



Published in final edited form as:

Ann Rheum Dis. 2017 June ; 76(6): 1150–1158. doi:10.1136/annrheumdis-2016-210645.

Transethnic meta-analysis identifies *GSDMA* and *PRDM1* as susceptibility genes to systemic sclerosis

A full list of authors and affiliations appears at the end of the article.

Abstract

Objectives—Systemic sclerosis (SSc) is an autoimmune disease characterised by skin and systemic fibrosis culminating in organ damage. Previous genetic studies including genome-wide association studies (GWAS) have identified 12 susceptibility loci satisfying genome-wide significance. Transethnic meta-analyses have successfully expanded the list of susceptibility genes and deepened biological insights for other autoimmune diseases.

Methods—We performed transethnic meta-analysis of GWAS in the Japanese and European populations, followed by a two-staged replication study comprising a total of 4436 cases and 14 751 controls. Associations between significant single nuclear polymorphisms (SNPs) and neighbouring genes were evaluated. Enrichment analysis of H3K4Me3, a representative histone mark for active promoter was conducted with an expanded list of SSc susceptibility genes.

Results—We identified two significant SNP in two loci, *GSDMA* and *PRDM1*, both of which are related to immune functions and associated with other autoimmune diseases ($p=1.4\times 10^{-10}$ and 6.6×10^{-10} , respectively). *GSDMA* also showed a significant association with limited cutaneous SSc. We also replicated the associations of previously reported loci including a non-GWAS locus, *TNFAIP3*. *PRDM1* encodes BLIMP1, a transcription factor regulating T-cell proliferation and plasma cell differentiation. The top SNP in *GSDMA* was a missense variant and correlated with gene expression of neighbouring genes, and this could explain the association in this locus. We found different human leukocyte antigen (HLA) association patterns between the two populations. Enrichment analysis suggested the importance of CD4-naïve primary T cell.

Conclusions—*GSDMA* and *PRDM1* are associated with SSc. These findings provide enhanced insight into the genetic and biological basis of SSc.

Correspondence to Dr Yannick Allanore, INSERM U1016/UMR 8104, Cochin Institute, Rheumatology A Department, Paris Descartes University, 75014 Paris, France; yannick.allanore@me.com or Dr Chikashi Terao, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Shogoin-Kawaharacho 54, Sakyo-ku, Kyoto 606-8507, Japan; a0001101@kuhp.kyoto-u.ac.jp.

Twitter Follow Soumya Raychaudhuri @soumya_boston

Contributors Wrote the paper: CT, YA. Performed imputation and analyses: CT, TK, MM Performed the experimental work: CT, AK, AO, MS, AC. Conceived and designed the study: CT, TM, FM, YA. Substantial contribution to acquired samples and creation of data in GWAS: CT, TK, PD, MK, YK, KO, TH, HY, TA, TT, MT, YT, MS, AT, AC, MM, TM, FM, YA. Substantial contribution to acquired samples and creation of data in replication study: CT, PD, JV, MH, MMC, KO, GR, AK, PA, TH, AO, EH, HY, PC, NH, MB, TA, PH, TT, MT, YT, MS, NA, HY, HF, ST, MH, MF, OI, TY, DG, YA, MJ, HE, HT, KT, SS, HI, SR, KL, PG, NT, VR, IM, GV, AC, TM, FM, YA. Contribution to collection of clinical information: CT, PD, JV, MK, MH, YK, MMC, KO, GR, PA, TH, AO, EH, PC, NH, MB, TA, PH, YT, AT, NA, HY, HF, ST, MH, MF, OI, TY, DG, YA, MJ, HE, HT, KT, SS, HI, VR, IM, GV, AC, TM, FM, YA. All authors revised and approved the manuscript to be published.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Ethics approval This study was approved by local ethical committees.

INTRODUCTION

Systemic sclerosis (SSc) is an orphan disease with high morbidity and mortality. It is composed of two main subsets, a limited cutaneous form (lcSSc) and a diffuse cutaneous form (dcSSc).¹ SSc is also characterised by production of specific autoantibodies, anticentromere antibody (ACA) and anti-Scl70 antibody. Severe complications in SSc include interstitial lung disease (ILD), digital ulcers (DU), renal crisis and pulmonary hypertension (PH), where fibrosis in tissues and vessel remodelling play fundamental roles.¹ Genetic and environmental elements are associated with the development of SSc.¹² While SSc is a heterogeneous disease, it has a significant genetic component.³ A total of 12 *non-HLA* loci showing significant associations ($p < 5.0 \times 10^{-8}$) were reported for their associations⁴⁻¹¹ (table 1).

In spite of a paradigm shift in the treatment of autoimmune diseases by biological agents,¹² treatment of SSc remains challenging and new molecular targets are still under investigation. Results of previous genome-wide association studies (GWAS) in other autoimmune diseases have successfully identified important pathways as molecular targets, leading to effective treatments.¹³⁻¹⁵ Similarly, GWAS on SSc may suggest novel targets for treatment.

To this end, transethnic meta-analysis of GWAS would be a promising way to identify unknown susceptibility genes which were difficult to detect in a single population due to lack of statistical power, different structure of linkage disequilibrium (LD) or different allele frequencies between populations. In fact, transethnic meta-analyses of GWAS for another autoimmune disease, rheumatoid arthritis (RA), have expanded lists of susceptibility genes and led to candidates for target cell types and molecules.¹⁶ However, most of the previous GWAS for SSc are mainly reported from the European population, with only one GWAS from the Asian population using 137 Korean patients.¹⁷ Thus, we performed GWAS for SSc using 716 Japanese cases,² and 1797 controls and performed transethnic meta-analysis of GWAS using the previous GWAS from the French population,⁵ with comparable numbers of subjects to Japanese GWAS (see online supplementary figure S1).

MATERIALS AND METHODS

Study design

The schematic view of the study design is illustrated in online supplementary figure S1. In brief, after a first analysis of Japanese and French GWAS data, we then performed two replication studies. In the first replication study, we used a Japanese cohort and a European cohort originating from several European countries. In the second replication study, we used Canadian and North American population with European decent. As for markers, we picked up a total of 33 single nucleotide polymorphisms (SNPs) for the first replication study based on the criteria of selection of candidate SNPs. We further selected seven SNPs fulfilling the criteria of selection for the second replication study.

Samples

A total of 1280 cases and 3660 controls in the Japanese population and 3156 cases and 11 091 controls in the European population were recruited. Break down of subjects are shown

in online supplementary tables S1 and S2. All case samples fulfilled the American College of Rheumatology classification criteria for SSc.¹⁸ Written informed consent was obtained from all the participants. This study was approved by local ethical committees.

Clinical information

Clinical information regarding subtypes of SSc defined by LeRoy *et al.*,¹⁹ but also ILD, PH, renal crisis, DU and possession of ACA, anti-Scl70 antibody and anti-RNA polymerase III antibody were collected. The clinical information was selected based on the importance of SSc outcome and the previous genetic studies identifying specific associations with SSc subtypes or phenotypes. Due to very low prevalence of renal crisis, PH and anti-RNA polymerase III antibody, we did not include these phenotypes for subtype-specific analysis. The availability of clinical information is shown in online supplementary table S1.

Genotyping

The French GWAS data were published previously and the methods are written elsewhere.⁵ Genotyping with a competitive allele-specific PCR system for replication in the European samples was performed in the LGC Genomics (Hoddesdon, UK). A part of the replication data in the European samples was obtained by imputation based on GWAS (see online supplementary tables S1 and S2). The Japanese samples in the GWAS and the first replication study were genotyped in Kyoto University and University of Tsukuba, Japan.

Imputation

Imputation and phasing were performed by MaCH software,²⁰ using the East Asian panel and European panel in the 1000 Genomes Project,²¹ as references for Japanese and French populations, respectively. After imputation, we performed quality control (see below). Imputation for the Japanese and French population was performed separately at Kyoto University in Japan and the INSERM UMR 1220 in France, respectively, and only summary statistics for the French imputation data were available due to restriction of data sharing policy of the control samples.

Quality control

We applied different quality control criteria in the two GWAS. The details are shown in online supplementary table S1. Since the current study, especially GWAS, had limited power to find signals in SNPs with low allele frequency (see online supplementary table S3), we filtered SNPs in each data set again after imputation based on allele frequency and used SNPs showing $r^2 > 0.5$ in the output of MaCH for the subsequent analyses (see online supplementary table S4). Since information of variants in the sex chromosomes was not available in the French GWAS, we focused on variants in the autosomal chromosomes.

Linkage disequilibrium between SNPs

LD structure was evaluated based on the 1000 Genomes data and our genotyping data. Statistical value for LD was calculated by Haploview²² or PLINK.²³

Selection of SNPs for the first replication study

For the first replication study, we picked up SNPs whose associations or the associations of other SNPs in the same region were not previously reported, and satisfying the following criteria: (1) whose p values for SSc susceptibility were $<1.0 \times 10^{-5}$ in the meta-analysis of the two GWAS, (2) whose p values for SSc susceptibility were $<2.0 \times 10^{-4}$ in the meta-analysis and whose p values were <0.05 in the text-based in silico analysis using Gene Relationships Across Implicated Loci (GRAIL) programme,²⁴ with use of previously reported genes in SSc as seeds or (3) whose p values for SSc subtypes were $<1.0 \times 10^{-6}$. When multiple SNPs in the same region ($r^2 > 0.5$) satisfied the above criteria, we picked up SNPs showing the best p values or SNPs for which probe and primer design for replication studies was not technically difficult.

Selection of SNPs for the second replication study

For the second replication study, we selected SNPs (1) whose p values in the meta-analysis of the GWAS and first replication studies were $<5.0 \times 10^{-6}$ or (2) whose p values in the GWAS were $<2.0 \times 10^{-5}$ and whose p values were <0.05 by GRAIL.

Calculation of variance explained by susceptibility SNPs

We evaluated variance explained by new susceptibility SNPs based on liability-scale threshold model. We assumed that there are underlying liability scores following normal distribution and that subjects having a liability score over a predefined threshold to develop SSc. We set prevalence of SSc as 0.05%. OR in the overall study was used as approximation of common relative risk between populations. We performed this estimation separately in each population using control allele frequencies in GWAS.

Associations between clinical manifestation and associated SNPs

After confirming the associations of the two SNPs with SSc, the associations between the two SNPs and SSc subtypes or clinical manifestations including ILD and DU were estimated. We did not perform GWAS for these clinical manifestations due to limited number of subjects who were positive for these manifestations. We further defined two phenotypes, fibrotic and vascular and performed association studies with these two phenotypes. Fibrotic phenotype includes dcSSc or severe lung disease defined by forced volume capacity (FVC) $<70\%$ or ILD in combination with FVC $<75\%$. Vascular phenotype is DU, pulmonary arterial hypertension (PAH) or renal crisis.

Amino acid conservation search

We assessed amino acid conservation for the residue of GSDMA altered by rs3894194 across vertebrates in combination with Genomic Evolutionary Rate Profiling (GERP) score by using UCSC Genome Browser. GERP score was calculated for the single amino acid residue in which positive score indicated conservation.

HLA imputation and analyses

HLA alleles and amino acids were imputed by SNP2HLA.²⁵ We used Asian reference panel and European reference panel for Japanese and French samples, respectively.

Functional annotation and biological insights

HaploReg V4.0 was used to assess functional annotation of the significant SNPs. The programme of functional enrichment by Trynka et al,¹⁵ was used for the enrichment analysis. We looked up the effects of SNPs on gene expression by previous expression quantitative trait loci (eQTL) data.²⁶²⁷

To get biological insight of SSc based on all SSc-associated loci, we also searched for all the SSc-associated loci: (1) missense mutations and functional annotation signals in SNPs in strong LD ($r^2 > 0.8$) with top SNPs in the loci, (2) the associations with the other diseases by GWAS catalogue using Gene names, (3) cis-eQTL signals based on the largest eQTL study,²⁶ with p values $< 1.0 \times 10^{-5}$, (4) H3K4me3 signals in CD4-naïve primary T cell with scores more than 0 calculated by the method mentioned above and (5) promoter histone marks in skin tissues based on results of HaploReg V4.0.

Statistical analysis

Logistic regression model was used for association studies. We used the first three principal components as covariates for the Japanese GWAS since additional PCs did not further improve the results. No inflation of p values was observed in French GWAS. We performed association studies using SSc, dcSSc, lcSSc, anti-Scl70 antibody(+)SSc and ACA(+)SSc as dependent variables. Inverse-variance method assuming fixed effects was applied to integrate the different association studies. Hardy-Weinberg equilibrium was assessed for the SNPs across the studies. SNPs showing association p values $< 5.0 \times 10^{-8}$ in the overall study were regarded as significant. Heterogeneity was evaluated for SNPs showing significant results using Cochran Q test. Interactive effects of two SNPs were evaluated by multiplicative model. Power calculation of the current study was conducted with use of 'Genetics Design' package of R software.

Since the first replication study in the European population was composed of French, Italian and German populations, we put covariates of the three populations as indicator variables. When we excluded imputation data to confirm the results by avoiding batch effects derived from different genotyping methods, German cases were combined with French cohorts because imputation data were used for all German controls.

Since individual imputation data for the French GWAS were not fully available due to restriction of data sharing policy in the control samples, we performed conditional analysis and HLA imputation using all of the case samples and a part of the control samples whose genotyping data were available.

Omnibus test to assess critical amino acid positions in the HLA region was conducted in each population and in the combined set as previously described.²⁸²⁹ When we analysed the combined set, the indicator variable of population was added as a covariate.

Statistical analysis was performed by PLINK or R statistical software. LocusZoom³⁰ was used to draw regional plots.

RESULTS

Japanese GWAS of SSc

We genotyped the Japanese cases and controls with five different Illumina Infinium arrays (see online supplementary table S1). After filtering samples based on quality control criteria, 700 cases and 1797 controls remained (see the Materials and methods section). To maximise the power to find new susceptibility loci, we performed imputation for this dataset with the East Asian panel in the 1000 Genome project²¹ as a reference. We identified rs12612769 in *STAT4* and rs9268636 near *HLA-DRA* showing significant associations ($p=4.7\times 10^{-8}$ and 9.6×10^{-10} , respectively, online supplementary figure S2A).

European GWAS for SSc and meta-analysis

Next, we used the previously published French GWAS containing 564 cases and 1776 controls and performed imputation with use of the European population panel in the 1000 Genomes Project European panel as reference (see online supplementary figure S2B). We conducted a transethnic meta-analysis of the two GWAS by the inverse-variance method assuming fixed effects for SNPs satisfying criteria of quality control (see online supplementary table S4). Since no evidence of population structure was obtained ($\lambda=1.05$, figure 1), we did not apply genomic control³¹ to correct statistics. As a result, we identified the *STAT4* region showing a significant association ($p=3.0\times 10^{-11}$, figure 1 and see online supplementary table S5). The *HLA* locus did not show a significant association ($p>1.3\times 10^{-7}$, figure 1) in spite of significant associations of the *HLA* locus in both populations (see online supplementary figure S2), suggesting different causative variants between the two populations. In fact, when we conducted HLA imputation using SNP2HLA, the different association patterns of amino acid positions were observed (see online supplementary table S6). The results of the susceptibility loci in the previous studies are shown in table 1. The risk alleles for all of the SNPs in the previous studies were the same in the meta-analysis or the French GWAS, suggesting replication of the previous findings and validity in the current study. In addition, we found an association in the *TNFAIP3* region whose association was previously reported without satisfying genome-wide significance level.³² All the variants in non-*HLA* region showing p value $<1.0\times 10^{-5}$ are shown in online supplementary table S5. We also performed SSc subtype GWAS according to the previous GWAS,⁶ namely, lcSSc, dcSSc, ACA(+)SSc and anti-Sc170(+)SSc (see online supplementary figure S3).

Selection of SNPs for the replication studies

We identified 33 SNPs in 33 novel candidates of susceptibility loci (see the Materials and methods section or see online supplementary figure S1). Twenty-seven out of the 33 SNPs were novel candidates of susceptibility loci to SSc. Among the remaining six SNPs, one and two were specific for limited and diffuse types, respectively, and two and one for possession of ACA and anti-Sc170 antibody only, respectively (see online supplementary table S7).

The two-staged replication studies

We recruited 564 cases and 1863 controls in the Japanese population and 1582 cases and 6694 controls in the European population for the first replication study (see online supplementary table S1 and S2). We found that rs3894194 in *GSDMA* showed an association beyond the significance level in the combined population. All the results for the 33 SNPs are shown in online supplementary table S7. We further recruited a total of 1010 cases and 2621 controls in the European population for the second replication study to validate the associations of the seven SNPs showing possible associations (see the Materials and methods or see online supplementary figure S1). As a result, rs3894194 in *GSDMA* kept its association (overall $p = 1.4 \times 10^{-10}$, table 2). rs4134466 in *PRDM1* in chromosome 6 also showed an association beyond the significance level (overall $p = 6.6 \times 10^{-10}$, table 2). The two SNPs did not display deviation from Hardy-Weinberg disequilibrium ($p > 0.037$) and heterogeneity ($p > 0.011$) across the studies. When we assessed the liability-scale variance explained by these two SNPs,³³ a total of 0.2% was explained in each population (see the Materials and methods section).

PRDM1 as a novel locus for SSc

rs4134466 is located 20 kbp downstream of *PRDM1*, also known as *BLIMP1*, encoding a transcription factor regulating T-cell proliferation and plasma cell differentiation.³⁴ The LD block spanning rs4134466 does not contain any other genes (figure 2A). The previous GWAS reported that this region was associated with other inflammatory conditions including RA,³⁵ systemic lupus erythematosus (SLE)³⁶ and inflammatory bowel disease (IBD).³⁷ When we searched for SNPs in the exonic region of *PRDM1* in strong LD with rs4134466, we could not find any coding variants in both Japanese and European populations. While *PRDM1* in chromosome 6 was adjacent to *ATG5*, a previously reported susceptibility gene to SSc,⁸ rs4134466 in *PRDM1* was not in strong LD with rs9373839 in *ATG5* showing the strongest susceptibility association in the previous study ($r^2 < 0.15$ in our study and the 1000 Genomes Project). In addition, rs9373839 was not polymorphic in the Japanese population. Thus, the association of rs4134466 was not driven by rs9373839. In fact, when we conditioned the association of rs4134466 on rs9373839 using imputation data of French GWAS, the effect size of rs4134466 risk allele did not change before and after conditioning (OR 1.102 and 1.105, before and after conditioning, respectively). Since the previous study of SLE GWAS³⁶ reported that rs65684331 in *PRDM1* is associated with SLE independently from rs2245214 in *ATG5*,³⁸ SSc seems to have multiple hits in this region as in SLE.

GSDMA as a novel locus for SSc

rs3894194 is a missense mutation of *GSDMA* altering an arginine residue to glutamine (p.R18Q). This amino acid residue is conserved across species with GERP score 3.34 (see online supplementary table S8). Estimation by PolyPhen-2³⁹ software suggest a benign effect of this variant. The LD block containing SNPs in LD with rs3894194 ($r^2 > 0.8$) harboured *LRRC3C* and this region is neighbouring *ORMDL3* and *GSDMB* (figure 2B). This region is a gene-rich region and reported to be associated with various immune-related diseases including RA¹⁶ and IBD.^{37,40} However, SNPs located in the LD block tagged by

rs3894194 have not been reported to be associated with other diseases. The RA-associated SNP (rs59716545) is in low LD with rs3894194 ($r^2=0.25$). *GSDMA* is associated with IBD in the previous study and the associated SNPs are in low LD with rs3894194 (rs2872507 or rs12946510, $r^2<0.38$). This region is also associated with asthma,⁴¹ but the effect of this SNP on asthma is opposite to that on IBD.⁴¹ This opposing effect seems to be true for asthma and SSc (OR of risk allele of this region: 1.26 and 1.18 in asthma and SSc, respectively).

Functional annotation of the two SNPs

Next, we assessed the effects of the two SNPs and the neighbouring SNPs on gene expression and functional annotation. We went through GTEx,²⁶ and found that rs3894194 in *GSDMA* showed a strong association with expression of *GSDMB* and *ORMDL3*, neighbouring genes to *GSDMA* ($p = 2.6 \times 10^{-12}$, figure 3A) and whose gene expression strongly correlated with each other. The association between gene expression and rs3894194 was also confirmed in the largest eQTL data²⁷ (see online supplementary table S9). We found that the associations between SSc and SNPs in the *GSDMA* locus correlated well with the associations of the SNPs with gene expressions of *GSDMB* and *ORMDL3* (figure 3B). Thus, the effect of the SNP on gene expression of *GSDMB* and *ORMDL3* in combination with amino acid alteration of the *GSDMA* protein seems to explain the association of this locus. HaploReg V4.0⁴² revealed that rs3894194 showed enhancer activity and enrichment of histone marks (see online supplementary table S10). While the previous eQTL studies²⁷ did not show associations between rs4134466 and gene expression, rs4134466 showed DNase hypersensitivity and methylation in various kinds of cells (see online supplementary table S10).

When we assessed interactive effects of the two SNPs on SSc susceptibility, we did not observe a significant effect ($p = 0.57$).

Subtype analyses for the two SNPs

When the associations of these two SNPs and the subtypes of SSc were analysed, rs3894194 in *GSDMA* showed a significant association with lcSSc (figure 3C). No other significant associations were observed (see online supplementary figure S4), but this study was underpowered to detect phenotype-specific associations. When we focused on SSc subtypes showing extreme phenotypes of fibrosis and vasculopathy (see the Materials and methods section), we did not find enhanced associations between the two SNPs and the subtypes (data not shown).

Enrichment analysis of histone modification

Next, based on the expanded list of susceptibility genes to SSc, we performed enrichment analysis of H3K4Me3, a representative histone modification mark that was shown to be enriched in autoimmune disease-related variants.¹⁵ We found that the susceptibility SNPs and the neighbouring SNPs in LD with them ($r^2>0.8$) showed suggestive enrichment of H3K4Me3 signal in CD4-naïve primary T cell or CD4 memory T cell (see online supplementary figure S5A). We also found that the suggestive enrichment signal in CD4-

naïve primary T cell was mainly brought about by the three SNPs in *GSDMA*, *PRDMI* and *TNFAIP3* found in the current study (see online supplementary figure S5B).

Functional annotation of susceptibility loci

The significant SSc-associated genes including the current results and *TNFAIP3* are summarised in online supplementary figure S6. We combined information of protein alteration, associations with other diseases and functional annotations. The development of promising drug targets by enrichment analysis based on the list may be challenging, but this table would be useful for candidates of future functional analyses and further expansion of SSc-associated loci.

DISCUSSION

This is the largest SSc GWAS from non-European populations and the first transethnic meta-analysis of SSc GWAS. We identified two novel susceptibility loci, namely *GSDMA* and *PRDMI*. Both loci were associated with other autoimmune diseases, consistent with overlapping susceptibility genes among various autoimmune diseases. We also replicated the associations of previously reported GWAS variants and provided evidence of association with *TNFAIP3*. To avoid possible batch effects due to different genotyping methods, we excluded all European subjects in the first replication study whose genotypes were imputed. The associations of the two SNPs remained significant ($p = 9.8 \times 10^{-9}$, data not shown).

We did not find a significant multiplicative interaction between the two SNPs. Since a previous study showed substantial interactive effects limited to the *HLA* loci,⁴³ it would be interesting to expand SSc cohorts and assess *HLA* interaction.

The enrichment analysis suggested possible involvement of CD4-naïve primary T cells with SSc. However, further expansion of susceptibility loci and convincing evidence of cell-type-specific enrichment are essential. We did not observe suggestive enrichment signal in CD19 primary cells, representing B cells. Interestingly, both SNPs showed evidence of associations of gene-expression including fibroblast or keratinocyte. Since previous loci were associated with gene expression especially in immune-related cells, the current findings would suggest importance of skin-residing cells on SSc pathophysiology. Cell-specific gene expression profile of fibroblast, keratinocyte or other fibrosis-related cell types including endothelial cells in combination with genetic data would be useful to address the importance and involvement of these cells and genes in SSc.

PRDM1, also known as B lymphocyte-induced maturation protein 1, is a transcript factor influencing a broad range of genes involved with cell proliferation and the immune system. PRDM1 is critical for epithelial and B cell differentiation,³⁴ and associated with other autoimmune diseases and haematopoietic malignancies. The association of this locus with SSc suggests a critical role of lymphocytes on SSc susceptibility. In fact, rs4134466 provided the highest score of H3K4me3 in CD4 (+)-naïve primary T cell among the SSc susceptibility variants. The first European replication study might suggest heterogeneity of this allele within the European population. Further expansion of subjects in subpopulations would clarify this point.

rs3894194 is a missense variant of *GSDMA* protein and associated with neighbouring gene expression. While it is not easy to pinpoint a causative variant, rs3894194 is a promising candidate of a causative SNP. *GSDMA* and *GSDMB* are strongly expressed in the skin and functional annotation revealed that rs3894194 has a regulatory effect of gene expression in various cell types including skin fibroblast. While rs3894194 also provided histone methylation in CD4(+)-naïve primary T cells, this locus may mainly demonstrate its susceptibility effect in the skin. The *GSDMA* locus showed a significant association with limited cutaneous SSc in spite of the reduced number of case subjects. This may suggest that this locus plays a more important role on developing lcSSc than dcSSc. However, since this locus also showed substantial associations with dcSSc, the results were inconclusive.

TNFAIP3 encodes A20 regulating tumour necrosis factor response by inhibiting nuclear factor- κ B (NF- κ B) activation. A20 also suppresses profibrotic signalling, relevant to SSc pathogenesis.⁴⁴ rs2230926 is a missense variant of *TNFAIP3* and associated with other rheumatic diseases.⁴⁵ The association of *TNFAIP3* as well as *TNIP1* supports NF- κ B involvement with SSc. However, we did not observe significant interactive effect of the two SNPs (data not shown).

Since the two populations substantially contributed to both the associations found in this study, the current findings indicate that transethnic meta-analysis is effective to identify unreported susceptibility loci to SSc, which comprise moderate effect sizes in each population. Furthermore, the current findings, especially rs4134466, would suggest that transethnic meta-analysis is effective by taking advantage of different allele frequencies and LD structure between the populations to discern unreported susceptibility signals from previously reported loci.

HLA and *STAT4* loci showed different association patterns between subtypes of SSc, suggesting genetic heterogeneity in SSc. While the association between *STAT4* and SSc was mainly driven by ACA(+) SSc, intracase analysis did not reveal significant difference in *STAT4* between ACA(+) and ACA(-) SSc ($p > 0.01$, data not shown). The *HLA* locus showed strong associations with antibody-positive SSc subtypes in spite of the reduced sample numbers even in intracase analyses. The associations of the *HLA* locus were attenuated in overall SSc, and this could be explained by different associations of the *HLA* locus between different SSc subtypes or different antibodies.⁴⁶ Our results also suggested different association patterns of the *HLA* locus between Japanese and European populations. It would be feasible to expand SSc to compare the genetic architectures between populations or subtypes.

The different arrays between cases and controls in the Japanese subjects reduced the number of preimputed and post-imputed markers. It would be feasible to rescan the control samples using the same arrays as the cases or take advantage of other controls which have used the same arrays to maximise power to find significant signals in future studies.

While it is still challenging to pinpoint a specific cell type contributing to SSc based on genetic findings, most of the susceptibility genes are immune-related and enrichment analysis suggested the importance of immune-related cells. Increasing samples for genetic

studies especially from non-European populations would increase SSc susceptibility loci, identify population-specific susceptibility loci, narrow down candidates of causative variants and clarify genetic architecture. Exome,⁴⁷ whole-genome or target deep sequencing might also be helpful. Clarification of genetic background of SSc by multiple approaches in combination with functional analyses would lead to the identification of possible therapeutic targets.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Chikashi Terao^{1,2,3,4,5}, Takahisa Kawaguchi¹, Philippe Dieude⁶, John Varga⁷, Masataka Kuwana⁸, Marie Hudson⁹, Yasushi Kawaguchi¹⁰, Marco Matucci-Cerinic¹¹, Koichiro Ohmura¹², Gabriela Riemekasten^{13,14}, Aya Kawasaki¹⁵, Paolo Airo¹⁶, Tetsuya Horita¹⁷, Akira Oka¹⁸, Eric Hachulla¹⁹, Hajime Yoshifuji¹², Paola Caramaschi²⁰, Nicolas Hunzelmann²¹, Murray Baron⁹, Tatsuya Atsumi¹⁷, Paul Hassoun²², Takeshi Torii²³, Meiko Takahashi¹, Yasuharu Tabara¹, Masakazu Shimizu¹, Akiko Tochimoto¹⁰, Naho Ayuzawa²⁴, Hidetoshi Yanagida²⁴, Hiroshi Furukawa^{15,25}, Shigeto Tohma²⁵, Minoru Hasegawa²⁶, Manabu Fujimoto²⁷, Osamu Ishikawa²⁸, Toshiyuki Yamamoto²⁹, Daisuke Goto³⁰, Yoshihide Asano³¹, Masatoshi Jinnin³², Hirahito Endo³³, Hiroki Takahashi³⁴, Kazuhiko Takehara³⁵, Shinichi Sato³¹, Hironobu Ihn³², Soumya Raychaudhuri^{3,4,5,36}, Katherine Liao³, Peter Gregersen³⁷, Naoyuki Tsuchiya¹⁵, Valeria Riccieri³⁸, Inga Melchers³⁹, Gabriele Valentini⁴⁰, Anne Cauvet⁴¹, Maria Martinez⁴², Tsuneyo Mimori¹², Fumihiko Matsuda¹, Yannick Allanore⁴³

Affiliations

¹Department of Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan

²Center for the Promotion of Interdisciplinary Education and Research, Kyoto University Graduate School of Medicine, Kyoto, Japan

³Division of Rheumatology, Immunology, and Allergy, Brigham and Women's Hospital, Boston, Massachusetts, USA

⁴Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA

⁵Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA

⁶Rheumatology Bichat Hospital, Paris, University, Paris, France

⁷Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

⁸Division of Rheumatology, Department of Internal Medicine, Keio University School of Medicine, Tokyo, Japan

- ⁹Jewish General Hospital and Lady Davis Research Institute, Montreal, Quebec, Canada
- ¹⁰Institute of Rheumatology, Tokyo Women's Medical University, Tokyo, Japan
- ¹¹Division of Rheumatology AOUC, Department of Experimental and Clinical Medicine, Department of Medical & Geriatrics Medicine, University of Florence, Firenze, Italy
- ¹²Department of Rheumatology and Clinical Immunology, Kyoto University Graduate School of Medicine, Kyoto, Japan
- ¹³Clinic for Rheumatology, University of Lübeck, Lübeck, Germany
- ¹⁴German Lung Center Borstel, Leibniz Institute, Germany
- ¹⁵Molecular and Genetic Epidemiology Laboratory, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan
- ¹⁶Rheumatology Unit, Spedali Civili, Brescia, Italy
- ¹⁷Division of Rheumatology, Endocrinology and Nephrology, Hokkaido University, Graduate School of Medicine, Sapporo, Japan
- ¹⁸The Institute of Medical Science, Tokai University, Isehara, Japan
- ¹⁹Internal Medicine Department, FHU Immune-Mediated Inflammatory Diseases and Targeted Therapies, Lille University, Lille, France
- ²⁰Rheumatology Department, University of Verona, Azienda Ospedaliera Universitaria Integrata, Italy
- ²¹Dermatology Department, University of Köln, Köln, Germany
- ²²Division of Pulmonary and Critical Care Medicine, Department of Medicine, Johns Hopkins University, Baltimore, Maryland, USA
- ²³Torii Clinic, Kyoto, Japan
- ²⁴Department of Rheumatology, National Hospital Organization, Utano National Hospital, Kyoto, Japan
- ²⁵Clinical Research Center for Allergy and Rheumatology, Sagamihara Hospital, National Hospital Organization, Sagamihara, Japan
- ²⁶Division of Medicine, Faculty of Medical Sciences, Department of Dermatology, University of Fukui, Fukui, Japan
- ²⁷Department of Dermatology, Faculty of Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan
- ²⁸Department of Dermatology, Gunma University Graduate School of Medicine, Gunma, Japan
- ²⁹Department of Dermatology, Fukushima Medical University, Fukushima, Japan

- ³⁰Department of Internal Medicine, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan
- ³¹Department of Dermatology, University of Tokyo Graduate School of Medicine, Tokyo, Japan
- ³²Department of Dermatology and Plastic Surgery, Faculty of Life Sciences, Kumamoto University, Kumamoto, Japan
- ³³Division of Rheumatology, Department of Internal Medicine, School of Medicine, Toho University, Tokyo, Japan
- ³⁴Department of Rheumatology and Clinical Immunology, Sapporo Medical University School of Medicine, Sapporo, Hokkaido, Japan
- ³⁵Department of Dermatology, Faculty of Medicine, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Kanazawa, Ishikawa, Japan
- ³⁶Arthritis Research UK Centre for Genetics and Genomics, Centre for Musculoskeletal Research, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK
- ³⁷Robert S. Boas Center for Genomics and Human Genetics, The Feinstein Institute for Medical Research, Manhasset, New York, USA
- ³⁸Sapienza University of Rome, Rome, Italy
- ³⁹University Medical Center, Freiburg, Germany
- ⁴⁰Department of Clinical and Experimental Medicine, Rheumatology Section, Second University of Naples, Naples, Italy
- ⁴¹INSERM U1016/UMR 8104, Cochin Institute, Paris Descartes University, Paris, France
- ⁴²INSERM U1220—IRSD—Batiment B Purpan Hospital Toulouse, Paris, France
- ⁴³Rheumatology A Department, INSERM U1016/UMR 8104, Cochin Institute, Paris Descartes University, Paris, France

Acknowledgements

We thank the investigators of the French Three-City (3C) cohort and in particular, Drs Philippe Amouyel, Christophe Lambert (Lille, France) and Luc Letenneur (Bordeaux), who gave us access to data from controls. We also thank Drs Monique Hinchcliffe and Benjamin Korman for sample collection and Mr Takeshi Iino for performing the replication in the Japanese population.

Funding This study was supported by JSPS KAKENHI Grant Number JP16H06251, KANAE foundation for the promotion of medical science, Research Project of Genetic Studies for Intractable Diseases, Nagao Memorial Fund, The Uehara Memorial Foundation, The John Mung Advanced Program, Kyoto University and Association des Sclerodermie de France, INSERM, CNRS, ATIP AVENIR Programme, Agence Nationale pour la Recherche (Project ANR-O8-GENO-OI6-I).

REFERENCES

1. Elhai M, Avouac J, Kahan A, et al. Systemic sclerosis: recent insights. *Joint Bone Spine* 2015;82:148–53. [PubMed: 25557659]
2. Terao C, Ohmura K, Kawaguchi Y, et al. PLD4 as a novel susceptibility gene for systemic sclerosis in a Japanese population. *Arthritis Rheum* 2013;65:472–80. [PubMed: 23124809]
3. Bossini-Castillo L, López-Isac E, Martin J. Immunogenetics of systemic sclerosis: defining heritability, functional variants and shared-autoimmunity pathways. *J Autoimmun* 2015;64:53–65. [PubMed: 26212856]
4. Radstake TR, Gorlova O, Rueda B, et al. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. *Nat Genet* 2010;42:426–9. [PubMed: 20383147]
5. Allanore Y, Saad M, Dieude P, et al. Genome-wide scan identifies TNIP1, PSORS1C1, and RHOB as novel risk loci for systemic sclerosis. *PLoS Genet* 2011;7: e1002091.
6. Gorlova O, Martin JE, Rueda B, et al. Identification of novel genetic markers associated with clinical phenotypes of systemic sclerosis through a genome-wide association strategy. *PLoS Genet* 2011;7:e1002178.
7. Martin JE, Broen JC, Carmona FD, et al. Identification of CSK as a systemic sclerosis genetic risk factor through Genome Wide Association Study follow-up. *Hum Mol Genet* 2012;21:2825–35. [PubMed: 22407130]
8. Mayes MD, Bossini-Castillo L, Gorlova O, et al. Immunochip analysis identifies multiple susceptibility loci for systemic sclerosis. *Am J Hum Genet* 2014;94:47–61. [PubMed: 24387989]
9. Bossini-Castillo L, Martin JE, Broen J, et al. A GWAS follow-up study reveals the association of the IL12RB2 gene with systemic sclerosis in Caucasian populations. *Hum Mol Genet* 2012;21:926–33. [PubMed: 22076442]
10. López-Isac E, Campillo-Davo D, Bossini-Castillo L, et al. Influence of TYK2 in systemic sclerosis susceptibility: a new locus in the IL-12 pathway. *Ann Rheum Dis* 2016;75:1521–6. [PubMed: 26338038]
11. López-Isac E, Bossini-Castillo L, Guerra SG, et al. Identification of IL12RB1 as a novel systemic sclerosis susceptibility locus. *Ann Rheum Dis* 2014;66:3521–3.
12. Terao C Genetic contribution to susceptibility and disease phenotype in rheumatoid arthritis. *Inflamm Regen* 2014;34:71–7.
13. Terao C, Yoshifuji H, Kimura A, et al. Two susceptibility loci to Takayasu arteritis reveal a synergistic role of the IL12B and HLA-B regions in a Japanese population. *Am J Hum Genet* 2013;93:289–97. [PubMed: 23830516]
14. Terao C, Yoshifuji H, Nakajima T, et al. Ustekinumab as a therapeutic option for Takayasu arteritis: from genetic findings to clinical application. *Scand J Rheumatol* 2015:1–3. [PubMed: 26303149]
15. Trynka G, Sandor C, Han B, et al. Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nat Genet* 2013;45:124–30. [PubMed: 23263488]
16. Okada Y, Wu D, Trynka G, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2014;506:376–81. [PubMed: 24390342]
17. Zhou X, Lee JE, Arnett FC, 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;60:3807–14. [PubMed: 19950302]
18. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. *Arthritis Rheum* 1980;23:581–90. [PubMed: 7378088]
19. LeRoy EC, Black C, Fleischmajer R, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *J Rheumatol* 1988;15:202–5. [PubMed: 3361530]
20. Li Y, Willer CJ, Ding J, et al. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 2010;34:816–34. [PubMed: 21058334]
21. Abecasis GR, Altshuler D, Auton A, et al., 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* 2010;467:1061–73. [PubMed: 20981092]

22. Barrett JC, Fry B, Maller J, et al. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–5. [PubMed: 15297300]
23. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–75. [PubMed: 17701901]
24. Raychaudhuri S, Plenge RM, Rossin EJ, et al. Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet* 2009;5:e1000534.
25. Jia X, Han B, Onengut-Gumuscu S, et al. Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS ONE* 2013;8:e64683.
26. Consortium GTEx. The genotype-tissue expression (GTEx) project. *Nat Genet* 2013;45:580–5. [PubMed: 23715323]
27. Westra HJ, Peters MJ, Esko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013;45:1238–43. [PubMed: 24013639]
28. Raychaudhuri S, Sandor C, Stahl EA, et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat Genet* 2012;44:291–6. [PubMed: 22286218]
29. Terao C, Yano K, Ikari K, et al. Brief report: main contribution of DRB1*04:05 among the shared epitope alleles and involvement of DRB1 amino acid position 57 in association with joint destruction in anti-citrullinated protein antibody-positive rheumatoid arthritis. *Arthritis Rheumatol* 2015;67:1744–50. [PubMed: 25777156]
30. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26:2336–7. [PubMed: 20634204]
31. Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999;55:997–1004. [PubMed: 11315092]
32. Dieude P, Guedj M, Wipff J, et al. Association of the TNFAIP3 rs5029939 variant with systemic sclerosis in the European Caucasian population. *Ann Rheum Dis* 2010;69:1958–64. [PubMed: 20511617]
33. Stahl EA, Wegmann D, Trynka G, et al. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat Genet* 2012;44:483–9. [PubMed: 22446960]
34. Ott G, Rosenwald A, Campo E. Understanding MYC-driven aggressive B-cell lymphomas: pathogenesis and classification. *Blood* 2013;122:3884–91. [PubMed: 24009228]
35. Raychaudhuri S, Thomson BP, Remmers EF, et al. Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. *Nat Genet* 2009;41:1313–18. [PubMed: 19898481]
36. Gateva V, Sandling JK, Hom G, et al. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. *Nat Genet* 2009;41:1228–33. [PubMed: 19838195]
37. Anderson CA, Boucher G, Lees CW, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 2011;43:246–52. [PubMed: 21297633]
38. Harley JB, Alarcón-Riquelme ME, Criswell LA, et al. International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEN). Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXX, KIAA1542 and other loci. *Nat Genet* 2008;40:204–10. [PubMed: 18204446]
39. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9. [PubMed: 20354512]
40. Barrett JC, Hansoul S, Nicolae DL, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn’s disease. *Nat Genet* 2008;40:955–62. [PubMed: 18587394]
41. Li X, Ampleford EJ, Howard TD, et al. Genome-wide association studies of asthma indicate opposite immunopathogenesis direction from autoimmune diseases. *J Allergy Clin Immunol* 2012;130:861–8.e7. [PubMed: 22694930]
42. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012;40:D930–934. [PubMed: 22064851]

43. Lenz TL, Deutsch AJ, Han B, et al. Widespread non-additive and interaction effects within HLA loci modulate the risk of autoimmune diseases. *Nat Genet* 2015;47:1085–90. [PubMed: 26258845]
44. Bhattacharyya S, Wang W, Graham LV, et al. A20 suppresses canonical Smad-dependent fibroblast activation: novel function for an endogenous inflammatory modulator. *Arthritis Res Ther* 2016;18:216. [PubMed: 27716397]
45. Shimane K, Kochi Y, Horita T, et al. The association of a nonsynonymous single-nucleotide polymorphism in TNFAIP3 with systemic lupus erythematosus and rheumatoid arthritis in the Japanese population. *Arthritis Rheum* 2010;62:574–9. [PubMed: 20112363]
46. Beretta L, Rueda B, Marchini M, et al. Analysis of class II human leucocyte antigens in Italian and Spanish systemic sclerosis. *Rheumatology (Oxford)* 2012;51:52–9. [PubMed: 22087014]
47. Gao L, Emond MJ, Louie T, et al. Identification of rare variants in ATP8B4 as a risk factor for systemic sclerosis by whole-exome sequencing. *Rheumatology (Oxford)* 2016;68:191–200.

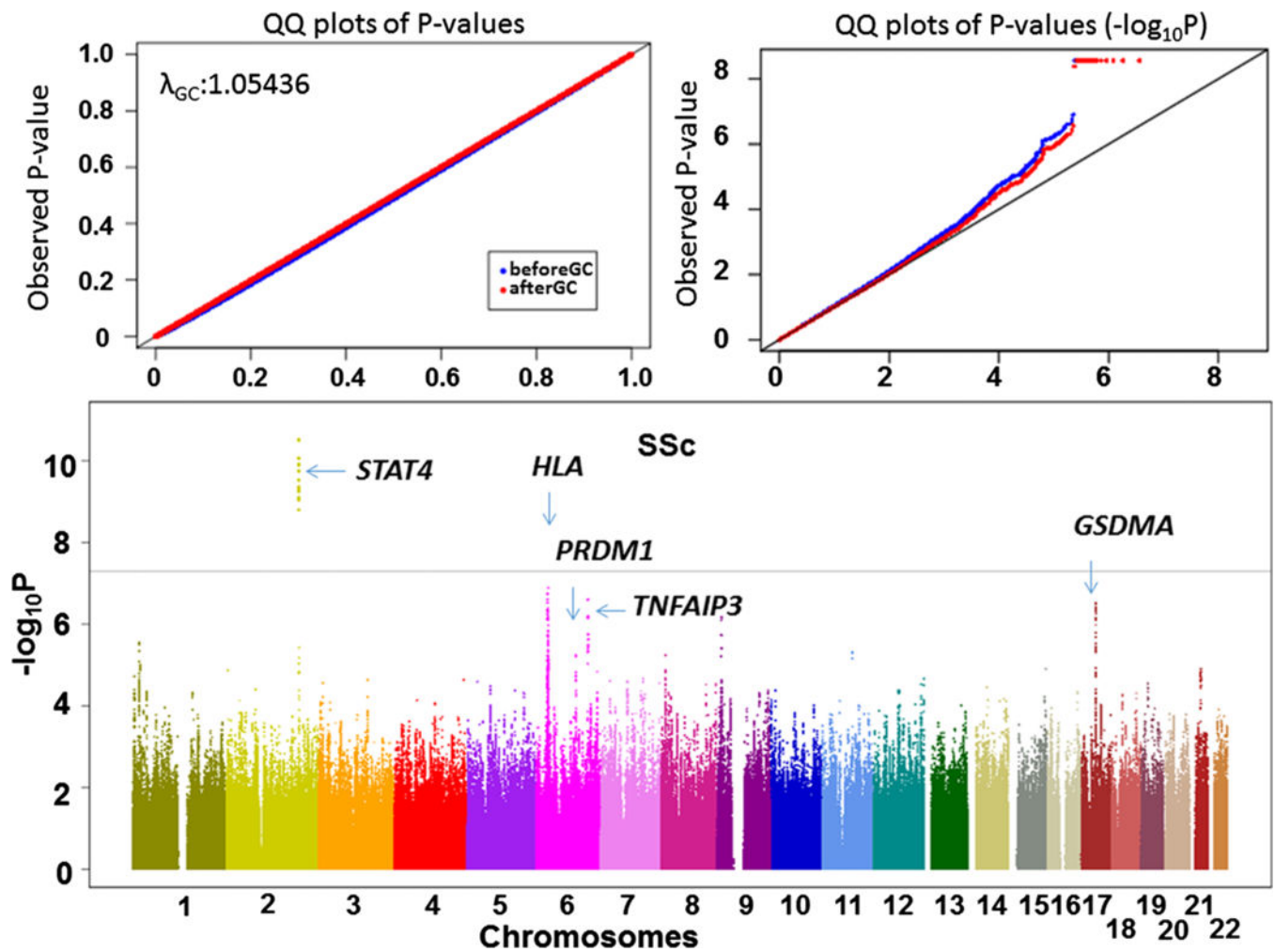


Figure 1.

Transethnic meta-analysis of genome-wide association studies (GWAS) revealed multiple susceptibility loci to systemic sclerosis (SSc). The results of the transethnic meta-analysis of GWAS are shown in the Manhattan plot and quantile-quantile (QQ) plot. The newly identified loci and previously reported loci with strong p values are indicated in the Manhattan plot. The horizontal line indicates the genome-wide significance level.

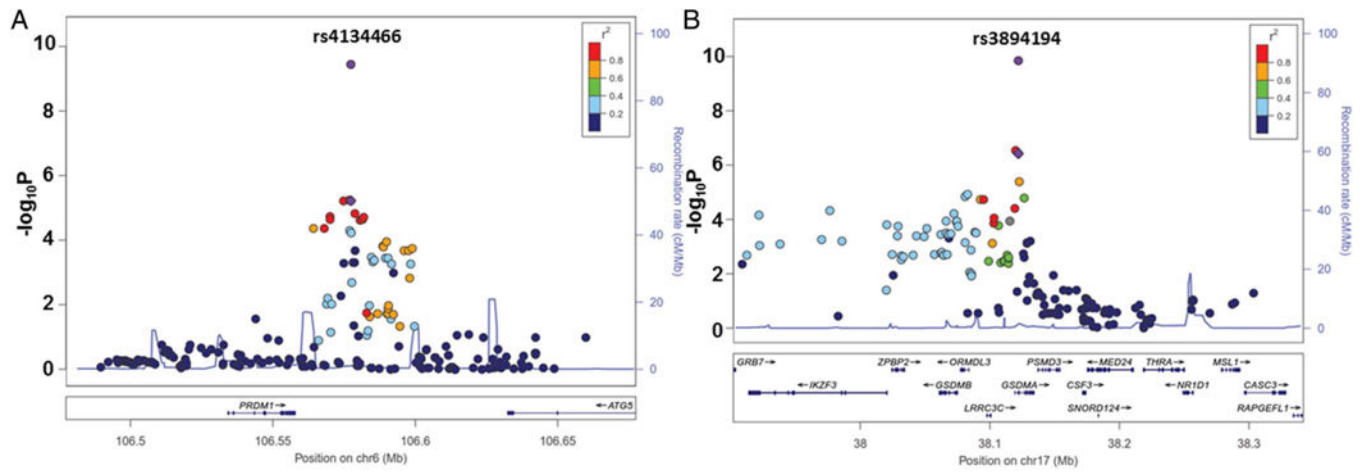


Figure 2.

Detailed plot for the two loci found in the current study. The detailed plots in chromosome 6 and 17 are shown for (A and B), respectively. The purple plots indicate the top SNPs in the combined results and GWAS meta-analysis for the upper and lower plots, respectively. The plots are drawn based on the linkage disequilibrium (LD) structure of East Asians by using LocusZoom as a representative.

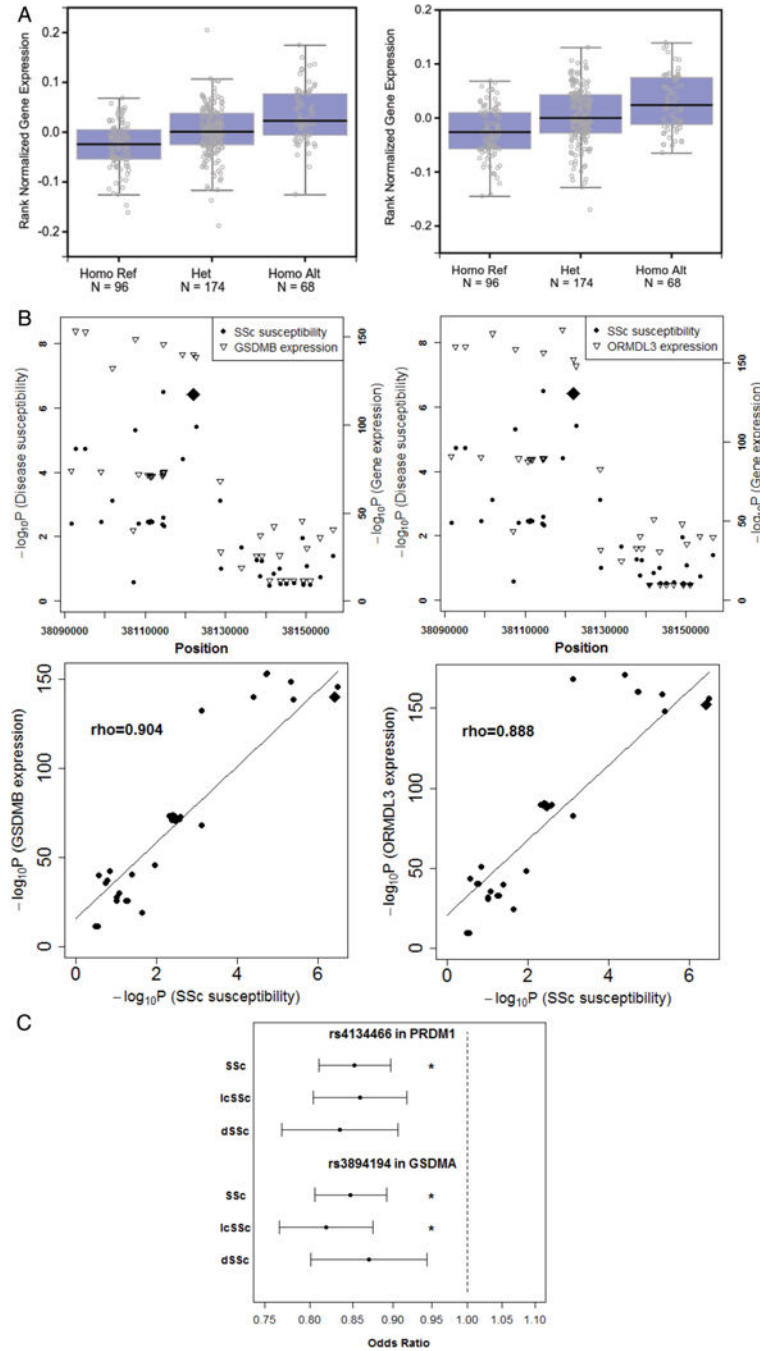


Figure 3. Correlation between the associations of variants in GSDMA region with systemic sclerosis (SSc) susceptibility and gene expression. (A) rs4134466 is associated with gene expression of *GSDMA*-neighbouring genes *GSDMB* (left) and *ORMDL3* (right). The box plots were obtained from GTEx data. (B) The associations between SSc susceptibility and single nuclear polymorphisms (SNPs) in chromosome 17 *GSDMA* locus are plotted together with the associations between the variants and expression of *GSDMB* (left) and *ORMDL3* (right). The gene expression data were obtained from Blood eQTL Browser. The correlation

plots are indicated in the lower panels. The black diamonds indicate rs4134466. (C) rs4134466 is associated with limited SSc. The associations between the two SNPs and the two subtypes of SSc are indicated. lcSSc, limited cutaneous SSc; dcSSc, diffuse cutaneous SSc.

Table 1

The results in the current study for the previous GWAS loci and *TNFAIP3*

Previously reported loci				Current GWAS meta-analysis				Japanese GWAS			French GWAS				
SNP	Chr	BP	Neighbouring gene	Risk SNP	β	SE	p Value	Risk SNP	β	SE	p Value	Risk SNP	β	SE	p Value
rs3790567	1	67822377	<i>IL12RB2</i>	A	0.112	0.057	0.050	A	0.116	0.085	0.17	A	0.109	0.077	0.16
rs2056626	1	167420425	<i>CD247</i>	T	0.101	0.059	0.086	T	-0.048	0.105	0.65	T	0.170	0.071	0.017
rs7574865	2	191964633	<i>STAT4</i>	T	0.332	0.054	5.3×10^{-10}	T	0.375	0.074	3.6×10^{-7}	T	0.285	0.078	0.00025
rs35677470	3	58183636	<i>DNASE1L3</i>	A	-	-	-	A	-	-	-	A	0.654	0.130	9.4×10^{-7}
rs77583790	3	159694053	<i>SCHIP1-IL12A</i>	A	-	-	-	A	-	-	-	A	0.345	0.285	0.24
rs2233287	5	150440097	<i>TNIP1</i>	A	-	-	-	A	-	-	-	A	0.440	0.107	3.7×10^{-5}
rs9373839	6	106655617	<i>ATG5</i>	C	-	-	-	C	-	-	-	C	0.246	0.083	0.0034
rs2230926	6	138196066	<i>TNFAIP3</i>	G	0.525	0.102	2.5×10^{-7}	G	0.352	0.130	0.0066	G	0.801	0.164	1.9×10^{-6}
rs10488631	7	128594183	<i>IRF5/TNPO3</i>	C	-	-	-	C	-	-	-	C	0.328	0.108	0.0024
rs11642873	16	85991705	<i>IRF8</i>	A	0.235	0.077	0.0024	A	0.195	0.132	0.14	A	0.256	0.096	0.0072
rs2304256	19	10475652	<i>TYK2</i>	C	0.103	0.054	0.058	C	0.079	0.074	0.29	C	0.132	0.080	0.095
rs2305743	19	18193191	<i>IL12RB1</i>	G	0.104	0.062	0.094	G	0.040	0.087	0.64	G	0.171	0.089	0.055
rs137894	22	50467524	<i>C5K</i>	C	-	-	-	C	-	-	-	C	-	-	-

Data not available due to lack of satisfying quality control criteria.

BP, base position; Chr, chromosome; GWAS, genome-wide association studies; SNP, single nuclear polymorphisms.

Table 2

The results of the seven SNPs selected for the second replication study

SNP	Ch	BP	Gene	AI/ A2	Pop	GWAS meta-analysis			Replication 1			Replication 2			Overall						
						A2Case	A2Cont	p Value	A2Case	A2Cont	p Value	A2Case	A2Cont	p Value	β	SE	p Value	β	SE	p Value	
rs10907300	1	18400980	<i>ICSEF21</i>	C/ A	Japanese	0.61	0.55	0.00025	0.60	0.57	0.038	0.60	0.57	0.038	0.088	0.026	0.00076	0.088	0.026	0.00076	1.09 (1.04 to 1.15)
rs6714060	2	74210442	<i>TET3</i>	C/ T	European Japanese	0.39 0.15	0.34 0.19	0.0031 0.0031	0.38 0.15	0.35 0.20	0.018 0.00041	0.37 0.40	0.37 0.40	0.012	-0.134	0.037	0.00024	-0.134	0.037	0.00024	0.87 (0.81 to 0.94)
rs4134466	6	106577368	<i>PRDM1</i>	A/ G	European Japanese	0.13 0.59	0.15 0.66	0.016 0.0058	0.13 0.60	0.14 0.66	0.34 0.00020	0.15 0.14	0.14 0.74	0.74	-0.160	0.026	6.6×10 ⁻¹⁰	-0.160	0.026	6.6×10 ⁻¹⁰	0.85 (0.81 to 0.90)
rs12676482	8	42174077	<i>IKBKB</i>	G/ A	European Japanese	0.57 0.20	0.63 0.15	0.00030 0.0013	0.62 0.19	0.62 0.16	0.35 0.013	0.59 0.64	0.59 7.2×10 ⁻⁵	0.30	0.221	0.051	1.4×10 ⁻⁵	0.221	0.051	1.4×10 ⁻⁵	1.25 (1.13 to 1.38)
rs2821195	9	11689891	<i>no gene</i>	C/ G	European Japanese	0.04 0.49	0.03 0.55	0.049 0.00047	0.03 0.50	0.03 0.51	0.44 0.63	0.05 0.04	0.05 0.30	0.30	-0.093	0.027	0.00048	-0.093	0.027	0.00048	0.91 (0.87 to 0.96)
rs12357548	10	63803472	<i>ARID5B</i>	G/ A	European Japanese	0.31 0.30	0.36 0.35	0.00038 0.00070	0.32 0.34	0.36 0.35	0.0062 0.54	0.37 0.35	0.37 0.080	0.080	-0.073	0.026	0.00053	-0.073	0.026	0.00053	0.93 (0.88 to 0.98)
rs3894194	17	38121993	<i>GSDMA</i>	G/ A	European Japanese	0.51 0.50	0.55 0.54	0.044 0.012	0.49 0.51	0.53 0.53	0.00072 0.28	0.51 0.48	0.51 0.0065	0.0065	-0.166	0.026	1.4×10 ⁻¹⁰	-0.166	0.026	1.4×10 ⁻¹⁰	0.85 (0.80 to 0.89)

β and SE are values for A2 allele.

A2Case, frequency of A2 allele in case; A2Cont, frequency of A2 allele in control; BP, base position; Ch, chromosome; Pop, population; GWAS, genome-wide association studies; SNP, single nuclear polymorphisms.