Hypomethylation and hypermethylation	PBMCs from monozygotic twins	X-chromosome genes	X-chromosome-derived proteins	Upregulation and downregulation	Selmi *et al.* (2012)[57]
	Hypomethylation	Female $CD4^+$ T cells	*CD40LG*	CD40L	Upregulation	Lian *et al.* (2012)[56]
	Hypomethylation	$CD4^+$ T cells	*CD70*	CD70	Upregulation	Jiang *et al.* (2012)[54]
	Hypermethylation	Fibroblasts, PBMCs, bleomycin-treated skin	*DKK1* and *SFRP1*	DKK1 and SFRP1	Downregulation	Dees *et al.* (2013)[60]
	Hypermethylation	Fibroblasts	*FLI1*	FLI1	Downregulation	Wang *et al.* (2006)[59]
	Hypermethylation	MVECs and skin	*BMPR2*	BMPR2	Downregulation	Wang *et al.* (2013)[61]
Enzyme affected	Decreased DNMT1, MBD3 and MBD4	$CD4^+$ T cells	Global hypomethylation	ND	Upregulation	Komura *et al.* (2008)[55]
	Increased DNMT1, MBD1 and MeCP2	Fibroblasts	Global hypermethylation	ND	Downregulation	Wang *et al.* (2006)[59]
Inhibitor*	5-aza-2'-deoxycytidine (DNMT)	Fibroblasts	Global hypomethylation, *FLI1*, *DKK1*, *SFRP1*	FLI1, DKK1, SFRP1	Upregulation	Wang *et al.* (2006)[59] Dees *et al.* (2013)[60]
Histone modification						
Pattern	H3 and H4 deacetylation	Fibroblasts	*FLI1*	FLI1	Downregulation	Wang *et al.* (2006)[59]
	H4 hyperacetylation and H3K9 hypomethylation	B cells	ND	ND	Upregulation	Wang *et al.* (2013)[69]
	Increased H3K27me3	Fibroblasts	ND	ND	Downregulation	Kramer *et al.* (2013)[68]
Enzyme affected	Increased p300 (histone acetyltransferase)	Fibroblasts and skin	Increased *COL1A2* acetylation	Collagen-α2(I)	Upregulation	Ghosh *et al.* (2013)[64]
	Increased HDAC1 and HDAC6	Fibroblasts	Global deacetylation	ND	Downregulation	Wang *et al.* (2006)[59]
	Decreased HDAC2 and HDAC7	B cells	Global histone H4 hyperacetylation	ND	Upregulation	Wang *et al.* (2013)[69]
	Increased KDM3a (histone demethylase), decreased SUV39H2 (histone methyltransferase)	B cells	Global histone H3K9 hypomethylation	ND	Upregulation	Wang *et al.* (2013)[69]
Inhibitor*	Trichostatin A (HDAC)	Fibroblasts and bleomycin-treated skin	Global hyperacetylation	ND	Upregulation	Wang *et al.* (2006)[59] Huber *et al.* (2007)[62] Hemmatazad *et al.* (2009)[63]
	3-deazaneplanocin A (histone methylation)	Fibroblasts and bleomycin-treated skin	Global inhibition of H3K27me3, *FOSL2*	FOSL2	Upregulation	Kramer *et al.* (2013)[68]

*The effect of chemicals modulating DNA methylation and histone modifications on fibroblasts in the context of SSc are also summarized; the target is indicated in parentheses. Abbreviations: BMPR2, bone morphogenetic protein receptor 2; DKK1, Dickkopf-related protein 1; DNMT, DNA methyltransferase; FLI1, Friend leukemia integration 1 transcription factor; FOSL2, FOS-like antigen 2; HDAC, histone deacetylase; KDM3A, lysine-specific demethylase 3A (also known as JmjC domain-containing histone demethylation protein 2A); MBD, methyl-CpG-binding domain protein; me, methylation; MeCP2, methyl-CpG-binding protein 2; MVECs, microvascular endothelial cells; ND, not determined; PBMCs, peripheral blood mononuclear cells; SFRP1, secreted frizzled-related protein 1; SSc, systemic sclerosis; SUV39H2, suppressor of variegation3-9 homologue 2.

and is thought to be important in SSc pathogenesis given that its blockade attenuates skin fibrosis and autoimmunity in the tight-skin mouse model.[55] Interestingly, *CD40LG* upregulation specifically occurred in female patients with SSc and correlated with hypomethylation of its promoter, suggesting that the altered methylation pattern of *CD40LG*—and possibly of other genes on the X chromosome—might contribute to the striking susceptibility of women to SSc.[56] Supporting this hypothesis, the DNA methylation profile of peripheral blood mononuclear cells from monozygotic twins discordant for SSc showed that only X chromosome sites were consistently either hypermethylated or hypomethylated.[57] However, these observations fail to demonstrate the functional consequence of the altered DNA methylation pattern on the phenotype of peripheral leukocytes from patients with SSc. Another report indicated that increased DNA methylation of regulatory sequences in *FOXP3* from $CD4^+$ T cells from patients with SSc affected the expression of this key transcription factor, which is required for the generation of regulatory T cells. This study was the first demonstration of a direct link between changes in the DNA methylation pattern and the number of $CD4^+$ regulatory T cells in SSc.[58] Interestingly, *FOXP3* hypermethylation has previously been suggested to be influenced by X chromosomal inactivation patterns in SSc.[4] These findings might suggest an epigenetically mediated loss of immune homeostasis in SSc development.

Table 2 | miRNAs dysregulated in SSc

miRNA	Expression in SSc	Tissue	Validated target genes	Gene product	Stimuli*	Role in fibrosis	References
miR-7	Upregulated	Fibroblasts, skin	*COL1A1*, *COL2A1*	Collagen α-1(I), collagen α-1(II)	TSP2	Antifibrotic	Kajihara *et al.* (2012)[87]
miR-21	Upregulated	Fibroblasts, skin, bleomycin-treated skin	*SMAD7*	SMAD7	TGF-β	Profibrotic	Zhu *et al.* (2012)[71] Zhu *et al.* (2013)[74]
miR-29a	Downregulated	Fibroblasts, skin, bleomycin-treated skin	*COL1A1*, *COL3A1*	Collagen α-1(I), collagen α-1(III)	TGF-β, PDGF-β, IL-4	Antifibrotic	Zhu *et al.* (2012)[71] Maurer *et al.* (2010)[78] Bhattacharyya *et al.* (2013)[79]
miR-30b	Downregulated	Skin, bleomycin-treated skin, serum	*PDGFRB*	PDGFR-β	TGF-β	Antifibrotic	Tanaka *et al.* (2013)[84]
miR-92a	Upregulated	Fibroblasts, serum	*MMP1*	MMP1	TGF-β	Profibrotic	Sing *et al.* (2012)[86]
miR-129-5p	Downregulated	Fibroblasts	*COL1A1*	Collagen α-1(I)	IL-17A	Antifibrotic	Nakashima *et al.* (2012)[83]
miR-142-3p	Upregulated	Serum	ND	ND	ND	ND	Makino *et al.* (2012)[90]
miR-145, miR-125b, miR-206	Downregulated	Skin	ND	ND	ND	ND	Li *et al.* (2012)[70] Zhu *et al.* (2012)[71]
miR-150	Downregulated	Fibroblasts, skin, serum	*ITGB3*	Integrin β3	ND	Antifibrotic	Honda *et al.* (2013)[73]
miR-196a	Downregulated	Fibroblasts, skin, hair shaft, skin of tight-skin mice	*COL1A1*, *COL2A1*	Collagen α-1(I), collagen α-1(II)	TGF-β	Antifibrotic	Honda *et al.* (2012)[72] Makino *et al.* (2013)[81] Wang *et al.* (2013)[82]
let-7a	Downregulated	Skin, serum	*COL1A1*, *COL2A1*	Collagen α-1(I), collagen α-1(II)	TGF-β	Antifibrotic	Makino *et al.* (2013)[80]
let-7g	Upregulated	Skin	ND	ND	ND	ND	Li *et al.* (2012)[70]

Not all evidence is reported at both the gene and protein level. *Stimuli demonstrated to be responsible for miRNA dysregulation in SSc or able to modulate miRNA expression in a similar manner to that occurring in SSc and therefore possibly responsible for their dysregulation. Abbreviations: miRNA, microRNA; MMP1, matrix metalloproteinase-1; ND, not determined; PDGF, platelet-derived growth factor; PDGFR-β, platelet-derived growth factor receptor β; SMAD7, mothers against decapentaplegic homolog 7; TGF-β, transforming growth factor β; TSP2, thrombospondin 2.

Altered DNA methylation in other cell types

In contrast to $CD4^+$ T cells, altered DNA methylation has been more robustly demonstrated in fibroblasts from patients with SSc. Skin biopsy samples from patients with SSc and derivative fibroblasts showed considerably higher levels of the methylation-regulating genes *DNMT1*, *MBD1* and *MECP2* than samples from healthy controls, indicative of global DNA hypermethylation in fibroblasts from patients with SSc. Consistent with this finding, increased DNA methylation was observed in the promoter region of *FLI1*, which encodes an inhibitor of collagen expression, and treatment with 5-aza-2'-deoxycytidine (an inhibitor of DNA methyltransferases; DNMTs) decreased the level of collagen in fibroblasts from patients with SSc, indicating that epigenetic mechanisms might mediate the fibrotic manifestations of SSc.[59] Hypermethylation in these fibroblasts also affects Wnt signalling, one of the central profibrotic pathways in SSc, by silencing the expression of the Wnt antagonists *DKK1* and *SFRP1*. Treatment with 5-aza-2'-deoxycytidine restored the expression of both genes, thereby blocking activation of the Wnt pathway and reducing bleomycin-induced fibrosis.[60]

Microvascular endothelial cells from patients with SSc were reported to show extensive CpG hypermethylation of the promoter of the gene encoding bone morphogenetic protein receptor type-2 (*BMPR2*), which is important in the pathogenesis of familial pulmonary arterial hypertension.[61] This abnormal methylation downregulated the expression of *BMPR2* and rendered microvascular endothelial cells more vulnerable to apoptosis induced by serum starvation and oxidation injury, potentially contributing to the endothelial cell apoptosis that characterizes SSc lesional skin.

In summary, these studies demonstrate that the modulation of DNA methylation, potentially by using DNMT inhibitors that are already approved for clinical use, might modulate the activity of cells involved in the pathogenesis in SSc, possibly ameliorating fibrosis in this condition.

Histone modifications

Histone acetylation in fibroblasts

The crucial role of histone acetylation in SSc pathogenesis became evident from studies analysing the effect of the histone deacetylase (HDAC) inhibitor trichostatin A (TSA) on gene expression in fibroblasts from patients with SSc. TSA promotes a state of histone hyperacetylation, which is associated with an increased rate of gene transcription that attenuates the expression of *COL1A1* and *FN1*, which encode collagen I and fibronectin, respectively, in both normal and SSc skin fibroblasts.[59,62] This process occurs under resting or stimulated conditions, in response to transforming growth factor (TGF)-β, IL-4 and platelet-derived growth factor (PDGF).[62] TSA also prevented the dermal accumulation of extracellular matrix (ECM) in a mouse model of bleomycin-induced skin fibrosis.[62] At least three

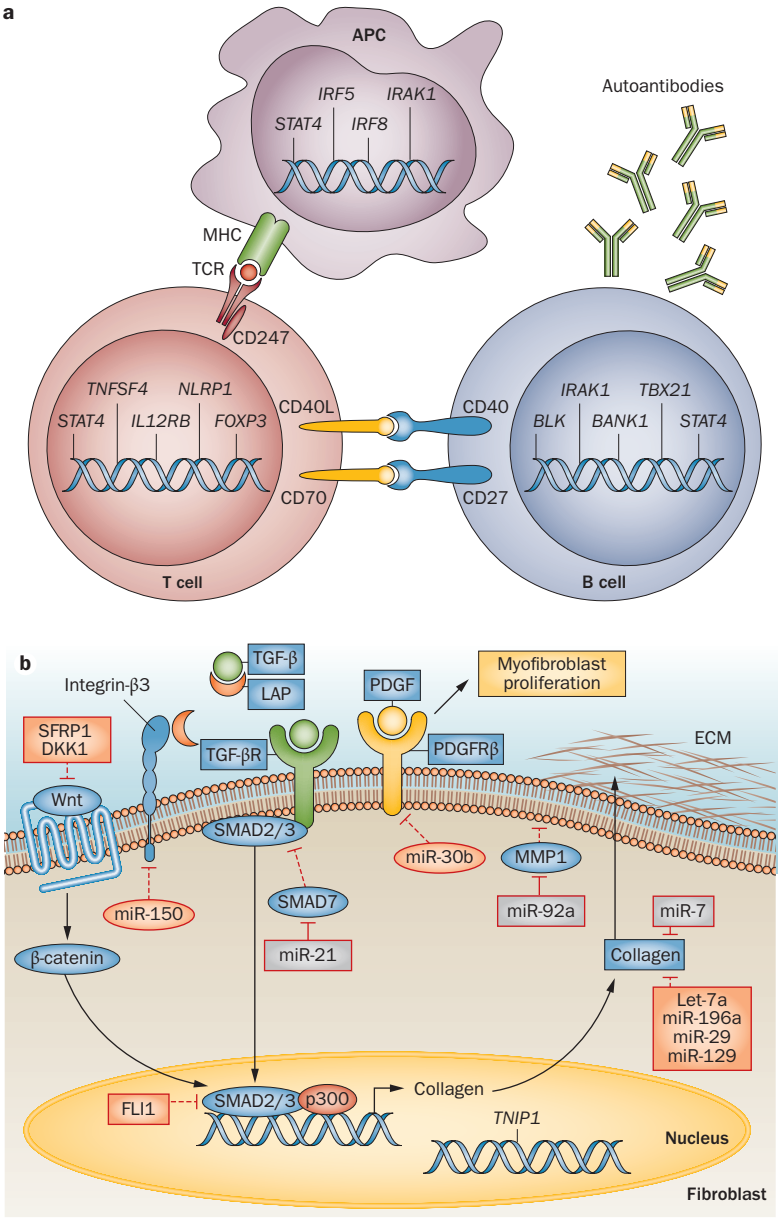

Figure 2 | SSc pathophysiology involves the dysregulation of multiple cell types. This cell dysregulation results in aberrant immune activation, vascular damage, fibroblast activation, excessive collagen and ECM deposition and skin fibrosis. Genetic and epigenetic factors known to affect the activity of **a** | APCs, T cells, B cells and **b** | fibroblasts are indicated in the figure: genes known to influence SSc susceptibility are indicated next to the DNA; miRNAs and molecules upregulated or downregulated in SSc at the epigenetic level are shown respectively in grey boxes with a red border and orange boxes with a red border. The solid lines indicate that the inhibition is enforced because of miRNA upregulation; whereas dashed lines indicate that the inhibition is lost because of miRNA downregulation. Abbreviations: APC, antigen-presenting cell; BANK1, B-cell scaffold protein with ankyrin repeats; BLK, B-lymphocyte kinase; CD40L, CD40 ligand; DKK1, Dickkopf-related protein 1; ECM, extracellular matrix; FLI1, Friend leukemia integration 1 transcription factor; IRAK1, IL-1 receptor-associated kinase 1; IRF, interferon regulatory factor; LAP, latency-associated peptide; miR, microRNA; MMP1, metalloproteinase-1; NLRP1, nucleotide-binding oligomerization domain, leucine-rich repeat and pyrin domain-containing protein 1; PDGF, platelet-derived growth factor; PDGFRβ, platelet-derived growth factor receptor β; SFRP1, secreted frizzled-related protein 1; SMAD, mothers against decapentaplegic homologue; STAT4, signal transducer and activator of transcription 4; TBX21, T-Box 21 transcription factor (T-bet); TCR, T-cell receptor; TGF-β, transforming growth factor β; TNIP1, TNFAIP3-interacting protein 1 protein 1; TNFSF4, TNF ligand superfamily member 4.

mechanisms can explain the inhibitory effect of TSA on fibroblast activation. First, TSA can restore the expression of the negative regulator of collagen expression *FLI1*, which is usually repressed in SSc by promoter hypoacetylation and DNA methylation (as described above).[59] Furthermore, TSA reportedly inhibits the activation of the transcription factors SMAD3 and SMAD4 by TGF-β, thereby interfering with the signalling pathway involved in tissue fibrosis.[62] In addition, the strong inhibition of *HDAC7* expression by TSA in turn downregulates the expression of *COL1A1* and *COL3A1*, both in resting cells and in those stimulated with TGF-β.[63]

These results suggest that HDAC inhibitors might be an effective treatment for skin fibrosis. However, given that HDACs are involved in crucial processes such as development and cell differentiation, treatment specificity should be carefully evaluated. In addition, as well as inhibiting the expression of collagen proteins in fibroblasts in SSc, TSA concurrently upregulated the expression of other profibrotic molecules, such as connective tissue growth factor and intercellular adhesion molecule-1, that might contribute substantially to the development of skin fibrosis.[63] Finally, the levels of expression of HDAC genes in SSc skin might also influence the effectiveness of HDAC inhibitors; however, reports from the literature are conflicting. Hemmatazad *et al.*[63] reported similar levels of expression of all HDAC genes (*HDAC1* to *HDAC11*) between fibroblast from patients with SSc and controls, whereas others[59] showed increased levels of HDAC1 and HDAC6 protein in fibroblasts in the context of SSc.

In contrast to HDACs, the protein level of histone acetyltransferase p300 has been shown to be consistently increased in fibroblasts and skin biopsy samples from patients with SSc in multiple studies compared with healthy samples.[64,65] Interestingly, maximal stimulation of collagen synthesis by TGF-β depends on the activity of p300,[66] which facilitates the access of TGF-β-activated SMAD2/3 to the *COL1A2* promoter by increasing its acetylation.[64] Given that TGF-β itself can induce p300 expression via the activation of Egr1, a profibrotic transcription factor that is also upregulated in biopsy samples from patients with SSc,[67] the overexpression of p300 is thought to promote an epigenetic feed-forward amplification of fibrosis in SSc in concert with TGF-β and Egr1.

Histone H3 Lys27 trimethylation in fibroblasts

Whereas DNMT and HDAC inhibitors have antifibrotic effects, 3-deazaneplanocin A, a potent inhibitor of trimethylation of histone H3 on lysine 27, stimulates fibroblasts to release collagen and exacerbates bleomycin-induced and TGF-β-receptor-induced experimental fibrosis *in vivo*,[68] indicating a clear role for this histone modification as a negative regulator of fibrosis. However, although the total levels of trimethylated histone H3 are increased in fibroblasts from patients with SSc as compared with those from healthy controls, they are not sufficient to block excessive collagen deposition by such fibroblasts.[68]

Histone modification in B cells

The alteration of histone modifications in SSc is not confined to fibroblasts. Global hyperacetylation of histone H4 and hypomethylation of H3 on lysine 9 have been detected in B cells from patients with SSc; these changes correlated with altered levels of the histone modification enzymes HDAC2 and SUV39H2 and clinical parameters, such as skin thickness and disease activity.[69] Even if these results are purely descriptive (as in, do not show functional analysis), they suggest that the abnormal histone modification pattern of autoimmune-related genes might contribute to B-cell activation and the development of autoimmunity.

miRNAs

Expression profiling has revealed the presence of a different miRNA pattern in skin[70,71] and fibroblasts cultured *ex vivo*[72,73] from patients with SSc compared with healthy controls. Functional studies subsequently demonstrated that differentially expressed miRNAs have a crucial role in SSc fibrosis by modulating multiple fibrosis-related genes such as those encoding collagens, metallopeptidases and integrins (Table 2).

miR-21

A detailed study by Zhu *et al.*[74] demonstrated that increased levels of miR-21 contribute to the upregulation of fibrosis-related genes (for example, *ACTA2*, *COL1A1*, *COL1A2*, and *FN1*) in fibroblasts from patients with SSc by negatively regulating SMAD7, an endogenous feedback inhibitor of TGF-β signalling, and therefore perpetuating the action of TGF-β. Consistent with these findings, the expression of miR-21 is also increased in other fibrotic conditions, such as post-ischaemic reperfusion, heart failure and idiopathic pulmonary fibrosis,[75,76] and in bleomycin-induced skin fibrosis; by contrast, the chemical inhibition of fibrosis by bortezomib restored the levels of miR-21 and Smad7 in bleomycin-treated mice.[74] The upregulation of miR-21 in the skin of patients with SSc is probably caused by exposure to TGF-β, which is known to induce the expression of this miRNA *in vitro*.[74–76]

miR-29a

The miR-29 family (miR-29a, miR-29b and miR-29c) is a class of antifibrotic miRNAs that oppose the profibrotic activity of miR-21. Decreased expression of miR-29 is associated with multiple fibrotic conditions, including cardiac, pulmonary and liver fibrosis,[77] and three distinct studies have clearly demonstrated that miR-29 levels are also decreased both in skin tissues and fibroblasts from patients with SSc that are cultured *ex vivo*.[71,78,79] The profibrotic stimuli TGF-β, PDGF or IL-4 inhibited miR-29a both *in vitro* and *in vivo*,[78] suggesting that they are also responsible for its downregulation in the skin of patients with SSc. Given that miR-29 was shown to directly repress the expression of *COL1A1*, *COL1A2* and *COL3A1*, low levels of this miRNA in patients with SSc are likely to lead to the uncontrolled accumulation of ECM proteins, thereby contributing to a positive-feedback loop that promotes fibrosis.[78]

Additional miRNAs

In addition to the well-established role of miR-21 and miR-29a in driving fibrosis in SSc, other miRNAs that are potentially involved in SSc pathogenesis include let-7a, miR-196a, miR-129-5p, miR-30b and miR-92a. Given that TGF-β can modulate the expression of these miRNAs *in vitro* and *in vivo* in a similar manner as observed in SSc-skin or fibroblasts,[72,73] their dysregulation is probably the consequence of the excessive exposure of fibroblasts in patients with SSc to TGF-β. Although reports that link these miRNAs to SSc are sporadic, functional studies indicate that they might be actively involved in the ongoing fibrosis in SSc. Downregulation of let-7a contributes to the accumulation of type I collagen in fibroblasts in the context of SSc, and the intermittent overexpression of let-7a in the skin by intraperitoneal miRNA injection improved skin fibrosis induced by bleomycin in mice.[80]

In fibroblasts from healthy donors, miR-196a functions to inhibit the production of type I collagen[72] and its expression is under the control of the discoidin domain receptor 2 (DDR2).[81] However, in fibroblasts from patients with SSc, the increased production of TGF-β directly inhibits the expression of both miR-196a and *DDR2*, thereby impairing the miR-196a-mediated negative feedback that controls type I collagen production and tissue fibrosis.[81] Interestingly, the observation of decreased levels of miR-196a in the hair shafts of patients with SSc[82] indicates that miR-196a might be considered as a noninvasive biomarker for SSc diagnosis.

Normally, IL-17A induces the expression of miR-129-5p, which is responsible for the IL-17-mediated inhibition of collagen expression at the post-transcriptional level, but in fibroblasts from patients with SSc, miR-129-5p is downregulated as a result of TGF-β inhibiting the expression of *IL17RA* (which encodes the IL-17 receptor type A) and, consequently, impairing IL-17 signalling.[83]

Decreased expression of miR-30b maintains the TGF-β-mediated induction of the PDGF receptor and fibroblast proliferation.[84,85] Moreover, the consequence of miR-92a overexpression might contribute to excessive collagen accumulation in fibroblasts from patients with SSc through the downregulation of MMP1.[86]

As well as a prominent role for TGF-β in regulating the expression of SSc-associated miRNAs, other factors can also influence the expression and, consequently, the activity of miRNAs in fibroblasts in SSc. Constitutive downregulation of miR-150, possibly through hypermethylation of its promoter, has been proposed to cause the upregulation of its direct target, β3-integrin, thus promoting the binding of latency associated peptide to the cell surface and the release of bioactive TGF-β.[73] By contrast, the decreased intracellular levels of the profibrotic factor thrombospondin 2 that occur in fibroblasts in the context of SSc result in the upregulation of miR-7. Although miR-7 inhibits type I collagen production, the potential negative-feedback circuit induced by the upregulation of miR-7 seems to be insufficient to normalize the excessive collagen synthesis in fibroblasts

from patients with SSc.[87] These results require further investigation to reconcile this apparent contradiction.

Taken together, these data indicate that the pathogenesis of SSc is influenced by aberrant gene expression arising from defective post-transcriptional control mediated by miRNAs. The modulation of miRNA expression by using miRNA inhibitors or mimics might, therefore, represent an attractive future opportunity for SSc treatment.

Circulating miRNAs

The detection of specific circulating miRNAs that correlate with SSc subset or disease severity could address the unmet need for reliable and accurate biomarkers of SSc. Although a comprehensive profile of circulating miRNAs is lacking, several studies have demonstrated that the levels of selected miRNAs are altered in the serum of patients with SSc.

Consistent with their expression in skin or fibroblasts in the context of SSc, the levels of miR-150, let-7a and miR-30a were, respectively, upregulated and downregulated in the serum of patients with SSc versus healthy controls, indicating that these miRNAs could be used as easily accessible biomarkers for SSc diagnosis.[73,80,84] Interestingly, patients with SSc who have the lowest serum levels of miR-150 showed an increased frequency of pitting scars or ulcers and the presence of ATA, whereas miR-30a levels were inversely correlated with modified Rodnan skin scores;[73,84] however, these correlations seem more sporadic than systematically associated with disease severity or a specific subclass.

The serum concentrations of let-7g, miR-21, miR-29a, miR-125b, miR-145, miR-206 and miR-196a are similar in patients with SSc and healthy controls, despite differences in the skin; some of them, however, correlate weakly with SSc clinical parameters.[72,88,89] For example, patients with lower miR-196a serum levels had a higher modified Rodnan skin score and an increased prevalence of pitting scars compared with those with normal miR-196a levels.[72]

Serum levels of miR-92a and miR-142-3p were markedly higher in patients with SSc than in healthy individuals or in patients with SLE, dermatomyositis or scleroderma spectrum disorders,[86,90] indicating that these miRNAs might provide useful diagnostic markers for the differentiation of SSc from other autoimmune diseases. Larger, and more detailed, studies will be required to better evaluate the potential of these miRNAs as diagnostic or prognostic biomarkers for SSc.

Conclusions

Despite extensive efforts, genetic research to date is not able to explain fully the development of SSc, nor has implementation of any of the findings from this research into clinical practice been achieved. The main gaps in the value of genetic research in SSc are in the field of clinical translation. A powerful genetic approach would be to investigate the predictive value of specific genetic markers for severe complications, such as pulmonary hypertension and pulmonary fibrosis. Baseline measurements of these genetic markers could indicate the need to monitor the potential development of these complications more tightly in patients at risk, or even allow for preventive care. Only a few genetic studies are currently addressing this issue, using follow-up data. This approach would position this field of research more prominently in clinical decision-making and would greatly enhance its scientific value.

A difficult concept to overcome is that SSc is still regarded as a multifactorial disease, such that clusters of genetic variants are thought to contribute to genetic susceptibility as a whole. Logically, the next step would be to apply gene–gene interaction analysis to investigate whether certain combinations of variants contribute to the overall risk of genetic susceptibility. This approach has the potential to identify cumulative effects, which might, therefore, account for a more substantial part of SSc disease risk than current individual genetic markers. Hopefully, techniques that address the high computational burden that is pivotal to perform these analyses will evolve alongside this strategy.

Of interest is the high degree of overlap between genetic risk factors for SSc and those for other autoimmune diseases,[91] known as 'shared autoimmunity'. This concept implies that additional factors might regulate the development of a distinct autoimmune disease on a similar genetic background. These factors are probably epigenetic modifications, environmental factors or rare genetic variants. Although likely to be involved in a rare disease like SSc, rare variants have not been investigated with a similar scrutiny to epigenetic modifications or environmental factors. The most powerful way to identify the role of such rare variants would be to perform genetic association studies in families with SSc, which are rare. New technologies, such as next-generation whole-exome or genome sequencing, should hopefully enable a more definitive answer as to what drives the pathogenesis of SSc at the genetic level, as should DNA methylation profiling. Investigating methylation patterns, which indicate genomic regions that contribute to the epigenetic inheritance involved in the pathogenesis of SSc, would be informative in a discordant monozygotic twin study, but also for comparing autoimmune diseases with similar genetic backgrounds.

Furthermore, analysis of the crosstalk between SNPs and epigenetic modifications might clarify how SSc-susceptibility SNPs located in noncoding regions can influence gene expression to ultimately lead to SSc pathogenesis. Disease-associated SNPs are increasingly recognized to have regulatory functions linked to epigenetics, given that they are enriched on the promoters of coding genes and in regions characterized by multiple chromatin marks.[92,93] Consistent with these findings, a decreased number of activating epigenomic marks in SSc-linked susceptibility genes have been reported.[94] Finally, the presence of SNPs can also influence DNA methylation patterns,[95] transcription-factor binding to promoter regions,[96] and the sequence of miRNAs and their recognition elements,[97,98] thereby potentially influencing gene and miRNA activity and leading to SSc development.

Several reports have demonstrated the power of a multiple-data-layers integration approach in identifying molecular signatures and networks that orchestrate both physiological and pathological processes, such as the immune response and cancer.[99,100] It is hoped that a similar systems-biology approach to further investigate the interplay between genetic, epigenetic and environmental factors will help to unravel the complex molecular circuitry that orchestrates SSc pathogenesis, ultimately leading to the generation of effective treatments for this chronic inflammatory condition.

Review criteria

A search for original articles published 2000–2014 was performed using PubMed with the terms "genetics", "SNPs", "GWAS", "family", "genome-wide", "microRNA", "methylation", "acetylation", "histone modifications", "chromatin" and "SSc or scleroderma or systemic sclerosis", alone and in combination. Publications based on primary findings in systemic sclerosis tissue compared with healthy or other autoimmune conditions were included; studies discussing only general fibrosis conditions or models were omitted. All articles identified were in English.

1. Varga, J. & Abraham, D. Systemic sclerosis: a prototypic multisystem fibrotic disorder. *J. Clin. Invest.* **117**, 557–567 (2007).
2. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. *Arthritis Rheum.* **23**, 581–590 (1980).
3. Steen, V. D. Autoantibodies in systemic sclerosis. *Semin. Arthritis Rheum.* **35**, 35–42 (2005).
4. Broen, J. C., Coenen, M. J. & Radstake, T. R. Deciphering the genetic background of systemic sclerosis. *Expert Rev. Clin. Immunol.* **7**, 449–462 (2011).
5. Joseph, C. G. *et al.* Association of the autoimmune disease scleroderma with an immunologic response to cancer. *Science* **343**, 152–157 (2014).
6. McCormic, Z. D., Khuder, S. S., Aryal, B. K., Ames, A. L. & Khuder, S. A. Occupational silica exposure as a risk factor for scleroderma: a meta-analysis. *Int. Arch. Occup. Environ. Health* **83**, 763–769 (2010).
7. Hudson, M. *et al.* Polyautoimmunity and familial autoimmunity in systemic sclerosis. *J. Autoimmun.* **31**, 156–159 (2008).
8. Englert, H. *et al.* Familial risk estimation in systemic sclerosis. *Aust. N. Z. J. Med.* **29**, 36–41 (1999).
9. Arora-Singh, R. K. *et al.* Autoimmune diseases and autoantibodies in the first degree relatives of patients with systemic sclerosis. *J. Autoimmun.* **35**, 52–57 (2010).
10. Arnett, F. C. *et al.* Familial occurrence frequencies and relative risks for systemic sclerosis (scleroderma) in three United States cohorts. *Arthritis Rheum.* **44**, 1359–1362 (2001).
11. Frech, T. *et al.* Heritability of vasculopathy, autoimmune disease, and fibrosis in systemic sclerosis: a population-based study. *Arthritis Rheum.* **62**, 2109–2116 (2010).
12. Feghali-Bostwick, C., Medsger, T. A., Jr & Wright, T. M. 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 (2003).
13. Broen, J. C., Coenen, M. J. & Radstake, T. R. Genetics of systemic sclerosis: an update. *Curr. Rheumatol. Rep.* **14**, 11–21 (2012).
14. Tsuchiya, N. *et al.* Association of *STAT4* polymorphism with systemic sclerosis in a Japanese population. *Ann. Rheum. Dis.* **68**, 1375–1376 (2009).
15. Rueda, B. *et al.* The *STAT4* gene influences the genetic predisposition to systemic sclerosis phenotype. *Hum. Mol. Genet.* **18**, 2071–2077 (2009).
16. Gourh, P. *et al.* Polymorphisms in *TBX21* and *STAT4* increase the risk of systemic sclerosis: evidence of possible gene-gene interaction and alterations in T_H1/T_H2 cytokines. *Arthritis Rheum.* **60**, 3794–3806 (2009).
17. Dieude, P. *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.* **60**, 2472–2479 (2009).
18. Avouac, J. *et al.* Inactivation of the transcription factor STAT-4 prevents inflammation-driven fibrosis in animal models of systemic sclerosis. *Arthritis Rheum.* **63**, 800–809 (2011).
19. Korman, B. D., Kastner, D. L., Gregersen, P. K. & Remmers, E. F. *STAT4*: genetics, mechanisms, and implications for autoimmunity. *Curr. Allergy Asthma Rep.* **8**, 398–403 (2008).
20. Redmond, W. L., Ruby, C. E. & Weinberg, A. D. The role of OX40-mediated co-stimulation in T-cell activation and survival. *Crit. Rev. Immunol.* **29**, 187–201 (2009).
21. Bossini-Castillo, L. *et al.* A replication study confirms the association of *TNFSF4* (*OX40L*) polymorphisms with systemic sclerosis in a large European cohort. *Ann. Rheum. Dis.* **70**, 638–641 (2011).
22. Gourh, P. *et al.* Association of the C8orf13-BLK region with systemic sclerosis in North-American and European populations. *J. Autoimmun.* **34**, 155–162 (2010).
23. Aliprantis, A. O. *et al.* Transcription factor T-bet regulates skin sclerosis through its function in innate immunity and via IL-13. *Proc. Natl Acad. Sci. USA* **104**, 2827–2830 (2007).
24. Lakos, G., Melichian, D., Wu, M. & Varga, J. Increased bleomycin-induced skin fibrosis in mice lacking the T_H1-specific transcription factor T-bet. *Pathobiology* **73**, 224–237 (2006).
25. Guarda, G. *et al.* T cells dampen innate immune responses through inhibition of NLRP1 and NLRP3 inflammasomes. *Nature* **460**, 269–273 (2009).
26. Dieude, P. *et al.* NLRP1 influences the systemic sclerosis phenotype: a new clue for the contribution of innate immunity in systemic sclerosis-related fibrosing alveolitis pathogenesis. *Ann. Rheum. Dis.* **70**, 668–674 (2011).
27. Orozco, G. *et al.* Study of the common genetic background for rheumatoid arthritis and systemic lupus erythematosus. *Ann. Rheum. Dis.* **70**, 463–468 (2011).
28. Ito, I. *et al.* Association of the FAM167A-BLK region with systemic sclerosis. *Arthritis Rheum.* **62**, 890–895 (2010).
29. Dymecki, S. M., Niederhuber, J. E. & Desiderio, S. V. Specific expression of a tyrosine kinase gene, *blk*, in B lymphoid cells. *Science* **247**, 332–336 (1990).
30. Yokoyama, K. *et al.* BANK regulates BCR-induced calcium mobilization by promoting tyrosine phosphorylation of IP(3) receptor. *EMBO J.* **21**, 83–92 (2002).
31. Gourh, P. *et al.* Association of the C8orf13-BLK region with systemic sclerosis in North-American and European populations. *J. Autoimmun.* **34**, 155–162 (2010).
32. Kozyrev, S. V. *et al.* Functional variants in the B-cell gene *BANK1* are associated with systemic lupus erythematosus. *Nat. Genet.* **40**, 211–216 (2008).
33. Orozco, G. *et al.* Study of functional variants of the *BANK1* gene in rheumatoid arthritis. *Arthritis Rheum.* **60**, 372–379 (2009).
34. Dieude, P. *et al. BANK1* is a genetic risk factor for diffuse cutaneous systemic sclerosis and has additive effects with *IRF5* and *STAT4*. *Arthritis Rheum.* **60**, 3447–3454 (2009).
35. Rueda, B. *et al. BANK1* functional variants are associated with susceptibility to diffuse systemic sclerosis in Caucasians. *Ann. Rheum. Dis.* **69**, 700–705 (2010).
36. Carmona, F. D. *et al.* New insight on the Xq28 association with systemic sclerosis. *Ann. Rheum. Dis.* **72**, 2032–2038 (2013).
37. Dieudé, P. *et al.* Evidence of the contribution of the X chromosome to systemic sclerosis susceptibility: association with the functional *IRAK1* 196Phe/532Ser haplotype. *Arthritis Rheum.* **63**, 3979–3987 (2011).
38. Zhou, X. *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.* **60**, 3807–3814 (2009).
39. Radstake, T. R. *et al.* Genome-wide association study of systemic sclerosis identifies *CD247* as a new susceptibility locus. *Nat. Genet.* **42**, 426–429 (2010).
40. Dieude, P. *et al.* Independent replication establishes the *CD247* gene as a genetic systemic sclerosis susceptibility factor. *Ann. Rheum. Dis.* **70**, 1695–1696 (2011).
41. Allanore, Y. *et al.* Genome-wide scan identifies *TNIP1*, *PSORS1C1*, and *RHOB* as novel risk loci for systemic sclerosis. *PLoS Genet.* **7**, e1002091 (2011).
42. Bossini-Castillo, L. *et al.* 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 (2013).
43. López-Isac, E. *et al.* A genome-wide association study follow-up suggests a possible role for *PPARG* in systemic sclerosis susceptibility. *Arthritis Res. Ther.* **16**, R6 (2014).
44. Bossini-Castillo, L. *et al.* A GWAS follow-up study reveals the association of the *IL12RB2* gene with systemic sclerosis in Caucasian populations. *Hum. Mol. Genet.* **21**, 926–933 (2012).
45. Martin, J. E. *et al.* A systemic sclerosis and systemic lupus erythematosus pan-meta-GWAS reveals new shared susceptibility loci. *Hum. Mol. Genet.* **22**, 4021–4029 (2013).
46. Mayes, M. D. *et al.* Immunochip analysis identifies multiple susceptibility loci for systemic sclerosis. *Am. J. Hum. Genet.* **94**, 47–61 (2014).

47. Gorlova, O. *et al.* Identification of novel genetic markers associated with clinical phenotypes of systemic sclerosis through a genome-wide association strategy. *PLoS Genet.* **7**, e1002178 (2011).
48. Katsumoto, T. R., Whitfield, M. L. & Connolly, M. K. The pathogenesis of systemic sclerosis. *Annu. Rev. Pathol.* **6**, 509–537 (2011).
49. Marie, I. *et al.* Prospective study to evaluate the association between systemic sclerosis and occupational exposure and review of the literature. *Autoimmun. Rev.* **13**, 151–156 (2014).
50. Mora, G. F. Systemic sclerosis: environmental factors. *J. Rheumatol.* **36**, 2383–2396 (2009).
51. Ballestar, E. Epigenetic alterations in autoimmune rheumatic diseases. *Nat. Rev. Rheumatol.* **7**, 263–271 (2011).
52. Meda, F., Folci, M., Baccarelli, A. & Selmi, C. The epigenetics of autoimmunity. *Cell Mol. Immunol.* **8**, 226–236 (2011).
53. Lei, W. *et al.* Abnormal DNA methylation in $CD4^+$ T cells from patients with systemic lupus erythematosus, systemic sclerosis, and dermatomyositis. *Scand. J. Rheumatol.* **38**, 369–374 (2009).
54. Jiang, H. *et al.* Demethylation of TNFSF7 contributes to CD70 overexpression in $CD4^+$ T cells from patients with systemic sclerosis. *Clin. Immunol.* **143**, 39–44 (2012).
55. Komura, K. *et al.* Blockade of CD40/CD40 ligand interactions attenuates skin fibrosis and autoimmunity in the tight-skin mouse. *Ann. Rheum. Dis.* **67**, 867–872 (2008).
56. Lian, X. *et al.* DNA demethylation of CD40L in $CD4^+$ T cells from women with systemic sclerosis: a possible explanation for female susceptibility. *Arthritis Rheum.* **64**, 2338–2345 (2012).
57. Selmi, C. *et al.* X chromosome gene methylation in peripheral lymphocytes from monozygotic twins discordant for scleroderma. *Clin. Exp. Immunol.* **169**, 253–262 (2012).
58. Wang, Y. Y. *et al.* DNA hypermethylation of the Forkhead box protein 3 (*FOXP3*) promoter in $CD4^+$ T cells of patients with systemic sclerosis. *Br. J. Dermatol.* http://dx.doi.org/10.1111/bjd.12913.
59. Wang, Y., Fan, P. S. & Kahaleh, B. Association between enhanced type I collagen expression and epigenetic repression of the *FLI1* gene in scleroderma fibroblasts. *Arthritis Rheum.* **54**, 2271–2279 (2006).
60. Dees, C. *et al.* The Wnt antagonists DKK1 and SFRP1 are downregulated by promoter hypermethylation in systemic sclerosis. *Ann. Rheum. Dis.* **73**, 1232–1239 (2013).
61. Wang, Y. & Kahaleh, B. Epigenetic repression of bone morphogenetic protein receptor II expression in scleroderma. *J. Cell Mol. Med.* **17**, 1291–1299 (2013).
62. Huber, L. C. *et al.* Trichostatin A prevents the accumulation of extracellular matrix in a mouse model of bleomycin-induced skin fibrosis. *Arthritis Rheum.* **56**, 2755–2764 (2007).
63. Hemmatazad, H. *et al.* Histone deacetylase 7, a potential target for the antifibrotic treatment of systemic sclerosis. *Arthritis Rheum.* **60**, 1519–1529 (2009).
64. Ghosh, A. K. *et al.* p300 is elevated in systemic sclerosis and its expression is positively regulated by TGF-β: epigenetic feed-forward amplification of fibrosis. *J. Invest. Dermatol.* **133**, 1302–1310 (2013).
65. Bhattacharyya, S. *et al.* Fibroblast expression of the coactivator p300 governs the intensity of profibrotic response to transforming growth factor β. *Arthritis Rheum.* **52**, 1248–1258 (2005).
66. Ghosh, A. K., Yuan, W., Mori, Y. & Varga, J. Smad-dependent stimulation of type I collagen gene expression in human skin fibroblasts by TGF-β involves functional cooperation with p300/CBP transcriptional coactivators. *Oncogene* **19**, 3546–3555 (2000).
67. Bhattacharyya, S. *et al.* Egr-1 induces a profibrotic injury/repair gene program associated with systemic sclerosis. *PLoS ONE* **6**, e23082 (2011).
68. Kramer, M. *et al.* Inhibition of H3K27 histone trimethylation activates fibroblasts and induces fibrosis. *Ann. Rheum. Dis.* **72**, 614–620 (2013).
69. Wang, Y. *et al.* Aberrant histone modification in peripheral blood B cells from patients with systemic sclerosis. *Clin. Immunol.* **149**, 46–54 (2013).
70. Li, H. *et al.* MicroRNA array analysis of microRNAs related to systemic scleroderma. *Rheumatol. Int.* **32**, 307–313 (2012).
71. Zhu, H. *et al.* MicroRNA expression abnormalities in limited cutaneous scleroderma and diffuse cutaneous scleroderma. *J. Clin. Immunol.* **32**, 514–522 (2012).
72. Honda, N. *et al.* TGF-β-mediated downregulation of microRNA-196a contributes to the constitutive upregulated type I collagen expression in scleroderma dermal fibroblasts. *J. Immunol.* **188**, 3323–3331 (2012).
73. Honda, N. *et al.* miR-150 down-regulation contributes to the constitutive type I collagen overexpression in scleroderma dermal fibroblasts via the induction of integrin β3. *Am. J. Pathol.* **182**, 206–216 (2013).
74. Zhu, H. *et al.* MicroRNA-21 in scleroderma fibrosis and its function in TGF-β-regulated fibrosis-related genes expression. *J. Clin. Immunol.* **33**, 1100–1109 (2013).
75. Vettori, S., Gay, S. & Distler, O. Role of MicroRNAs in fibrosis. *Open. Rheumatol. J.* **6**, 130–139 (2012).
76. Jiang, X., Tsitsiou, E., Herrick, S. E. & Lindsay, M. A. MicroRNAs and the regulation of fibrosis. *FEBS J.* **277**, 2015–2021 (2010).
77. Peng, W. J. *et al.* MicroRNA-29: a potential therapeutic target for systemic sclerosis. *Expert Opin. Ther. Targets.* **16**, 875–879 (2012).
78. Maurer, B. *et al.* MicroRNA-29, a key regulator of collagen expression in systemic sclerosis. *Arthritis Rheum.* **62**, 1733–1743 (2010).
79. Bhattacharyya, S. *et al.* Toll-like receptor 4 signaling augments transforming growth factor-β responses: a novel mechanism for maintaining and amplifying fibrosis in scleroderma. *Am. J. Pathol.* **182**, 192–205 (2013).
80. Makino, K. *et al.* The downregulation of microRNA let-7a contributes to the excessive expression of type I collagen in systemic and localized scleroderma. *J. Immunol.* **190**, 3905–3915 (2013).
81. Makino, K. *et al.* Discoidin domain receptor 2-microRNA 196a-mediated negative feedback against excess type I collagen expression is impaired in scleroderma dermal fibroblasts. *J. Invest Dermatol.* **133**, 110–119 (2013).
82. Wang, Z. *et al.* Detection of hair-microRNAs as the novel potent biomarker: evaluation of the usefulness for the diagnosis of scleroderma. *J. Dermatol. Sci.* **72**, 134–141 (2013).
83. Nakashima, T. *et al.* Impaired IL-17 signaling pathway contributes to the increased collagen expression in scleroderma fibroblasts. *J. Immunol.* **188**, 3573–3583 (2012).
84. Tanaka, S. *et al.* Alteration of circulating miRNAs in SSc: miR-30b regulates the expression of PDGF receptor β. *Rheumatology (Oxford)* **52**, 1963–1972 (2013).
85. Yamakage, A., Kikuchi, K., Smith, E. A., LeRoy, E. C. & Trojanowska, M. Selective upregulation of platelet-derived growth factor α receptors by transforming growth factor β in scleroderma fibroblasts. *J. Exp. Med.* **175**, 1227–1234 (1992).
86. Sing, T. *et al.* microRNA-92a expression in the sera and dermal fibroblasts increases in patients with scleroderma. *Rheumatology. (Oxford)* **51**, 1550–1556 (2012).
87. Kajihara, I. *et al.* Increased accumulation of extracellular thrombospondin-2 due to low degradation activity stimulates type I collagen expression in scleroderma fibroblasts. *Am. J. Pathol.* **180**, 703–714 (2012).
88. Koba, S. *et al.* Expression analysis of multiple microRNAs in each patient with scleroderma. *Exp. Dermatol.* **22**, 489–491 (2013).
89. Kawashita, Y. *et al.* Circulating miR-29a levels in patients with scleroderma spectrum disorder. *J. Dermatol. Sci.* **61**, 67–69 (2011).
90. Makino, K. *et al.* Circulating miR-142-3p levels in patients with systemic sclerosis. *Clin. Exp. Dermatol.* **37**, 34–39 (2012).
91. Broen, J. C. & Radstake, T. R. How birds of a feather flock together: genetics in autoimmune diseases. *Expert Rev. Clin. Immunol.* **7**, 127–128 (2011).
92. Schaub, M. A., Boyle, A. P., Kundaje, A., Batzoglou, S. & Snyder, M. Linking disease associations with regulatory information in the human genome. *Genome Res.* **22**, 1748–1759 (2012).
93. Maurano, M. T. *et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012).
94. Dozmorov, M. G., Wren, J. D. & Alarcón-Riquelme, M. E. Epigenomic elements enriched in the promoters of autoimmunity susceptibility genes. *Epigenetics* **9**, 276–285 (2013).
95. Bell, J. T. *et al.* DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol.* **12**, R10 (2011).
96. Kasowski, M. *et al.* Variation in transcription factor binding among humans. *Science* **328**, 232–235 (2010).
97. Saunders, M. A., Liang, H. & Li, W. H. Human polymorphism at microRNAs and microRNA target sites. *Proc. Natl Acad. Sci. USA* **104**, 3300–3305 (2007).
98. Sethupathy, P. & Collins, F. S. MicroRNA target site polymorphisms and human disease. *Trends Genet.* **24**, 489–497 (2008).
99. Li, S. *et al.* Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. *Nat. Immunol.* **15**, 195–204 (2014).
100. Werner, H. M., Mills, G. B. & Ram, P. T. Cancer systems biology: a peek into the future of patient care? *Nat. Rev. Clin. Oncol.* **11**, 167–176 (2014).
101. Farina, A. *et al*. Epstein–Barr virus infection induces aberrant TLR activation pathway and fibroblast-myofibroblast conversion in scleroderma. *J. Invest. Dermatol.* **134**, 954–964 (2014).

Acknowledgements

J.C.A.B. is sponsored by the VENI Laureate from the Dutch Association of Research (NWO) and M.R. is funded by a Marie Curie Intra-European Fellowship.

Author contributions

J.C.A.B. and M.R. researched data for the article. All authors provided substantial contributions to discussions of its content, wrote the article and undertook review and/or editing of the manuscript before submission.

Supplementary information is linked to the online version of the paper at www.nature.com/nrrheum.