











Unraveling the Genetics of Shared Clinical and Serological Manifestations in Patients With Systemic Inflammatory Autoimmune Diseases

Matteo Bianchi,¹  Sergey V. Kozyrev,¹  Antonella Notarnicola,²  Johanna K. Sandling,¹ Mats Pettersson,¹ Dag Leonard,¹ Christopher Sjöwall,³  Iva Gunnarsson,⁴ Solbritt Rantapää-Dahlqvist,⁵  Anders A. Bengtsson,⁶ Andreas Jönsen,⁶ Elisabet Svenungsson,⁴  Helena Enocsson,³  Marika Kvarnström,⁴ Helena Forsblad-d'Elia,⁷ Sara Magnusson Bucher,⁸ Katrine B. Norheim,⁹ Eva Baecklund,¹ Roland Jonsson,¹⁰ Daniel Hammenfors,¹¹ Per Eriksson,³ Thomas Mandl,¹² Roald Omdal,¹³ Leonid Padyukov,⁴ Helena Andersson,¹⁴ Øyvind Molberg,¹⁴ Louise Pyndt Diederichsen,¹⁵ Ann-Christine Syvänen,¹ Marie Wahren-Herlenius,¹⁶  Gunnel Nordmark,¹  Ingrid E. Lundberg,⁴  Lars Rönnblom,¹ and Kerstin Lindblad-Toh,¹⁷ with the DISSECT consortium and the ImmunoArray consortium

Objective. Systemic inflammatory autoimmune diseases (SIADs) such as systemic lupus erythematosus (SLE), primary Sjögren disease (pSS), and idiopathic inflammatory myopathies (myositis) are complex conditions characterized by shared circulating autoantibodies and clinical manifestations, including skin rashes, among others. This study was aimed at elucidating the genetics underlying these common features.

Methods. We performed targeted DNA sequencing of coding and regulatory regions from approximately 1,900 immune-related genes in a large cohort of 2,292 well-characterized Scandinavian patients with SIADs with SLE, pSS, and myositis as well as 1,252 controls. A gene-based functionally weighted genetic score for aggregate testing of all genetic variants, including rare variants, was complemented by in silico functional analyses and in vitro reporter experiments.

Results. Case–control association analysis detected known and potentially novel genetic loci in agreement with previous genetic and transcriptomics findings linked to the SIAD autoimmune background. Intriguingly, case–case comparisons between patient subgroups with and without specific autoantibodies revealed that the subgroups defined by antinuclear antibodies and anti–double-stranded DNA antibodies have unique genetic profiles reflecting their heterogeneity. When focusing on clinical features, we overall showed that dual-specificity phosphatase 1 (*DUSP1*) protective genetic variants lead to increased gene expression and potentially to anti-inflammatory effects on the SIAD-associated skin phenotype. This is consistent with recent genetic findings on eczema and with the previously reported down-regulation of the MAPK signaling-related gene *DUSP1* in other skin disorders.

Conclusion. Together, this suggests common molecular mechanisms potentially underlying overlapping clinical manifestations shared among different disorders and informs clinical heterogeneity, which could be translated to improve disease diagnostic and treatment, also in more generalized disease frameworks.

INTRODUCTION

Systemic inflammatory autoimmune diseases (SIADs) such as systemic lupus erythematosus (SLE), primary Sjögren

disease (pSS), and idiopathic inflammatory myopathies (myositis) are chronic heterogeneous rare conditions in which inflammation and the primary pathogenic events can be restricted to different organ systems and co-occur with specific

Supported by the Swedish Research Council for Medicine and Health (grants Dnr 2018-02399, 2018-02535, and 2016-01254), the Swedish Rheumatism Association, King Gustav V's 80-year Foundation, the Swedish Society of Medicine, the Ingegerd Johansson Donation, the Swedish Heart Lung Foundation, The Swedish Cancer Society, the Knut and Alice Wallenberg Foundation, the Norwegian Research Council, and the Torsten Söderberg Foundation. ALF funding was received from Stockholm County, KID funding was received from the Karolinska Institutet, and a

Wallenberg Scholar Award as well as a Distinguished Professor award were given to Dr Lindblad-Toh.

¹Matteo Bianchi, PhD, Sergey V. Kozyrev, PhD, Johanna K. Sandling, PhD, Mats Pettersson, PhD, Dag Leonard, MD, PhD, Eva Baecklund, MD, PhD, Ann-Christine Syvänen, PhD, Gunnel Nordmark, MD, PhD, Lars Rönnblom, MD, PhD: Uppsala University, Uppsala, Sweden; ²Antonella Notarnicola, MD, PhD: Karolinska University Hospital, Stockholm, Sweden; ³Christopher Sjöwall, MD, PhD, Helena Enocsson, PhD, Per Eriksson, MD, PhD: Linköping

serological profiles.^{1–3} Patients with SIADs show remarkable similarities, such as overexpression of type I interferon (IFN)–regulated genes (ie, IFN signature), identical circulating autoantibodies mainly targeting nuclear antigens (eg, antinuclear antibodies [ANA], SSA-Ro, SSB-La), and analogous clinical manifestations and comorbidities, including skin rashes, arthritis, different types of malignancies, and renal and lung involvement, which develop regardless of the primary disease diagnosis.⁴ In addition, a number of shared genetic loci encompassing genes involved in innate and adaptive immune responses have been identified as disease risk factors, thus overall highlighting a partially shared genetic etiology in patients with SIADs.^{5–8}

The genetic background of patients with SIADs has been primarily explored via genome-wide association and targeted genotyping studies, as well as meta-analyses, unraveling disease associations mainly driven by common genetic variation.^{5,6,8} Conversely, the contribution of the full spectrum of genetic variation has not been comprehensively evaluated in patients with SIADs, especially when considering their shared serological and clinical features. For this purpose, the detection of those genetic variants rarer in the population is required, as well as large and inclusive datasets in which different patient cohorts are combined.⁹ This has to be ideally complemented by the implementation of algorithms testing the cumulative effect of different genetic variants, especially of those in the lower range of allele frequency in sizable genetic regions.¹⁰ In fact, it has been shown that multiple causal variants underlie genetic associations in complex diseases¹¹ and that rare variants also contribute to disease heritability.¹² Using next-generation targeted DNA sequencing of regulatory and coding regions in a large cross-disease sample cohort including Scandinavian patients with SLE, pSS, and myositis as well as healthy controls, the objective of the present study was to examine the contribution of the whole spectrum of genetic variants in defining the genetic background underlying these patients with SIADs and potentially also their secondary shared serological and clinical manifestations, thus allowing the

identification of genetic profiles conceivably relevant for patient stratification in the clinic.

PATIENTS AND METHODS

For detailed materials and methods description, see Supplementary Materials and Methods. Members of the DISSECT and ImmunoArray consortia can be found in Appendixes A and B, respectively.

Study participants. The study cohort consisted of healthy controls as well as patients diagnosed with SLE, pSS, and myositis, from whom blood specimens and detailed clinical and serological information were collected at Scandinavian rheumatology clinics (ie, Sweden, Denmark, and Norway) in compliance with individual consent and regional ethical permits. Patient characteristics, as well as their clinical and serological profiles, have been previously described in details.^{13–15} Serological and clinical patient subgroups were defined based on the occurrence of the same circulating autoantibodies in the former and to inclusively combine clinical manifestations and comorbidities into collective features related to the same organ/tissue or comprising the same diagnosis in the latter (Table 1).

Targeted DNA sequencing and bioinformatics analysis. A comprehensive description of the capture array design and implementation, as well as the targeted sequencing of the samples included in this study, has been previously provided.^{13–16} Briefly, the targeted array included 1,853 genes involved in immune function, inflammation, and autoimmunity for which both coding and potentially regulatory regions were captured. After mapping the sequencing reads to the human GRCh37 reference genome and initial data quality control (QC), variant discovery, and joint genotyping of the whole study cohort of patients and healthy controls was performed under the GATK Best Practices framework (GATK version 3.3.0), followed by variant and individual QC, including relatedness, population stratification, and ancestry assessment.

University, Linköping, Sweden; ⁴Iva Gunnarsson, MD, PhD, Elisabet Svenungsson, MD, PhD, Marika Kvarnström, MD, PhD, Leonid Padyukov, MD, PhD, Ingrid E. Lundberg, MD, PhD: Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden; ⁵Solbritt Rantapää-Dahlqvist, MD, PhD: Umeå University, Umeå, Sweden; ⁶Anders A. Bengtsson, MD, PhD, Andreas Jönsen, MD, PhD: Lund University and Skåne University Hospital, Lund, Sweden; ⁷Helena Forsblad-d'Elia, MD, PhD: University of Gothenburg, Gothenburg, Sweden; ⁸Sara Magnusson Bucher, MD: Örebro University, Örebro, Sweden; ⁹Katrine B. Norheim, MD, PhD: Stavanger University Hospital, Stavanger, Norway; ¹⁰Roland Jonsson, DMD, PhD: University of Bergen, Bergen, Norway; ¹¹Daniel Hammenfors, MD: Haukeland University Hospital, Bergen, Norway; ¹²Thomas Mandl, MD, PhD: Lund University, Sweden; ¹³Roald Omdal, MD, PhD: Stavanger University Hospital, Stavanger, Norway, and University of Bergen, Bergen, Norway; ¹⁴Helena Andersson, MD, PhD, Øyvind Molberg, MD, PhD: Oslo University Hospital, Oslo, Norway; ¹⁵Louise Pyndt Diederichsen, MD, PhD: Odense University Hospital, Odense, Denmark, and

Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; ¹⁶Marie Wahren-Herlenius, MD, PhD: Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden, and Broegelmann Research Laboratory, University of Bergen, Bergen, Norway; ¹⁷Kerstin Lindblad-Toh, PhD: Uppsala University, Uppsala, Sweden, and Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts.

Additional Supplementary information cited in this article can be found online in the Supporting Information section (<http://onlinelibrary.wiley.com/doi/10.1002/art.42988>).

Author disclosures are available at <https://onlinelibrary.wiley.com/doi/10.1002/art.42988>.

Address correspondence via email to Matteo Bianchi, PhD, at matteo.bianchi@imbim.uu.se; or to Kerstin Lindblad-Toh, PhD, at kersli@broadinstitute.org.

Submitted for publication January 13, 2024; accepted in revised form September 3, 2024.

Single-marker and gene-based aggregate association analyses. Logistic regression was implemented in PLINK version 1.9.0 to perform single-marker case-control analysis of common single-nucleotide variants (SNVs; minor allele frequency [MAF] ≥ 0.01) comparing all patients with SIADs with all control individuals. Gene-based aggregate association testing of all SNVs, including rare variants, was performed using a method that was inspired by the simplest type of genetic burden test, which sums up all detected mutations in each individual.¹⁷ In our study, we assigned a genetic score (GS) to each individual and gene. For every SNV, the corresponding Combined Annotation Dependent Depletion (CADD) score¹⁸ was multiplied by the number of variant (ie, alternate) alleles in each individual. The sum of all products of all SNVs for each gene and for each individual was then defined as the GS. The differential distribution of GSs between patients with SIADs and the control group, as well as their statistical significance, was tested using logistic regression in R. GS-based aggregate testing was also extended to the serological and clinical patient subgroups via case-control analysis and case-case analysis. In order to provide additional support to our GS-based approach, the results for the analysis contrasting all patients with SIADs and controls were also evaluated in comparison to those generated by the optimized sequence kernel association test (SKAT-O). In addition to including sex as a covariate based on its well-established status of risk factor for SIADs, all association analyses and models included the four most significant population principal components (PCs) as covariates. Four PCs significantly capture the intrinsic variation present in our cohort, which subsequently results in a system with homogenous structure and without any apparent stratification, cluster, or batch effect (Supplementary Figure 1). Statistical significance was evaluated after Bonferroni and Benjamini and Hochberg (false discovery rate [FDR]) correction ($\alpha = 0.05$).

In silico variant functional annotation and reporter assays for dual-specificity phosphatase 1 genetic variants. Using publicly available databases, catalogs, and computational tools, we prioritized genes of interest and in silico functionally annotated SNVs. For dual-specificity phosphatase 1 (*DUSP1*), reporter constructs were generated for SNVs or haplotypes as described by Roy et al¹⁹ (Supplementary Data 1), transfected in HeLa cells and analyzed by quantitative real-time polymerase chain reaction (qRT-PCR) or the dual-luciferase reporter assay. Experiments were repeated three times with five replicates for each construct, and statistical analyses were performed with unpaired Student's *t*-test and one-way analysis of variance. The endogenous *DUSP1* gene expression was analyzed by qRT-PCR after stimulation of HeLa cells with prednisolone.

All participants provided informed consent to participate in the study, and the study was approved by the regional ethics board in Uppsala (Dnr 2015/450 and 2016/155). Sequencing and genotype data at the individual level are not publicly available due to them containing information that could compromise research participant

privacy and consent. However, they are available from the corresponding author upon reasonable request and on a collaborative basis, protecting the participants' personal information. The bioinformatics scripts used to perform all the analyses described in this study are available at https://github.com/teone182/Dissect_CrossDisease.

RESULTS

The primary dataset with a total of 2,292 Scandinavian patients with SIADs with SLE ($n = 935$), pSS ($n = 906$), and myositis ($n = 451$), as well as 1,252 control individuals, was characterized by call rates $>98\%$ and included 400,491 SNVs. Of these, 294,001 were rare variants (MAF < 0.01).

Cross-disease single-variant and gene-based aggregate association analysis. To test whether we could detect novel SIAD risk loci, we performed both single-variant association analysis and gene-based aggregate association testing. Comparing all patients with SIADs and the control individuals using 106,490 common markers, single-variant analysis confirmed the major histocompatibility complex (MHC), *IRF5*, and *NCF2* loci as key players in systemic inflammatory autoimmunity. The strongest association was observed for an upstream variant of *HLA-DRB1* (rs7775984-G, adjusted $P = 2.92 \times 10^{-48}$; odds ratio [OR] 3.5, 95% confidence interval [CI] 3.0–4.1; Supplementary Figures 2 and 3; Supplementary Tables 1 and 2).

We further extended the comparison of the patient and control groups by implementing gene-based aggregate testing (results summarized in Supplementary Table 3). When comparing all patients ($n = 2,292$) and the control individuals ($n = 1,252$), the GS-based aggregate analysis identified Bonferroni-significant associations mainly in the MHC region (20 genes, top hits *C2*, *HLA-C*, *MSH5*, *TNXB*, Bonferroni $P < 3.59 \times 10^{-13}$). The other non-MHC significant aggregate associations hit *IRF5* (Bonferroni $P = 4.94 \times 10^{-6}$) and *YDJC* (Bonferroni $P = 0.029$), both known as SIAD-associated genes, as well as the novel *MAP3K6*, *SLC5A6*, and *CGREF1* (Bonferroni $P = 0.0051$, 0.011, and 0.037, respectively). At the FDR statistical threshold, several non-MHC genes showed aggregate associations, including those previously implicated in SIAD *PHRF1* (FDR $P = 0.0048$), *ETS1* (FDR $P = 0.0052$), and *UBE2L3* (FDR $P = 0.031$), as well as the novel *MX1* (FDR $P = 0.049$).

Our GS-based analysis detected both previously known and novel potential loci, thus confirming the validity of the test and its noteworthy performance. The comparison of the results obtained for all the genes analyzed with the GS-based method and SKAT-O indicated a marked correlation (Spearman rho 0.53, $P < 2.2 \times 10^{-16}$; Supplementary Figure 4). Out of the 1,795 genes tested, 2.6% ($n = 46$) and 3.4% ($n = 61$) were FDR significant in the GS and SKAT-O analysis (Supplementary Tables 3 and 4), respectively, with 35 showing overlap. Among these 35 overlapping genes, 12 were non-MHC genes. When only the set of 46 FDR-significant genes in the GS-

Table 1. Basic characteristics of the study participants*

	SLE	pSS	Myositis	Control	Total
Individuals	935	906	451	1,252	3,544
Females, n (%)	806 (86)	845 (93)	300 (67)	1,080 (86)	–
Mean age at diagnosis (min to max), y	39 (3–85)	47 (14–90)	51 (7–88)	NA	–
Mean age (min to max), y	NA	NA	NA	54 (19–88)	–
Origin, n (%)					
Sweden	935 (100)	687 (76)	256 (57)	1,043 (83)	2,921
Denmark	0 (0)	0 (0)	68 (15)	0 (0)	68
Norway	0 (0)	219 (24)	127 (28)	209 (17)	555
Serological subgroup, n (%)					
ANA ^a	920 (98)	681 (75)	178 (39)	ND	1,779
dsDNA ^b	438 (47)	29 (3)	ND	ND	467
SSA	343 (37)	644 (71)	120 (27)	ND	1,107
Ro52	244 (26)	155 (17)	101 (22)	ND	500
Ro60	313 (33)	273 (30)	27 (6)	ND	613
SSB	200 (21)	388 (43)	11 (2)	ND	599
RNP	177 (19)	43 (5)	21 (5)	ND	241
RF	125 (13)	297 (33)	10 (2)	ND	432
SM ^c	92 (10)	9 (1)	1 (0)	ND	102
Clinical subgroup, n (%)					
Arthritis ^d	732 (78)	168 (19)	153 (34)	NA	1,053
Cancer ^e	77 (8)	41 (5)	91 (20)	NA	209
Cardiac involvement ^f	176 (19)	0 (0)	57 (13)	NA	233
Hematologic disorder ^g	600 (64)	341 (38)	0 (0)	NA	941
Lung involvement ^h	23 (2)	41 (5)	183 (41)	NA	247
Raynaud phenomenon	153 (16)	253 (28)	124 (27)	NA	530
Renal disorder ⁱ	331 (35)	19 (2)	0 (0)	NA	350
Skin involvement ^j	794 (85)	94 (10)	208 (46)	NA	1,096

* Percentages were calculated on the total number of patients with each of the three diseases considered (ie, SLE, pSS, and myositis) and the control group. ACR, American College of Rheumatology; ANA, antinuclear antibodies; dsDNA, double-stranded DNA; max, maximum; min, minimum; NA, not applicable; ND, not determined; pSS, primary Sjögren disease; RF, rheumatoid factor; SIAD, systemic inflammatory autoimmune disease; SLE, systemic lupus erythematosus; SM, Smith.

^a ANA positivity refers to the ACR-11^{52,53} classification criterion for patients with SLE.

^b dsDNA positivity refers to the ACR-10B^{52,53} classification criterion for patients with SLE.

^c SM positivity refers to the ACR-10C^{52,53} classification criterion for patients with SLE.

^d Arthritis positivity refers to the ACR-5^{52,53} classification criterion for patients with SLE.

^e Cancer included any malignancy or lymphoma in patients with SLE, lymphoma in patients with pSS, and any malignancy in patients with myositis. A total of 41% of these patients developed the malignancy after SIAD diagnosis, 21% before SIAD diagnosis, and 38% were lacking data.

^f Cardiac involvement included pericarditis (ACR-6B)^{52,53} or myocardial infarction in patients with SLE, any heart dysfunction due to myositis in patients with myositis.

^g Hematologic disorder included hematologic disorder (ACR-9)^{52,53} in patients with SLE and anemia, leucopenia, lymphopenia, or thrombocytopenia in patients with pSS.

^h Lung involvement included interstitial lung disease in patients with SLE, pSS, and myositis.

ⁱ Renal disorder included lupus nephritis (ACR-7)^{52,53} in patients with SLE and interstitial nephritis in patients with pSS.

^j Skin involvement included malar rash (ACR-1)^{52,53}, discoid rash (ACR-2)^{52,53}, photosensitivity (ACR-3)^{52,53}, oral ulcer (ACR-4)^{52,53}, or alopecia in patients with SLE; dermal vasculitis in patients with pSS; and heliotrope rash, Gottron papules or sign, erythroderma, periungual erythema, mechanics hands, calcinosis, ulceration, or vasculitis in patients with myositis.

based analysis was considered, the GS- and SKAT-O-derived *P* values had a substantially stronger correlation (Spearman rho 0.69, $P < 4.34 \times 10^{-9}$; Supplementary Figure 5). Overall, this suggests that the two methods appear to be equivalent in terms of association outcomes. However, we note that in the context of our study, the GS-based approach might compare favorably with SKAT-O in interpretation and translational potential, providing additional information on the size and direction of effect for each tested gene in relation to any phenotype or set of individuals considered.

Aggregate testing of the cross-SIADs major serological patient subgroups. Because a hallmark of SLE,

pSS, and myositis is represented by shared circulating autoantibodies directed against nuclear antigens and immunoglobulins, we extended the GS-based test to those subsets of patients presenting with a specific autoantibody regardless of their primary SIAD (Table 1; Figure 1). GS-based case-control aggregate analysis of the major serological subgroups with SIADs mainly revealed strong Bonferroni-significant associations with MHC genes (Figure 2; Supplementary Tables 5 and 6). To more accurately dissect the genetic background underlying the occurrence of shared autoantibodies, we sought to control for the contribution of generalized SIAD genetic components, among which the MHC locus represents the strongest factor. For each defined

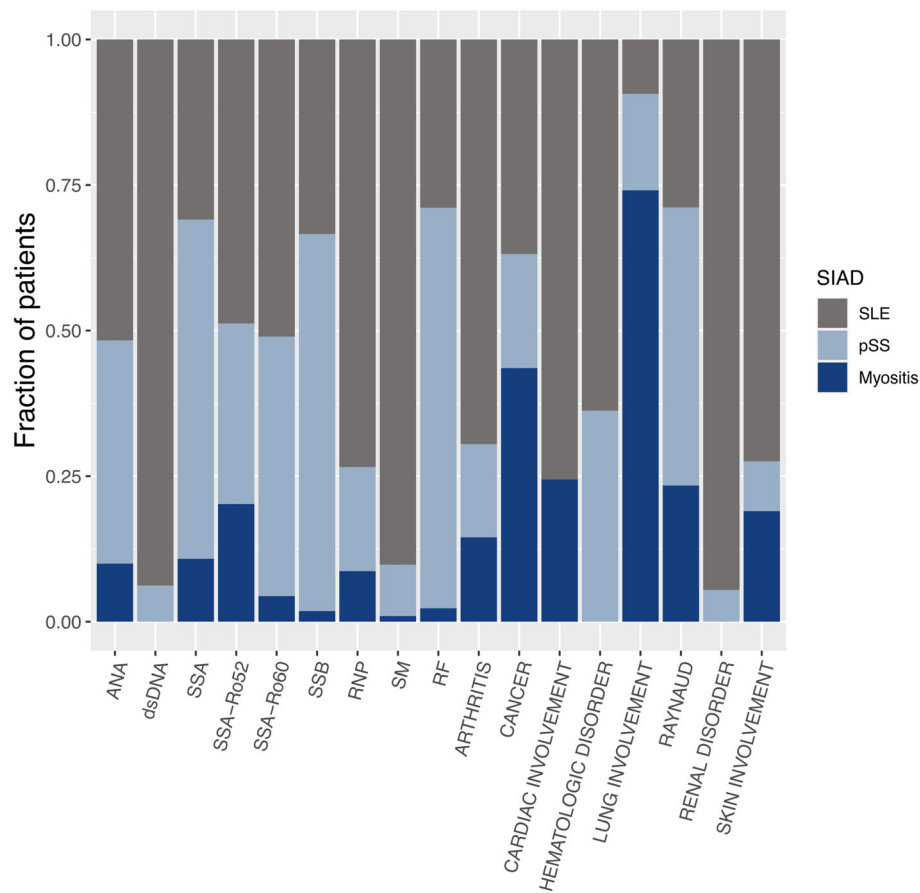


Figure 1. Bar plot showing the stratification of the different serological and clinical patient subgroups defined by the occurrence of autoantibodies and clinical features shared among patients with SIADs. Every subgroup is further categorized based on the prevalence of patients with SLE, pSS, or myositis. For additional numerical information, see Table 1 and Supplementary Table 1. ANA, antinuclear antibodies; dsDNA, double-stranded DNA; pSS, primary Sjögren disease; RF, rheumatoid factor; SIAD, systemic inflammatory autoimmune disease; SLE, systemic lupus erythematosus; SM, Smith.

serological subgroup, we therefore performed GS-based case–case aggregate association analysis by contrasting patients who were autoantibody positive against those who were autoantibody negative (Figure 2; Supplementary Tables 5 and 6). Although the RF, SSA, SSA-Ro52, SSA-Ro60, and SSB subgroups showed Bonferroni-corrected significant associations with the MHC region and no associations outside this locus, the subgroup of patients who were double-stranded DNA (dsDNA) positive displayed the opposite pattern, with a few associations emerging only in non-MHC genes (eight genes, top hit *RXRA*, Bonferroni $P = 7.39 \times 10^{-5}$). Interestingly, the association signals for the subgroup of patients who were ANA positive were weak and approximately in the same range of magnitude as in both MHC and non-MHC genes (median Bonferroni $P = 0.011$ and 0.0063 , respectively), with *ISG15* (Bonferroni $P = 0.0063$) being the only gene detected outside MHC (Figure 2). Finally, no association with either MHC or non-MHC genes was detected in the subgroup of patients who were RNP and SM positive, presumably due to their small sample size (Figure 2).

Characterization of clinical features shared among patients with SIADs. To better understand the etiology of those clinical manifestations and comorbidities shared among SIADs, we performed the same GS-based case–control and case–case aggregate association analyses but this time focused on those subsets of patients presenting with aggregate clinical conditions regardless of their primary SIAD diagnosis (Figure 1; Table 1; Supplementary Tables 7 and 8). Besides a few weak associations with non-MHC genes, GS-based case–control analysis revealed pervasive associations with the MHC region (Figure 3), reflecting the typical immunogenetic background of such disorders. In contrast, case–case analysis completely abolished MHC associations and detected genetic signals exclusively located in non-MHC loci. Apart from a few potentially pleiotropic genes associated with more than one aggregate clinical and/or serological feature across the case–control and/or case–case analyses (ie, *CEP131*, *COL61A*, *CXCL5*, *GNA13*, *IL17REL*, *ISG15*, *JUN*, *LSP1*, *PRM3*, *RXRA*, *TH*, and *TOGARAM2*), the case–case analysis also

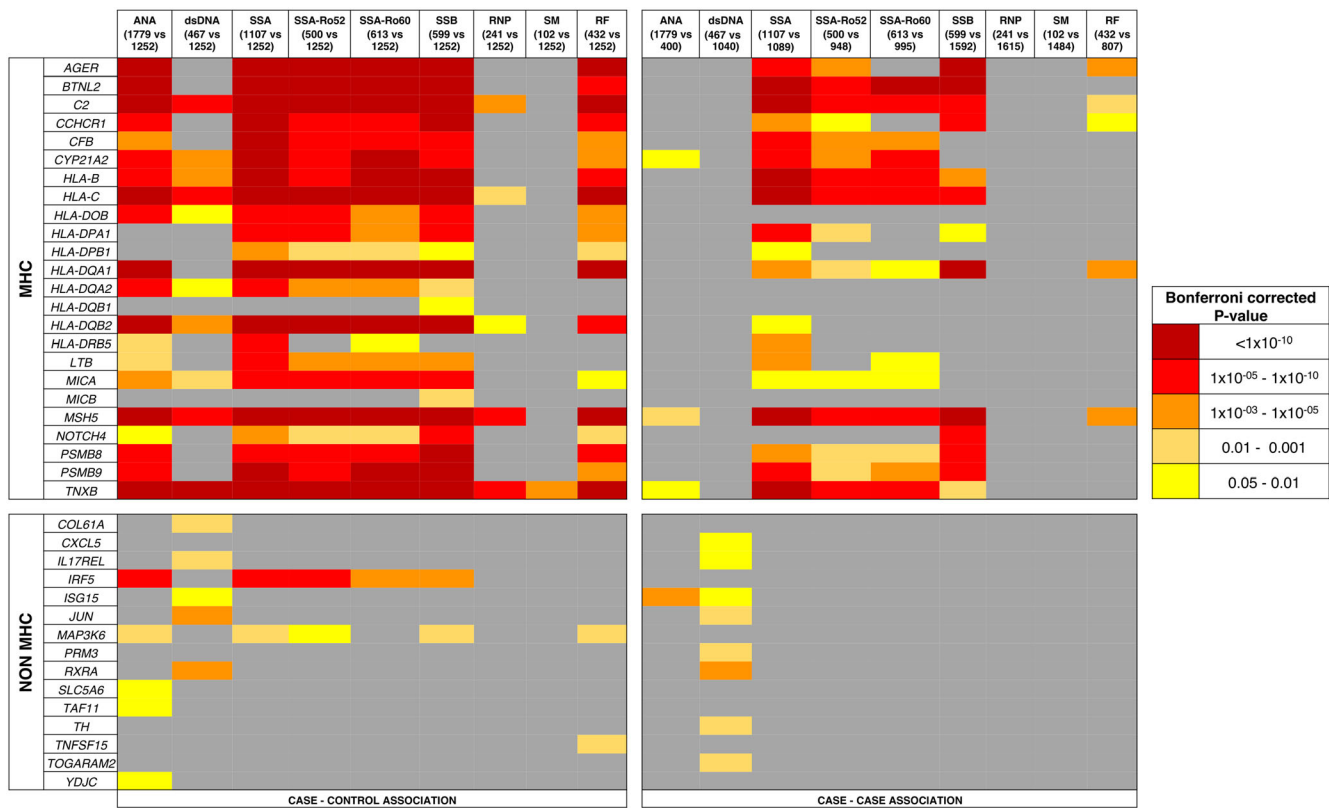


Figure 2. Heatmap showing the results of the genetic score-based aggregate case-control and case-case association testing for the subgroups of patients with SIAD defined based on the presence of shared autoantibodies. For each autoantibody indicated at the top, the results of the case-control analysis derive from the contrast between the patients with SIADs who were autoantibody positive and the control individuals, whereas the results of the case-case analysis derive from the contrast between the patients with SIADs who were autoantibody positive and those who were autoantibody negative (patients with missing data were not considered in the analysis). The number of samples of each group contrasted is shown under the autoantibody at the top. The results are further categorized into those pertaining genes located in the MHC locus and those located outside this region. The Bonferroni-corrected *P* values were obtained after correcting the association raw *P* values from the logistic regression model for the 1,795 genes tested. ANA, antinuclear antibody; dsDNA, double-stranded DNA; MHC, major histocompatibility complex; RF, rheumatoid factor; SIAD, systemic inflammatory autoimmune disease; SM, Smith.

highlighted genetic signals that were exclusively associated to unique subsets of patients with a particular clinical aggregate subphenotype (ie, signals not detected in additional aggregate clinical and/or serological subgroups in any of the case-control and case-case comparisons). This potentially indicates genetic contributions restricted to the clinically defined patient subgroups, as well as presumably independent on the genetic background of patients with SIADs (Figure 3). Protein kinase C zeta (*PRKCZ*; Bonferroni *P* = 0.015) and transcription factor 7 (*TCF7*) (Bonferroni *P* = 0.0020) were detected as uniquely associated with arthritis and cancer, respectively. After examining patients with SIADs with skin involvement in contrast to those without such manifestation, exclusive associations were detected in three genes (top hit *DUSP1*, Bonferroni *P* = 0.00022; *CDKN2A*, Bonferroni *P* = 0.0076; *BRDT*, Bonferroni *P* = 0.035). The analysis considering patients with cardiac or lung involvement, Raynaud phenomenon, and renal disorder did not show any aggregate significant associations.

The gene *DUSP1* as a candidate for protection against skin lesions. Although the detected case-case aggregate associations reflected functionally weighted burdens of all variants in the considered gene spaces, we investigated whether we could further pinpoint genes and variants of particular interest in the context of the etiology of each clinical subphenotype. To overcome the inherent difficulties linked to the prospective interpretation of pleiotropy and to pragmatically prioritize candidate genes and variants, we focused on *DUSP1*, which was the strongest genetic signal exclusively associated to a unique clinical subphenotype (ie, skin involvement) in the case-case analyses, despite *PRKCZ* and *TCF7* being also uniquely associated with arthritis and cancer, respectively. *DUSP1* emerged as the most remarkable candidate for further in-depth analysis given that (i) it has recently been implicated in another skin-related disease—eczema,²⁰ (ii) the genetic variation pattern and the protective genetic effect against the SIAD-associated clinical skin phenotype observed in our study (OR 0.53, 95% CI 0.42–0.68; Supplementary Table 8) are consistent with the results by Grosche et al,²⁰

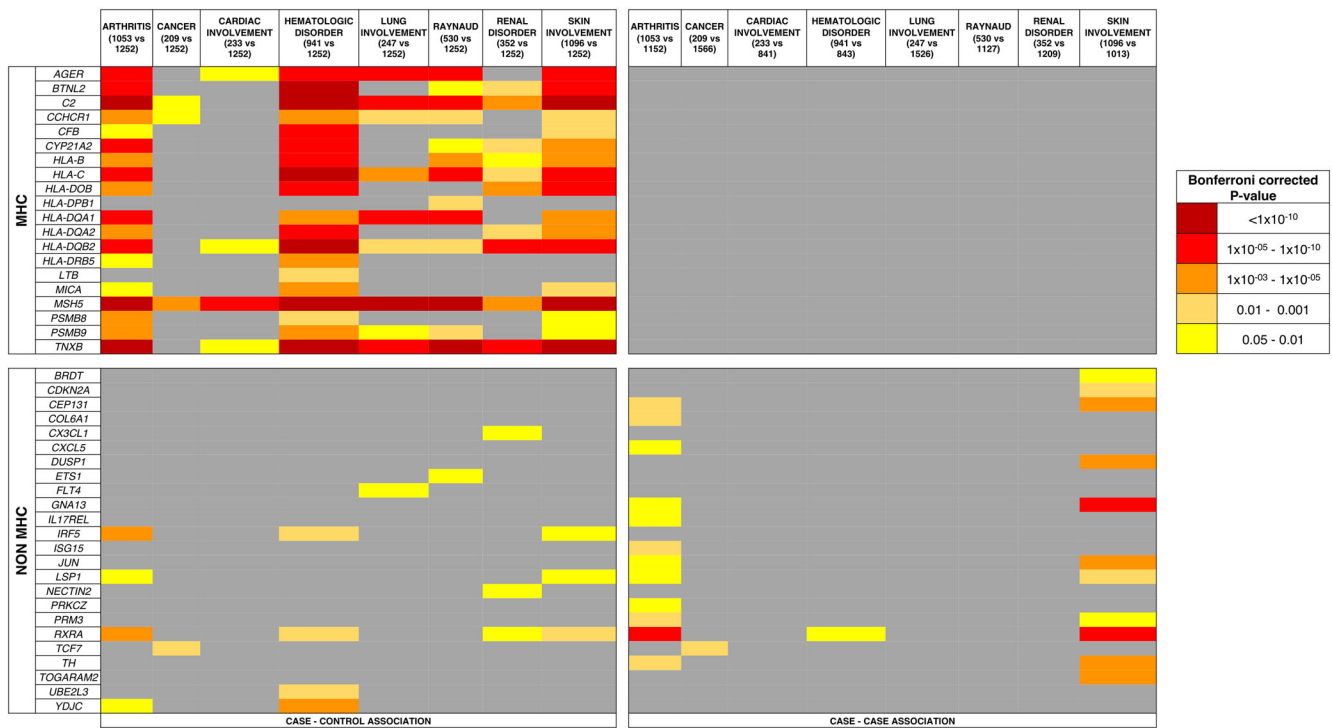


Figure 3. Heatmap showing the results of the genetic score–based aggregate case–control and case–case association testing for the subgroups of patients with SIAD defined based on the presence of shared clinical manifestations and comorbidities. For each clinical feature indicated at the top, the results of the case–control analysis derive from the contrast between the patients with SIADs who were positive for the feature and the control individuals, whereas the results of the case–case analysis derive from the contrast between the patients with SIADs who were positive and negative for the feature (patients with missing data are not considered in the analysis). The number of samples of each group contrasted is shown under the feature at the top. The results are further categorized into those pertaining genes located in the MHC locus and those located outside this region. The Bonferroni-corrected P values were obtained after correcting the association raw P values from the logistic regression model for the 1,795 genes tested. MHC, major histocompatibility complex; SIAD, systemic inflammatory autoimmune disease.

and (iii) the *DUSP1* SNVs detected in our study ($n = 32$) overlap with predicted or experimentally validated regulatory elements as well as with evolutionary constrained positions, thus indicating a likely functional role (Figures 4A–C and 5; Supplementary Tables 9 and 10). Of these 32 variants, 12 showed a protective genetic effect, 12 a risk effect, and 8 showed neutrality (Figure 4B). Furthermore, haplotype analysis identified a presumably protective haplotype block spanning the entire *DUSP1* gene (Supplementary Figure 6).

Confirmation of the regulatory function of *DUSP1* variants. To experimentally evaluate the regulatory potential of the 32 *DUSP1* variants detected in this study, we generated 49 reporter constructs carrying haplotypes identified in the patients and constituting 18 test regions in total (Figure 4D and E) and assessed them in HeLa cells normally expressing *DUSP1* (Supplementary Figure 7). We examined the expression levels of the reporters in unstimulated cells and cells stimulated with prednisolone, known to up-regulate *DUSP1*.²¹ For all variants, except for those immediately downstream of the *DUSP1* 3'UTR (3'UTR-SNVs, region 5; Supplementary Figure 8; Supplementary Table 11), we observed a substantial up-regulation of the

reporters' transcription compared to the control, indicating strong enhancer potential of all regions in HeLa cells (Supplementary Figure 9; Supplementary Table 11), thereby confirming their regulatory potential suggested by the in silico functional annotations. Notably, treatment with prednisolone resulted in either negative or neutral effect on the transcription levels of the majority of the reporter regions compared to the unstimulated conditions (Supplementary Figures 10 and 11; Supplementary Table 11).

A total of 11 of the 18 tested enhancer regions (61%) showed statistically significant differences in expression levels between allelic or haplotypic constructs in unstimulated cells. The effect alleles of regions 2 and 9 increased transcription of the reporter, whereas those of regions 4, 6, 10, 11, and 14 led to down-regulation. Three regions (ie, 15, 16, and 17) with haplotypes harboring two or more variants showed a complex outcome in which some effect alleles were associated with expression up-regulation and some others with down-regulation. Despite not reaching statistical significance, the effect alleles of regions 1, 3, 8, 12, and 13 led to stable expression up-regulation. Finally, regions 7 and 18 overlap with SINE and Alu repeat elements, respectively, and consistently did not show allelic differences in expression levels.

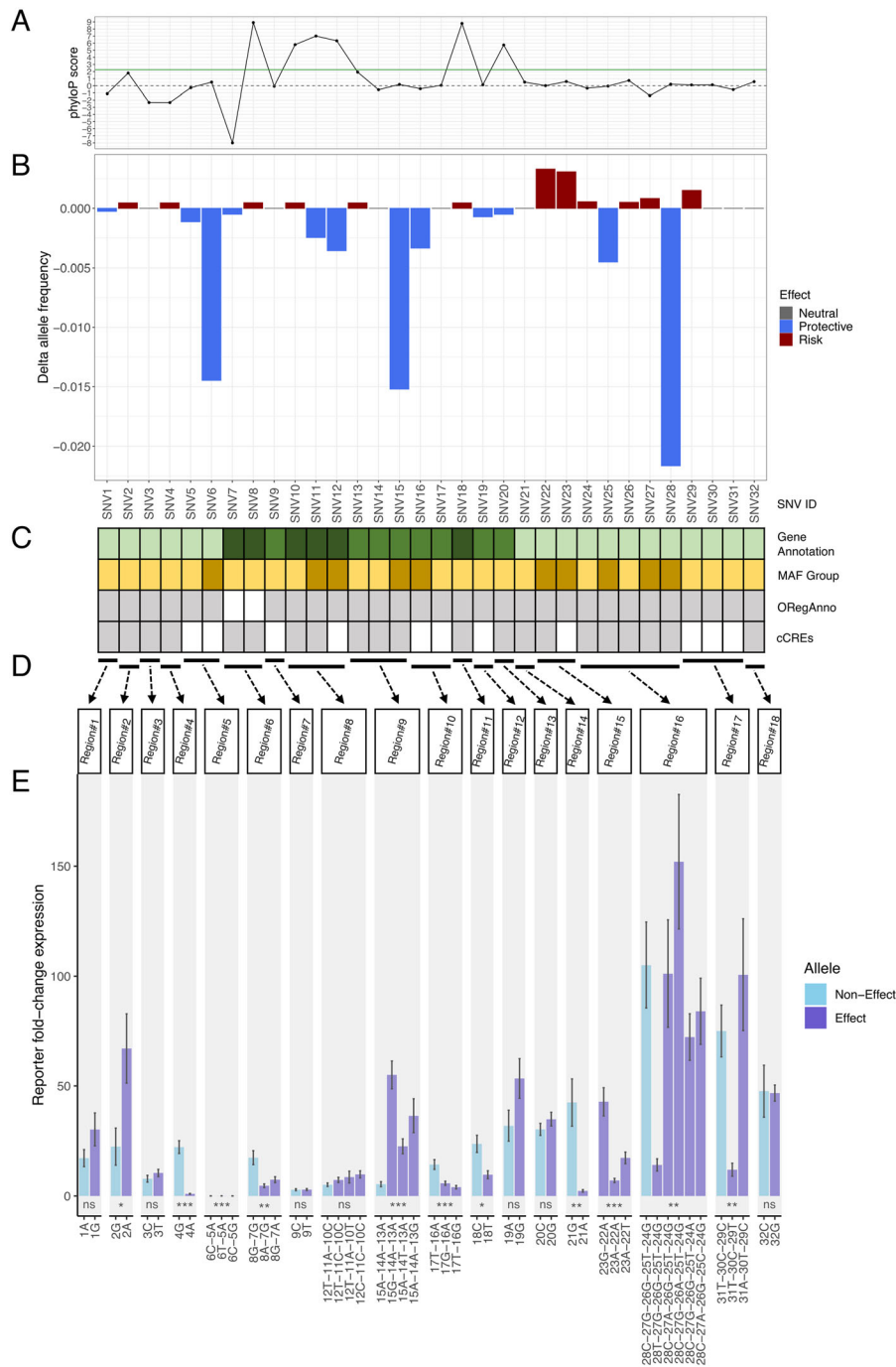


Figure 4. Genetic effects, in silico annotations, and reporter assay results for each *DUSP1* variant detected in our study. (A) Zoonomia Project evolutionary constraint metric.²⁶ (B) Delta allele frequency obtained as the minor (ie, effect) allele frequency difference between the patients with systemic inflammatory autoimmune diseases with and without skin involvement. (C) “Gene annotation”: regional annotation (pale green, upstream/downstream; green, intronic; dark green, exonic); “MAF group”: brown, common; yellow, rare; and “ORegAnno” and “cCREs”: gray, overlap with ORegAnno/ENCODE cCREs; white, no overlap. (D) The *DUSP1* variants included in each of 18 enhancer regions tested with reporter assay are depicted with horizontal black lines and connected with dashed arrows to the corresponding region. (E) Expression levels of all 49 reporter allelic constructs constituting the 18 total tested enhancer regions. Reporter constructs are shown on the x-axis and identified as SNV ID (only numerical) and allele for single-variant construct or multiple SNV IDs (only numerical) and alleles for haplotype constructs. Alleles and construct orientation refer to the negative strand and are relative to the transcription start site, respectively. Noneffect and effect alleles refer to the major and minor alleles, respectively. For details, see Supplementary Materials and Methods. Unpaired Student’s *t*-test and one-way analysis of variance were used for single-variant and haplotype constructs, respectively. **P* < 0.05, ***P* < 0.01, ****P* < 0.001. cCRE, candidate cis-regulatory element; *DUSP1*, dual-specificity phosphatase 1; ENCODE, Encyclopedia of DNA Elements; ID, identifier; MAF, minor allele frequency; ns, nonsignificant; ORegAnno, Open Regulatory Annotation database; SNV, single-nucleotide variant.

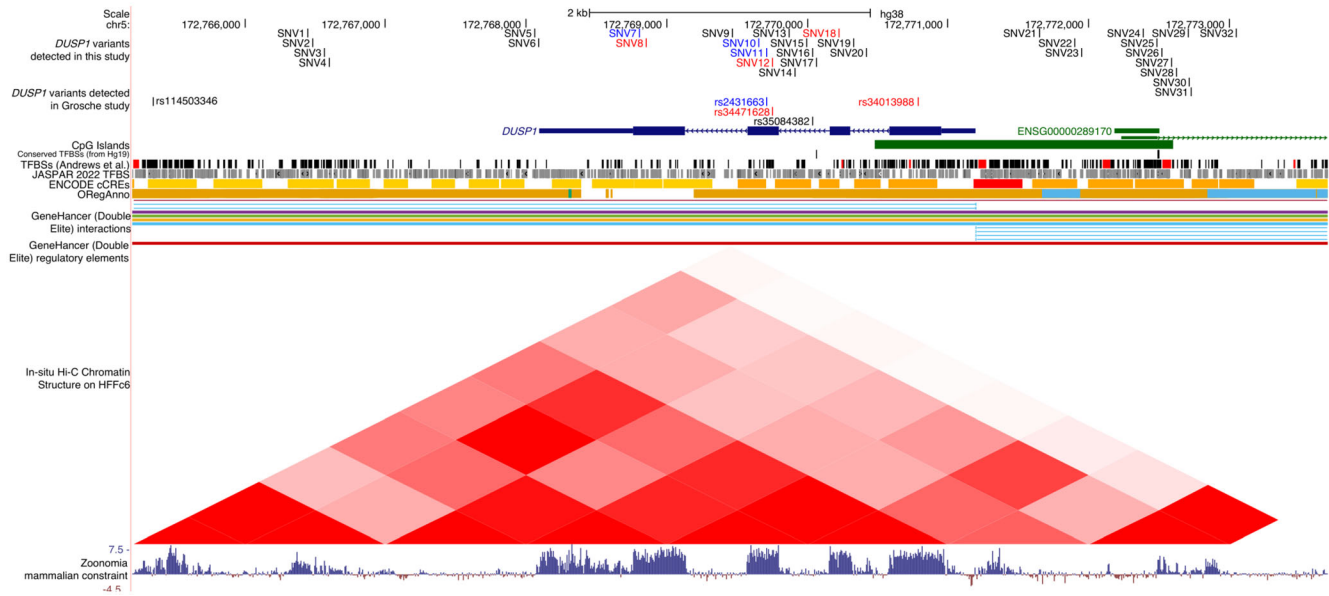


Figure 5. UCSC Genome Browser view (Hg38) of the *DUSP1* locus. *DUSP1* variants were detected and tested in this study (“*DUSP1* variants detected in this study”). Nonsynonymous and synonymous variants are indicated in red and blue, respectively. The same color labels apply for the *DUSP1* variants detected in the study by Grosche et al²⁰ (“*DUSP1* variants detected in Grosche study”). “Conserved TFBSs (from Hg19)” are conserved TFBSs lifted over to GRCh38. “TFBSs (Andrews et al)” are TFBSs sourced from the Zoonomia project²¹. CTCF TFBSs are colored in red. “JASPAR 2022 TFBSs” are experimentally defined TFBSs for eukaryotes. “ENCODE cCREs” are ENCODE Registry of cCREs in the human genome. “ORegAnno” is the open regulatory region associated with active gene expression. “GeneHancer (double elite) interactions” are highly filtered interactions between regulatory elements and genes. The light blue interactions are associations between *DUSP1* and cis-regulatory elements. “GeneHancer (double elite) regulatory elements” are highly filtered regulatory elements. “In situ Hi-C chromatin structure on HFFc6” is a zoomed-in heatmap of chromatin folding data from in situ HFFc6 cell lines. “Zoonomia mammalian constraint” consists of mammalian evolutionary constraint phyloP scores (Christmas et al²¹). cCRE, candidate cis-regulatory element; e; ENCODE, Encyclopedia of DNA Elements; ORegAnno, Open Regulatory Annotation database; SNV, single-nucleotide variant; TFBS, transcription factor binding site.

Bridging *DUSP1* genetic variation and reporters’ transcriptional activity.

Having available genetic and transcriptional activity data on *DUSP1* variants and the corresponding reporter regions, respectively, we further sought to explore their potential functional links. Considering the known anti-inflammatory effect of *DUSP1* overexpression,^{22–26} we reasoned that genetic variants with a protective effect on skin involvement would theoretically enhance *DUSP1* expression in the tested enhancer regions, whereas those categorized as risk variants would consequently down-regulate *DUSP1* gene expression. When comparing the genetic variant effect (in the case of a haplotypic construct, the variant with the greatest genetic effect was used as index) and the corresponding reporter allelic effect (ie, direction of *DUSP1* expression), we observed positive correlation in 11 regions (61%), including two protective exonic variants—SNV11 (rs2431663) and SNV12 (rs34471628)—identified previously²⁰ (Figure 4E; Supplementary Tables 11 and 12). Differently from the highly protective genetic effect of SNV15 correlating well with the statistically significant increase of reporter expression of the corresponding region 9, SNV28 in region 16 showed the greatest protective genetic effect but was individually associated to a decreased expression of the reporter. Region 16 showed

the most complex effect on transcription, with variable levels of expression depending on the tested haplotype. Interestingly, in foreskin fibroblast cells, this region appeared to be located in a separate topologically associated domain (TAD) comprising regions 14 to 18 (ie, from SNV21 to SNV32) and distinct from the one overlapping the locus extending from region 1 to region 13 (ie, from SNV1 to SNV20; Figure 5). The former also overlaps with several flanking CCCTC-binding factor (CTCF) binding sites as well as different GeneHancer interaction domains (Figure 5). Although SNV8 (region 6) was not an index variant in our analysis, this SNV and SNV18 (region 11) are evolutionarily constrained²¹ nonsynonymous risk variants and were also consistently associated to a marked down-regulation of the corresponding reporter’s expression.

DISCUSSION

Here, we exploited the power of next-generation sequencing coupled with a large and detailed collection of the clinical and serological data from patients with SIADs to investigate how genetic variation from the full allele frequency spectrum contributes to disease susceptibility and shared subphenotypes.

Although to a different extent and significance level, the single-variant and gene-based case-control association design contrasting all patients with SIADs and healthy individuals largely replicated known associations with shared loci controlling antigen processing and presentation (MHC), type I IFN- and proinflammatory cytokine gene-related signaling (*IRF5*, *UBE2L3/YDJC*, *PHRF1*, *SH2B3*), lymphocytes functional differentiation (*ETS1*), and reactive oxygen species metabolism (*NCF2*), all in the context of perturbed innate and adaptive immune response background typical of SIADs. Among the potentially novel loci identified via gene-based analysis, *MAP3K6*, *CGREF1*, and *MX1* are related to the type I IFN pathway activation,^{27–29} supporting a role in the pathophysiology of this pathway in all three SIADs. Although the type I IFN-mediated immune response has been widely observed and convincingly demonstrated at the transcriptomic level,^{30–32} these results further support a multilevel genetic dysregulation of this pathway in the three SIADs in our study.

Consistent with previous research indicating strong associations between *HLA* alleles and the occurrence of nuclear autoantibodies in patients with SLE, pSS, and myositis,^{33–36} our study indicated the MHC as the major susceptibility locus associated with autoantibody positivity. Differently from the subgroup of patients who were SSA/SSB positive in whom MHC appeared as the key and strongest genetic locus associated with both generalized autoimmune predisposition (ie, from case-control analysis) and specific autoantibody positivity (ie, from case-case analysis), the subgroups of patients who were ANA positive and dsDNA positive showed distinct association profiles in which MHC associations were weakened and abolished in the corresponding case-case analysis, respectively. The molecular heterogeneity of ANA together with the small sample size of the subset of patients who were ANA negative might explain the weak and diluted MHC associations in the case-case analysis of this serological subgroup. Outside the MHC locus, the case-case analysis identified an association with *ISG15*, a key IFN-stimulated gene, with the IFN signature previously linked to ANA positivity in individuals lacking an SIAD diagnosis.³⁷ In agreement with previous research in patients with SLE,³⁸ heterogeneity of MHC associations might also explain the lack of statistically significant signals within this locus when comparing the subgroups of patients who were dsDNA positive (dominated by patients with SLE) and those who were dsDNA negative. However, several statistically significant associations were detected outside the MHC region. From these observations, it is tempting to speculate that the production of dsDNA autoantibodies in patients with SIADs could at least in part be linked to environmental factors, such as previous bacterial infections, exerting their disease-triggering effects across a range of genetic susceptibility backgrounds. It has been demonstrated that antibodies targeting the highly immunogenic bacterial Z-DNA cross-react with self B-DNA, present in the serum.³⁹

Demonstrating the likely independence of the signals from the shared SIAD autoimmune background, the absence of

associations with the MHC locus from the comparison between patients with and without clinical subphenotypes suggested the possibility to pinpoint genetic factors and genes related to processes exclusively associated with each clinical subgroup of patients with SIADs. *PRKCZ*, exclusively associated with arthritis in the present study, has been implicated in patients with rheumatoid arthritis.^{40,41} *TCF7*, here uniquely linked to cancer, encodes TCF1, a key molecular player in different types of malignancies.^{42–45}

Finally, the strongest exclusive association was between *DUSP1* and skin involvement. Besides negatively regulating proinflammatory and T cell-mediated immune responses,^{46,47} this gene has recently been described as having a protective effect against eczema in a meta-analysis of approximately 20,000 patients and 380,000 controls from 21 different populations.²⁰ Consistently, here, we detected a gene-based aggregate negative association between *DUSP1* and skin manifestations in patients with SIADs. Using reporter assays, we confirmed that *DUSP1* variants were indeed functional and exerted cis-regulatory effects on gene expression.

Moreover, we found that the majority of the variants comply with the hypothesis that protective and risk variants would be linked to *DUSP1* higher and lower expression, respectively. Strikingly, two variants identified also in the study by Grosche et al²⁰ (ie, SNV11 and SNV12) followed our hypothesis, together with two evolutionary constrained nonsynonymous risk variants, one of which (ie, SNV8; p.Val338Met) introduces a predicted detrimental change at the end of the highly conserved C-terminal protein tyrosine phosphatase domain of *DUSP1*.⁴⁸ A mismatch between a specific genetic effect and reporter expression observed for some of the variants might be partly due to the disruption of the optimal local architecture of the cis-regulatory sequence in short but variant-rich cloned DNA fragments. This is presumably the case of the variant with the greatest protective effect (ie, SNV28) and its corresponding reporter region that is characterized by an extremely complex expression pattern. In this region, cell-type specific alternative regulation and functional rewiring mechanisms could also underlie such discrepancies, as suggested by the presence of distinct TADs, CTCF-binding sites, and different functional interaction domains.

In our study, we did not detect any notable induction of the reporters' expression upon prednisolone treatment despite the two-fold up-regulation of *DUSP1* in HeLa cells. This might be due to the lack of high-affinity glucocorticoid-response elements (GREs) for direct binding of glucocorticoid receptors in the cloned fragments or, alternatively, the GRE cell-specific mode of action.²⁶ Besides cross-replicating each other, the independently identified associations of *DUSP1* with SIAD-associated skin phenotype and eczema suggest the implication of this gene in skin lesions, regardless of the primary disease etiology. This is also confirmed by other reports describing *DUSP1* down-regulation in lesional psoriatic skin compared to paired nonlesional psoriatic

skin²⁴ and in dermal mesenchymal stem cells from patients with psoriasis compared to healthy controls²³ as well as the correlation between *DUSP1* overexpression and protection against inflammatory skin fibrosis in a mouse model with scleroderma²⁵ and better overall survival in skin cutaneous melanoma.⁴⁹ Intriguingly, a recent machine learning analysis of gene expression from patients with cutaneous SLE, psoriasis, eczema, and systemic sclerosis revealed that skin biopsies from all these conditions displayed, compared to controls, shared features positively correlated with their characteristic inflammatory processes.⁵⁰ Consequently, the hypothesis of *DUSP1* high expression as having a generalized protective effect against nonspecific skin conditions seems plausible.

In the context of our study, the known association of *DUSP1* with C-reactive protein (CRP) levels is of note. The supplementary analysis of a subset of our cohort of patients with SIADs indicated genetically supported higher levels of CRP in patients without skin phenotype. Additionally, this showed consistency with the substantial contribution of type I IFN activation to skin clinical manifestations, which is also accompanied by a decrease in CRP levels (for details, see Supplementary Figure 12).⁵¹

Despite overall replicating the *DUSP1* findings of a previous independent megastudy²⁰ and therefore demonstrating the validity of our analysis, the present work is primarily limited by the lack of a direct replication cohort. However, our cohort represents, to the best of our knowledge, the largest and most unique resource combination of patients with SLE, pSS, and myositis and homogeneously integration of coding and regulatory genetic variation from the full spectrum of allele frequency. Nonetheless, further research is warranted to validate these results in larger and more diverse cohorts and to explore the functional significance of the identified genetic variants in disease pathogenesis, as well as the shared and unique features of patients with SIADs. Moreover, we acknowledge that the *DUSP1* findings should be functionally validated with reporter assays in additional and more relevant *in vitro* models, such as skin-derived cell lines. At the same time, it is worth considering that such experiments might also be hampered by inherent technical difficulties (eg, cells refractory to transfection and/or exogenous reporter expression). In this study, discovery power might also be limited by the intrinsic design of our targeted array and GS-based method, which assays only a fraction of the genome and assumes the same variant direction of effect, respectively.

In summary, we detected known and potentially novel genetic loci associated with patients with SIADs that are congruent with previous genetic and transcriptomics analyses of such disorders. In parallel, we delineated unified and differential genetic signatures and profiles associated with subsets of patients with SIADs with shared serological and clinical features. This might partially explain the clinical heterogeneity found in patients who fulfilled the same disease classification criteria, thus being relevant for their molecular stratification. Moreover, in keeping with

previous findings on other skin disorders, we highlighted a potentially functional role of *DUSP1* genetic variants in the context of SIAD-related cutaneous manifestations. Together, this suggests that common molecular mechanisms may underlie pathogenic conditions that are shared among different inflammatory disorders, which could possibly be translated to effectively improve treatment, also in more generalized disease frameworks.

ACKNOWLEDGMENTS

We would like to thank all patients and control individuals taking part in this study for making this research possible. We thank Michael Dong for providing the Zoonomia phyloP and PhastCons evolutionary constraint scores. We also thank Gregory Andrews for providing the transcription factor binding sites dataset generated as part of the Zoonomia project. DNA sequencing and genotyping was performed at the SNP&SEQ Technology Platform in Uppsala. The facility is part of the National Genomics Infrastructure Sweden and Science for Life Laboratory. Computations were performed on resources provided by the Swedish National Infrastructure for Computing through Uppsala Multi-disciplinary Center for Advanced Computational Science under projects SNIC SENS 2017142 and 2017107.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Bianchi confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Helsinki Declaration requirements.

REFERENCES

- Bengtsson AA, Ronnblom L. Systemic lupus erythematosus: still a challenge for physicians. *J Intern Med* 2017;281(1):52–64.
- Dalakas MC. Inflammatory muscle diseases. *N Engl J Med* 2015;372(18):1734–1747.
- Mariette X, Criswell LA. Primary Sjogren's syndrome. *N Engl J Med* 2018;378(10):931–939.
- Moutsopoulos HM. Autoimmune rheumatic diseases: one or many diseases? *J Transl Autoimmun* 2021;4:100129.
- Acosta-Herrera M, Kerick M, Gonzalez-Serna D, et al; Myositis Genetics Consortium; Scleroderma Genetics Consortium. Genome-wide meta-analysis reveals shared new loci in systemic seropositive rheumatic diseases. *Ann Rheum Dis* 2019;78(3):311–319.
- Langefeld CD, Ainsworth HC, Cunninghame Graham DS, et al. Trans-ancestral mapping and genetic load in systemic lupus erythematosus. *Nat Commun* 2017;8:16021.
- Lessard CJ, Li H, Adrianto I, et al. Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjogren's syndrome. *Nat Genet* 2013;45(11):1284–1292.
- Rothwell S, Cooper RG, Lundberg IE, et al. Dense genotyping of immune-related loci in idiopathic inflammatory myopathies confirms HLA alleles as the strongest genetic risk factor and suggests different genetic background for major clinical subgroups. *Ann Rheum Dis* 2016;75(8):1558–1566.

9. Kilpinen H, Barrett JC. How next-generation sequencing is transforming complex disease genetics. *Trends Genet* 2013;29(1):23–30.
10. Lee S, Abecasis GR, Boehnke M, et al. Rare-variant association analysis: study designs and statistical tests. *Am J Hum Genet* 2014;95(1):5–23.
11. Abell NS, DeGorter MK, Gludemans MJ, et al. Multiple causal variants underlie genetic associations in humans. *Science* 2022;375(6586):1247–1254.
12. Zuk O, Schaffner SF, Samocha K, et al. Searching for missing heritability: designing rare variant association studies. *Proc Natl Acad Sci U S A* 2014;111(4):E455–E464.
13. Bianchi M, Kozyrev SV, Notarnicola A, et al. Contribution of rare genetic variation to disease susceptibility in a large Scandinavian myositis cohort. *Arthritis Rheumatol* 2022;74(2):342–352.
14. Sandling JK, Pucholt P, Hultin Rosenberg L, et al. Molecular pathways in patients with systemic lupus erythematosus revealed by gene-centred DNA sequencing. *Ann Rheum Dis* 2021;80(1):109–117.
15. Thorlacius GE, Hultin-Rosenberg L, Sandling JK, et al. Genetic and clinical basis for two distinct subtypes of primary Sjogren's syndrome. *Rheumatology (Oxford)* 2021;60(2):837–848.
16. Eriksson D, Bianchi M, Landegren N, et al. Extended exome sequencing identifies BACH2 as a novel major risk locus for Addison's disease. *J Intern Med* 2016;280(6):595–608.
17. Morgenthaler S, Thilly WG. A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a cohort allelic sums test (CAST). *Mutat Res* 2007;615(1–2):28–56.
18. Rentzsch P, Witten D, Cooper GM, et al. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res* 2019;47(D1):D886–D894.
19. Roy A, Sakthikumar S, Kozyrev SV, et al. Using evolutionary constraint to define novel candidate driver genes in medulloblastoma. *Proc Natl Acad Sci U S A* 2023;120(33):e2300984120.
20. Grosche S, Marenholz I, Esparza-Gordillo J, et al. Rare variant analysis in eczema identifies exonic variants in DUSP1, NOTCH4 and SLC9A4. *Nat Commun* 2021;12(1):6618.
21. Christmas MJ, Kaplow IM, Genereux DP, et al. Evolutionary constraint and innovation across hundreds of placental mammals. *Science* 2023;380(6643):eabn3943.
22. Abraham SM, Lawrence T, Kleiman A, et al. Antiinflammatory effects of dexamethasone are partly dependent on induction of dual specificity phosphatase 1. *J Exp Med* 2006;203(8):1883–1889.
23. Chang WJ, Niu XP, Hou RX, et al. LITAF, HHEX, and DUSP1 expression in mesenchymal stem cells from patients with psoriasis. *Genet Mol Res* 2015;14(4):15793–15801.
24. Kjellerup RB, Johansen C, Kragballe K, et al. The expression of dual-specificity phosphatase 1 mRNA is downregulated in lesional psoriatic skin. *Br J Dermatol* 2013;168(2):339–345.
25. Scotece M, Hamalainen M, Leppanen T, et al. MKP-1 deficiency exacerbates skin fibrosis in a mouse model of scleroderma. *Int J Mol Sci* 2023;24(5):4668.
26. Vockley CM, D'Ippolito AM, McDowell IC, et al. Direct GR binding sites potentiate clusters of TF binding across the human genome. *Cell* 2016;166(5):1269–1281.e19.
27. Haller O, Kochs G. Human MxA protein: an interferon-induced dynamin-like GTPase with broad antiviral activity. *J Interferon Cytokine Res* 2011;31(1):79–87.
28. Hou P, Yang K, Jia P, et al. A novel selective autophagy receptor, CCDC50, delivers K63 polyubiquitination-activated RIG-I/MDA5 for degradation during viral infection. *Cell Res* 2021;31(1):62–79.
29. Kim N, Kukkonen S, Del Pilar Martinez-Viedma M, et al. Tat engagement of p38 MAP kinase and IRF7 pathways leads to activation of interferon-stimulated genes in antigen-presenting cells. *Blood* 2013;121(20):4090–4100.
30. Bave U, Nordmark G, Lovgren T, et al. Activation of the type I interferon system in primary Sjogren's syndrome: a possible etiopathogenic mechanism. *Arthritis Rheum* 2005;52(4):1185–1195.
31. Crow MK, Kirou KA, Wohlgemuth J. Microarray analysis of interferon-regulated genes in SLE. *Autoimmunity* 2003;36(8):481–490.
32. Pinal-Fernandez I, Casal-Dominguez M, Derfoul A, et al. Identification of distinctive interferon gene signatures in different types of myositis. *Neurology* 2019;93(12):e1193–e1204.
33. Gottenberg JE, Busson M, Loiseau P, et al. In primary Sjogren's syndrome, HLA class II is associated exclusively with autoantibody production and spreading of the autoimmune response. *Arthritis Rheum* 2003;48(8):2240–2245.
34. Leclair V, Galindo-Feria AS, Rothwell S, et al. Distinct HLA associations with autoantibody-defined subgroups in idiopathic inflammatory myopathies. *EBioMedicine* 2023;96:104804.
35. Morris DL, Fernando MM, Taylor KE, et al. MHC associations with clinical and autoantibody manifestations in European SLE. *Genes Immun* 2014;15(4):210–217.
36. O'Hanlon TP, Carrick DM, Targoff IN, et al. Immunogenetic risk and protective factors for the idiopathic inflammatory myopathies: distinct HLA-A, -B, -Cw, -DRB1, and -DQA1 allelic profiles distinguish European American patients with different myositis autoantibodies. *Medicine (Baltimore)* 2006;85(2):111–127.
37. Wither J, Johnson SR, Liu T, et al. Presence of an interferon signature in individuals who are anti-nuclear antibody positive lacking a systemic autoimmune rheumatic disease diagnosis. *Arthritis Res Ther* 2017;19(1):41.
38. Chung SA, Taylor KE, Graham RR, et al. Differential genetic associations for systemic lupus erythematosus based on anti-dsDNA autoantibody production. *PLoS Genet* 2011;7(3):e1001323.
39. Spencer DM, Svenungsson E, Gunnarsson I, et al. The expression of antibodies to Z-DNA in the blood of patients with systemic lupus erythematosus: relationship to autoantibodies to B-DNA. *Clin Immunol* 2023;255:109763.
40. Nakano K, Whitaker JW, Boyle DL, et al. DNA methylome signature in rheumatoid arthritis. *Ann Rheum Dis* 2013;72(1):110–117.
41. Webster AP, Plant D, Ecker S, et al. Increased DNA methylation variability in rheumatoid arthritis-discordant monozygotic twins. *Genome Med* 2018;10(1):64.
42. Jeannot G, Boudousquie C, Gardiol N, et al. Essential role of the Wnt pathway effector Tcf-1 for the establishment of functional CD8 T cell memory. *Proc Natl Acad Sci U S A* 2010;107(21):9777–9782.
43. Osman A, Yan B, Li Y, et al. TCF-1 controls T(reg) cell functions that regulate inflammation, CD8(+) T cell cytotoxicity and severity of colon cancer. *Nat Immunol* 2021;22(9):1152–1162.
44. Sade-Feldman M, Yizhak K, Bjorgaard SL, et al. Defining T cell states associated with response to checkpoint immunotherapy in melanoma. *Cell* 2018;175(4):998–1013.e20.
45. Zhang J, Lyu T, Cao Y, et al. Role of TCF-1 in differentiation, exhaustion, and memory of CD8(+) T cells: a review. *FASEB J* 2021;35(5):e21549.
46. Chuang HC, Tan TH. MAP4K family kinases and DUSP family phosphatases in T-cell signaling and systemic lupus erythematosus. *Cells* 2019;8(11):1433.
47. Jeffrey KL, Camps M, Rommel C, et al. Targeting dual-specificity phosphatases: manipulating MAP kinase signalling and immune responses. *Nat Rev Drug Discov* 2007;6(5):391–403.
48. Camps M, Nichols A, Arkinstall S. Dual specificity phosphatases: a gene family for control of MAP kinase function. *FASEB J* 2000;14(1):6–16.
49. Singh MK, Altameemi S, Lares M, et al. Role of dual specificity phosphatases (DUSPs) in melanoma cellular plasticity and drug resistance. *Sci Rep* 2022;12(1):14395.

50. Martinez BA, Shrotri S, Kingsmore KM, et al. Machine learning reveals distinct gene signature profiles in lesional and nonlesional regions of inflammatory skin diseases. *Sci Adv* 2022;8(17):eabn4776.
51. Enocsson H, Karlsson J, Li HY, et al. The complex role of C-reactive protein in systemic lupus erythematosus. *J Clin Med* 2021;10(24):5837.
52. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1997;40(9):1725.
53. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1982;25(11):1271-1277.

APPENDIX A: DISSECT CONSORTIUM

Members of the DISSECT Consortium are as follows: Lars Rönnblom (Department of Medical Sciences, Rheumatology, Uppsala University, Uppsala, Sweden), Kerstin Lindblad-Toh (Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden; Broad Institute of MIT and Harvard, Cambridge, Massachusetts), Marie Wahren-Herlenius (Division of Rheumatology, Department of Medicine, Solna, Karolinska Institutet, Solna, Sweden; Karolinska University Hospital, Stockholm, Sweden; Broegelmann Research Laboratory, Department of Clinical Science, University of Bergen, Bergen, Norway), Gunnel Nordmark (Department of Medical Sciences, Rheumatology, Uppsala University, Uppsala, Sweden), Ingrid E. Lundberg (Division of Rheumatology, Department of Medicine, Solna, Karolinska Institutet, Solna, Sweden; Department of Gastro, Dermatology and Rheumatology, Karolinska University Hospital, Stockholm, Sweden), Ann-Christine Syvänen (Department of Medical Sciences, Molecular Medicine and Science for Life Laboratory, Uppsala University, Uppsala, Sweden), Johanna K. Sandling (Department of Medical Sciences, Rheumatology, Uppsala University, Uppsala, Sweden), Mateo Bianchi (Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden), Pascal Pucholt (Department of Medical Sciences, Rheumatology, Uppsala University, Uppsala, Sweden), Sergey V. Kozyrev (Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden), Maija-Leena Eloranta (Department of Medical Sciences, Rheumatology, Uppsala University, Uppsala, Sweden), Dag Leonard (Department of Medical Sciences, Rheumatology, Uppsala University, Uppsala, Sweden), Andreas Jönsen (Department of Clinical Sciences Lund, Rheumatology, Lund University, Skåne University Hospital, Lund, Sweden), Anders A. Bengtsson (Department of Clinical Sciences Lund, Rheumatology, Lund University, Skåne University Hospital, Lund, Sweden), Iva Gunnarsson (Division of Rheumatology, Department of Medicine, Solna, Karolinska Institutet, Solna, Sweden; Karolinska University Hospital, Stockholm, Sweden), Elisabet Sveinsson (Division of Rheumatology, Department of Medicine, Solna, Karolinska Institutet, Solna, Sweden; Karolinska University Hospital, Stockholm, Sweden), Leonid Padyukov (Division of Rheumatology, Department of Medicine, Solna, Karolinska Institutet, Solna, Sweden; Karolinska University Hospital, Stockholm, Sweden), Solbritt Rantapää-Dahlqvist (Department of Public Health and Clinical Medicine/Rheumatology, Umeå University, Umeå, Sweden), Christopher Sjöwall (Division of Inflammation and Infection, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden), Helena Enocsson (Division of Inflammation and Infection, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden), Roland Jonsson (Broegelmann Research Laboratory, Department of Clinical Science, University of Bergen, Bergen, Norway), Roald Omdal (Clinical Immunology Unit, Department of Internal Medicine, Stavanger University Hospital, Stavanger, Norway;

Broegelmann Research Laboratory, Department of Clinical Science, University of Bergen, Bergen, Norway), Daniel Hammenfors (Department of Rheumatology, Haukeland University Hospital, Bergen, Norway), Elke Theander (Department of Rheumatology, Skåne University Hospital Malmö/Lund University, Lund, Sweden), Eva Baecklund (Division of Rheumatology, Department of Medical Sciences, Uppsala University, Uppsala, Sweden), Gudny Ella Thorlacius (Division of Rheumatology, Department of Medicine, Solna, Karolinska Institutet, Solna, Sweden; Karolinska University Hospital, Stockholm, Sweden), Helena Forsblad-d'Elia (Department of Rheumatology and Inflammation Research, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden), Johan G. Brun (Department of Rheumatology, Haukeland University Hospital, University of Bergen, Bergen, Norway), Juliana Imgenberg-Kreuz (Department of Medical Sciences, Rheumatology, Uppsala University, Uppsala, Sweden), Karl A. Brokstad (Broegelmann Research Laboratory, Department of Clinical Science, University of Bergen, Bergen, Norway), Kathrine Skarstein (The Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, Bergen, Norway), Katrine Brække Norheim (Department of Rheumatology, Stavanger University Hospital, Stavanger, Norway; Institute of Clinical Science, University of Bergen, Bergen, Norway), Lilian Vasaitis (Division of Rheumatology, Department of Medical Sciences, Uppsala University, Uppsala, Sweden), Malin V. Jonsson (Section for Oral and Maxillofacial Radiology, Department of Clinical Dentistry, University of Bergen, Bergen, Norway), Marika Kvarnström (Division of Rheumatology, Department of Medicine, Solna, Karolinska Institutet, Solna, Sweden; Karolinska University Hospital, Stockholm, Sweden; Academic Specialist Center, Center for Rheumatology, Stockholm Health Services, Region Stockholm, Stockholm, Sweden), Per Eriksson (Division of Inflammation and Infection, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden), Sara Magnusson Bucher (Department of Rheumatology, Faculty of Medicine and Health, Örebro University, Örebro, Sweden), Silke Appel (Broegelmann Research Laboratory, Department of Clinical Science, University of Bergen, Bergen, Norway), Svein Joar Johnsen (Department of Rheumatology, Stavanger University Hospital, Stavanger, Norway), Thomas Mandl (Division of Rheumatology, Department of Clinical Sciences, Lund University, Malmö, Sweden), Lara Adnan Aqrabi (Department of Oral Surgery and Oral Medicine, Institute of Clinical Odontology, University of Oslo, Oslo, Norway; Department of Health Sciences, Kristiania University College, Oslo, Norway), Janicke Liaaen Jensen (Department of Oral Surgery and Oral Medicine, Institute of Clinical Odontology, University of Oslo, Oslo, Norway), Øyvind Palm (Department of Rheumatology, Oslo University Hospital, Oslo, Norway), Albin Björk (Division of Rheumatology, Department of Medicine, Solna, Karolinska Institutet, Solna, Sweden; Karolinska University Hospital, Stockholm, Sweden), Marie Fischer (Division of Rheumatology, Department of Medicine, Solna, Karolinska Institutet, Solna, Sweden; Karolinska University Hospital, Stockholm, Sweden), Antonella Notarnicola (Division of Rheumatology, Department of Medicine, Solna, Karolinska Institutet, Solna, Sweden; Department of Gastro, Dermatology and Rheumatology, Karolinska University Hospital, Stockholm, Sweden), Anna Tjärnlund (Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden), Maryam Dastmalchi (Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden), Louise Pyndt Diederichsen (Department of Rheumatology, Odense University Hospital, Odense, Denmark; Center for Rheumatology and Spine Diseases, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark), Maria Lidén, (Department of Medical Sciences, Rheumatology, Uppsala University, Uppsala, Sweden), Øyvind Molberg (Department of Rheumatology, Oslo University Hospital, Oslo, Norway; Institute of Clinical Medicine, University of Oslo, Oslo, Norway), Thomas Skogh (Division of Inflammation and Infection, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden), Awat Jalal (Rheumatology Clinic, Örebro University Hospital, Örebro, Sweden), Balsam Hanna (Department of

Rheumatology, Sahlgrenska University Hospital, Gothenburg, Sweden), Helena Hellström (Department of Rheumatology, Falun, Sweden), Tomas Husmark (Department of Rheumatology, Falun, Sweden), Åsa Häggström (Division of Rheumatology, Department of Medicine, Kalmar Hospital, Kalmar, Sweden), Anna Svärd (Department of Rheumatology, Falun, Sweden; Center for Clinical Research Dalarna, Uppsala University, Uppsala, Sweden), Jennifer Meadows (Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden), Jessika Nordin (Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden), Fabiana H. G. Farias (Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden; Department of Psychiatry, Washington University, St. Louis, Missouri), Johanna Dahlqvist (Department of Medical Sciences, Uppsala University, Uppsala, Sweden; Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden), Argyri Mathioudaki (Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden), and Daniel Eriksson (Department of Medicine, Solna, Karolinska Institutet, Solna, Sweden; Department of Endocrinology, Metabolism and Diabetes, Karolinska University Hospital, Stockholm, Sweden).

APPENDIX B: IMMUNOARRAY CONSORTIUM

Members of the ImmunoArray Consortium are as follows: Kerstin Lindblad-Toh (Science for Life Laboratory, Department of Medical

Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden; Broad Institute of MIT and Harvard, Cambridge, Massachusetts), Gerli Rosengren Pielberg (Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden), Anna Lobell (Office for Medicine and Pharmacy, Uppsala University, Uppsala, Sweden), Åsa Karlsson (Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden), Eva Murén (Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden), Göran Andersson (Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden), Kerstin M. Ahlgren (Department of Surgical Sciences, Uppsala University, Uppsala, Sweden), Lars Rönnblom (Division of Rheumatology, Department of Medical Sciences, Uppsala University, Uppsala, Sweden), Majja-Leena Eloranta (Division of Rheumatology, Department of Medical Sciences, Uppsala University, Uppsala, Sweden), Nils Landegren (Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden; Centre for Molecular Medicine, Department of Medicine, Solna, Karolinska Institute, Solna, Sweden), Olle Kämpe (Center for Molecular Medicine, Department of Medicine, Solna, Karolinska Institutet, Stockholm, Sweden; Department of Endocrinology, Metabolism and Diabetes, Karolinska University Hospital, Stockholm, Sweden; Science for Life Laboratory, Department of Medical Sciences, Uppsala University, Uppsala, Sweden; K.G. Jebsen Center for Autoimmune Diseases, University of Bergen, Norway), and Peter Söderkvist (Division of Cell Biology, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden).