Association of *NCF1* polymorphism with systemic lupus erythematosus and systemic sclerosis but not with ANCA-associated vasculitis in a Japanese population

Nozomi Yokoyama¹,², Aya Kawasaki¹,², Takashi Matsushita³, Hiroshi Furukawa⁴,¹,²,⁴, Yuya Kondo⁴, Fumio Hirano⁷,⁸, Ken-ei Sada⁹, Isao Matsumoto⁶, Makio Kusao¹⁰, Hirofumi Amano¹⁰, Shouhei Nagaoka¹¹, Keigo Setoguchi¹², Tatsuo Nagai¹³, Kota Shimada⁶,¹⁴, Shoji Sugii¹⁴, Atsushi Hashimoto¹⁶, Toshihiro Matsui¹⁶, Akira Okamoto¹⁵, Noriyuki Chiba¹⁶, Eiichi Suematsu¹⁷, Shigeru Ohno¹⁸, Masao Katayama¹⁹, Kiyoshi Migtai¹⁰, Hajime Kono²¹, Minoru Hasegawa²², Shigeto Kobayashi²³, Hidehiro Yamada²⁴,²⁵, Kenji Nagasaka²⁶, Takahiko Sugihara⁷,²⁷, Kunihiro Yamagata²⁸, Shoichi Ozaki²⁴, Naoto Tamura¹⁰, Yoshinari Takasaki¹⁰,²³, Hiroshi Hashimoto²⁹, Hirofumi Makino³⁰, Yoshihiro Arimura³¹, Masayoshi Harigai³², Shinichi Sato³³, Takayuki Sumida⁶, Shigeto Tohma⁶,⁵, Kazuhiko Takehara³, Naoyuki Tsuchiya¹,²

Genome-wide association studies of systemic lupus erythematosus (SLE) in Chinese and Korean populations demonstrated strong association of single nucleotide polymorphisms (SNPs) located in the...
**GTF2I-NCF1 region, rs73366469 (GTF2I), rs117026326 (GTF2I), rs80346167(GTF2IRD1) and rs201802880 (NCF1).** This region has also been associated with susceptibility to Sjögren syndrome and rheumatoid arthritis; however, association studies with systemic sclerosis (SSc) and ANCA-associated vasculitis (AAV) have not been reported. Here we made an attempt to confirm their associations with SLE in the Japanese population, to find the primarily associated SNP, and to investigate whether these SNPs are also associated with susceptibility to SSc and AAV. By genotyping these four SNPs on 842 SLE, 467 SSc, 477 AAV patients and 934 healthy controls, striking association was confirmed in Japanese SLE. In addition, these SNPs were significantly associated with susceptibility to SSc, but not with AAV. Conditional logistic regression analysis revealed that the association of NCF1 rs201802880, a missense SNP encoding p.Arg90His, can account for the association of other SNPs by linkage disequilibrium. These results suggested that GTF2I-NCF1 region is associated with susceptibility to multiple autoimmune rheumatic diseases but not with AAV, and the primarily associated variant may be the missense SNP in NCF1.

Autoimmune diseases are caused by a combination of multiple genetic and environmental factors, but the precise mechanisms of their development are largely unestablished. Genome wide association study (GWAS) is an efficient approach to identify the genetic factors of such complex disorders. GWAS of autoimmune rheumatic diseases including rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), systemic sclerosis (SSc) and ANCA-associated vasculitis (AAV) unanimously demonstrated that the strongest association signal is present within the major histocompatibility complex (MHC) until 2013, when GWAS of Sjogren’s syndrome (SS) in the Chinese population surprisingly demonstrated striking associations of single nucleotide polymorphisms (SNPs), rs73366469 (T>C), rs117026326 (C>T) and rs80346167 (G>A), in a region encoding general transcription factors GTF2I and GTF2IRD1, which were even stronger than that of the MHC region.

Subsequently, Immunogenetics and replication studies in Chinese and Korean populations also demonstrated strong association of the SNPs located at GTF2I region with susceptibility to SLE. Furthermore, this region was also reported to be associated with susceptibility to RA in Japanese and Korean populations. This region has also been shown to be associated with susceptibility to SLE in European American populations, albeit more weakly. Thus, the region appears to be one of the strongest genetic factors for multiple autoimmune rheumatic diseases in East Asian populations.

GTF2I encodes general transcription factor II-I (TFII-I). TFII-I usually localizes in the cytoplasm. It is translocated into the nucleus after activation by growth factors, B cell and T cell receptor triggering factors, and endoplasmic reticulum stress. In the nucleus, TFII-I binds to promoter regions of target genes and promotes transcription. In addition, cytoplasmic TFII-I regulates surface expression of Ca^{2+} channel protein TRPC3. Thus, TFII-I has relevant functions to autoimmune diseases.

On the other hand, NCF1 gene encoding neutrophil cytoplasmic factor 1, a subunit of NADPH oxidase, is one of the responsible genes for chronic granulomatous disease, and is located close to GTF2I and GTF2IRD1 genes. A naturally occurring reduction-of-function polymorphism of Ncf1 has been positionedly identified to be associated with severity of pristane-induced arthritis in rats. Subsequently, introduction of Ncf1 mutation in mice has been shown to be associated with arthritis, autoimmune encephalomyelitis, and also lupus-like phenotypes with glomerulonephritis and type I interferon signature. In humans, a missense variant (p.Arg90His, rs201802880) in NCF1, leading to reduction-of-function of NADPH oxidase, has also been associated with susceptibility to SLE. The NCF1 and GTF2I region variants are in linkage disequilibrium (LD), and two studies strongly suggested that the causative variant of this region is the NCF1 missense variant. However, because of the complicated genomic configuration of this region and the presence of NCF1 copy number variation (CNV) and highly homologous pseudogenes (NCF1B and NCF1C), further studies from various populations will be informative in establishing the genetic contribution of each variant of this chromosomal region.

SLE and SSc are both characterized by antinuclear antibodies, and a small proportion of patients exhibit symptoms of both diseases (SSc-SLE overlap syndrome). In a recent cohort study of SSc in Toronto, the prevalence of SSc-SLE overlap was 6.8%. Similarly, although rare, co-occurrence of SLE and AAV has been reported especially in MPO-ANCA positive AAV, and a concept of SLE-AAV overlap syndrome has been proposed. Such co-occurrence suggests the presence of shared genetic factors. With respect to the overlap of susceptibility alleles, out of 116 non-HLA loci associated with SLE with P < 5 x 10^{-8} in a large-scale Immunochip analysis (based on the summary statistics downloaded from the NHGRI-EBI GWAS Catalog for study5 downloaded on 07/23/2019) and 18 confirmed SSc susceptibility loci, 10 loci were shared by SLE and SSc. As for AAV, only three loci (PTPN22, PRTN3, SERPINA1) have been confirmed as susceptibility loci except for HLA, among which only PTPN22 is shared with SLE. Thus, a rather small proportion of SLE susceptibility loci appear to be shared with SSc and AAV. To distinguish the susceptibility loci shared by multiple autoimmune rheumatic diseases and those specific for each disease will eventually lead us to deeper understanding of pathogenesis of these diseases.

Although GTF2I-NCF1 region associations have been reported in SLE, SS and RA, association studies have not been reported for SSc and AAV. In addition, to our knowledge, association study between this region and SLE has not been reported in the Japanese population. In this study, we examined whether the SNPs in GTF2I-NCF1 region are associated with susceptibility to SSc and AAV in addition to SLE. We also made an attempt to identify which SNP plays the primary role among these four SNPs.
Table 1. Associations between the SNPs and SLE, SSc and AAV under the additive model (case-control analysis). Case-control association analysis was performed using logistic regression test under the additive model using R software. P values (P) and odds ratios (OR) were adjusted for sex. FDR P values (Q) were calculated by Benjamini-Hochberg method. The numbers of the samples in the parentheses show the numbers of those analyzed for rs201802880, after the exclusion of non-amplified samples at the nested PCR process. MAF; minor allele frequency. CI; confidence interval.

### Results

**Association of GTF2I-NCF1 region SNPs with overall SLE and SSc.** First, we examined whether the GTF2I-NCF1 region SNPs are also associated with susceptibility to SLE in the Japanese population. The previously reported risk alleles at the four SNPs were strikingly increased in patients with SLE in comparison with healthy controls also in the Japanese population (Table 1).

Next we performed the association tests of these SNPs with SSc. When compared with healthy controls, the same alleles as in SLE were significantly associated with SSc (Table 1). Among the SNPs, NCF1 SNP rs201802880 showed the strongest associations with susceptibility to SLE and SSc (SLE: uncorrected P value $P_{uncorr} = 3.77 \times 10^{-4}$, FDR P value $Q = 8.29 \times 10^{-4}$; Odds Ratio [OR] = 3.57, 95%CI 2.99-4.28; SSc: $P_{uncorr} = 2.40 \times 10^{-4}$, $Q = 0.0011$, OR = 1.50, 95%CI 1.21-1.87, both under the additive model).

In contrast, significant association was not detected in AAV (Table 1). The statistical power to detect association in AAV was calculated to be 0.91% (rs73366469), 48.4% (rs117026326), 73.4% (rs80346167) and 73.4% (rs201802880) for the risk allele with the OR of 1.3.

### Primary role of NCF1 rs201802880 among the four SNPs.** Next we constructed the LD plot of the SNPs of 876 healthy control samples using Haploview 4.2 software. All of the four SNPs were found to be in LD; however, LD between NCF1 rs201802880 and GTF2I SNPs was moderate (Fig. 1).

To determine the primarily associated SNP among the four, conditional logistic regression test was performed with adjustment by each SNP. Notably, the associations of rs201802880 remained significant when conditioned on other SNPs. In contrast, when conditioned on rs201802880, no significant difference remained in other SNPs (Table 2). Therefore, NCF1 rs201802880 was considered to be primarily associated with SLE and SSc, while the associations of rs73366469, rs117026326 and rs80346167 were thought to be secondarily caused by LD with the OR of 1.3.

**Association of NCF1 rs201802880 with clinical characteristics of SLE and SSc.** Finally, we tested whether NCF1 rs201802880 is associated with specific clinical characteristics of SLE and SSc. Patients with SLE were stratified according to the age of onset (< 20 years or ≥ 20 years), presence of renal disorders, neurological disorders, anti-dsDNA, anti-Sm and anti-RNP antibodies, and patients with SSc according to diffuse cutaneous SSc (dcSSc) or limited cutaneous SSc (lcSSc), presence or absence of anti-topoisomerase I antibody (ATA), anti-centromere antibody (ACA), and interstitial lung disease (ILD), and case-case analysis was performed. As shown in Table 3, rs201802880 A allele was significantly enriched in the patients with SLE with the age of onset < 20 years as compared with the patients with the age of onset ≥ 20 years.

Among the SSc patients, 23 were complicated by RA, SS and/or SLE. Because SLE, SS and RA were already previously reported risk alleles at the four SNPs were strikingly increased in patients with SLE in comparison with healthy controls also in the Japanese population (Table 1).

### Discussion

In this study, GTF2I-NCF1 region SNPs were strikingly associated with susceptibility to SLE also in the Japanese population. More importantly, the same alleles were found to be associated with susceptibility to SSc for the first time. On the other hand, association was not detected in AAV. Taken together with previous observations on RA and SS, GTF2I-NCF1 region represents a shared genetic factor for multiple autoimmune rheumatic diseases, but not for AAV.

NCF1 is located adjacent to GTF2I and GTF2IRD1, and variants in these genes are in LD. The genomic structure of NCF1 region is extremely complicated due to presence of two pseudogenes highly homologous to NCF1. Two recent studies performed careful association analysis of the GTF2I-NCF1 region with SLE,
and reported that a missense mutation in NCF1, rs201802880, may be the primarily associated variant in this region\textsuperscript{10,11}. Our findings on SLE are consistent with these studies. Taken together with the functional role of Ncf1 mutation shown by the rodent models\textsuperscript{7–9,18}, it is considered that NCF1 rs201802880 plays a causal role
of this shared genetic risk factor among multiple autoimmune rheumatic diseases. GTF2I-NCF1 associations of other three SNPs. Further studies on rs201802880 appears to be primarily associated and could account for the genetic NCF1 in the Japanese population. In addition, the same alleles were also associated with susceptibility to SSc, but not patients were classified according to the European Medicines Agency (EMEA) algorithm. Autoantibody profiles (ILD) was made by site investigators based on chest radiography and/or thoracic computed tomography. AAV as in SLE. This lack of association is unlikely to be caused by lack of detection power, because our sample size observation, because the role of neutrophil extracellular traps (NETs) has been strongly implicated in AAV as well. The risk allele rs201802880 A (the same allele is denoted as NCF1 – 339T in Olsson et al.) was shown to be associated with reduced function of NADPH oxidase, leading to the reduced production of reactive oxygen species (ROS). Interestingly, the reduced production of ROS has recently been shown to be associated with autoimmune diseases with elevated interferon response in rodents and humans, especially SLE, suggesting a regulatory role of ROS against autoimmunity. The present study also detected that the susceptibility allele rs201802880 A is significantly enriched in SLE patients with younger age of onset, which is consistent with the previous observations in the European population that the age at diagnosis of SLE was significantly younger in the patients carrying the susceptibility allele.

On the other hand, lack of association of GTF2I-NCF1 region with susceptibility to AAV was an unexpected observation, because the role of neutrophil extracellular traps (NETs) has been strongly implicated in AAV as well as in SLE. This lack of association is unlikely to be caused by lack of detection power, because our sample size had 73.4% detection power for a risk allele at NCF1 rs201802880 with OR of 1.3, and we did not observe even a trend for association. These results suggested that it is unlikely that this allele has substantial genetic contribution to overall AAV, although the possibility that the genetic effect of NCF1 plays a role in granulomatosis with angiitis (GPA) or proteinase 3-ANCA positive AAV which are rare in the Japanese population cannot be excluded at this point.

In view of the complexity of this genomic region, as well as potential functional relevance of both GTF2I/GTF2IRD1 and NCF1, further studies are required to dissect the genetic contribution of this region and to determine whether a single causally associated variation can account for the genetic effect, or multiple variants are independently involved.

In conclusion, the association between GTF2I-NCF1 region SNPs and susceptibility to SLE was replicated in the Japanese population. In addition, the same alleles were also associated with susceptibility to SSc, but not with AAV. Furthermore, NCF1 rs201802880 appears to be primarily associated and could account for the genetic associations of other three SNPs. Further studies on GTF2I-NCF1 region are required to establish the effect size of this shared genetic risk factor among multiple autoimmune rheumatic diseases.

### Methods

#### Subjects.
Genomic DNA samples from SLE (n = 842, 66 males [7.8%]), SSc (n = 467, 50 males [10.7%]), AAV (n = 477, 190 males [39.8%]) and healthy controls (n = 934, 364 males [39.0%]) were genotyped. All patients and healthy controls are unrelated Japanese, recruited at universities and rheumatology centers in Japan. SLE and SSc patients fulfilled the American College of Rheumatology classification criteria for each disease. Presence or absence of renal disorders and neurological disorders in SLE was classified by the same criteria. dcSSc and lcSSc were determined by ELISA.

| Additive model for A allele |   |   | OR (95% CI) |
|----------------------------|---|---|-------------|
|                            | P | Q |             |
| SLE onset < 20 years vs ≥ 20 years | 0.0033 | 0.0095 | 1.48 (1.14–1.93) |
| renal disorder present vs absent | 0.85 | 0.89 | 0.98 (0.80–1.20) |
| neurologic disorder present vs absent | 0.70 | 0.89 | 1.05 (0.81–1.36) |
| anti-dsDNA antibody present vs absent | 0.027 | 0.059 | 1.35 (1.04–1.77) |
| anti-Sm antibody present vs absent | 0.85 | 0.89 | 1.02 (0.82–1.27) |
| anti-RNP antibody present vs absent | 0.80 | 0.89 | 1.03 (0.81–1.31) |
| SSc dcSSc vs lcSSc | 0.45 | 0.76 | 0.87 (0.59–1.25) |
| ATA present vs absent | 0.94 | 0.89 | 1.01 (0.68–1.49) |
| ACA present vs absent | 0.63 | 0.89 | 1.09 (0.77–1.54) |
| ILD present vs absent | 0.65 | 0.89 | 1.08 (0.77–1.52) |

Table 3. Association study of NCF1 rs201802880 and clinical characteristics of SLE and SSc (case–case analysis). Genotypes of rs201802880 in SLE and SSc patients with and without specific clinical characteristics were compared using logistic regression analysis under the additive model for A allele with adjustment for sex. Significant enrichment of A allele was observed among SLE patients with onset of < 20 years as compared with onset of ≥ 20 years. FDR P values (Q) were calculated by Benjamini–Hochberg method. dcSSc: diffuse cutaneous SSc, lcSSc: limited cutaneous SSc, ATA: anti-topoisomerase I antibody, ACA: anti-centromere antibody, ILD: interstitial lung disease. Q: FDR P value.
Ethics statement. This study was reviewed and approved by the Ethics Committees of University of Tsukuba, and of the following institutes where the subjects were recruited (in alphabetical order): Aichi Medical University, Asahikawa Medical University, Ehime University, Fukuoka University, Hamamatsu University, Hokkaido University, Hyogo University, Iwate Prefectural Central Hospital, Jichi Medical University, Juntendo University, Kagawa University, Kanazawa University, Kawasaki Municipal Hospital, Kitano Hospital, Kitsato University Hospital, Kobe University Hospital, Kyorin University, Kyoto Prefectural University, Kyoto University, Kyoudo Hospital, Kyushu University, Nagasaki University, Nagoya City University, Nagoya University, National Hospital Organization Kyushu Medical Center, National Hospital Organization Himeji Medical Center, National Hospital Organization Morioka Medical Center, National Hospital Organization Nagoya Medical Center, National Hospital Organization Sagamihara Hospital, National Hospital Organization Shimoshizu National Hospital, Okayama University, Okayama Saiseikai General Hospital, Ome Municipal General Hospital, Saitama Medical Center Hospital, Sendai Shakaikoken Hospital, Shimane University, St. Marianna University, Teikyo University, Tenri Hospital, The University of Miyazaki, The University of Tokyo, Toho University, Tokyo Medical and Dental University, Tokyo Medical University Hachioji Medical Center, Tokyo Metropolitan Komagome Hospital, Tokyo Metropolitan Geriatric Hospital and Institute of Gerontology, Tokyo Metropolitan Tama Medical Center, Tokyo Women's Medical University, University of Tsukuba, Yokohama City Minato Red Cross Hospital, Yokohama City University Medical Center, Yokohama Minami Kyosai Hospital. This study was conducted in accordance with the principles of the Declaration of Helsinki and the Ethical Guidelines for Human Genome/Gene Analysis Research implemented by Ministry of Education, Culture, Sports, Science and Technology, Ministry of Health, Labour and Welfare, and Ministry of Economy, Trade and Industry, of Japan. Informed consent was obtained from all subjects.

Nested polymerase chain reaction for the genotyping of rs201802880. Because of the presence of high homologous NCF1 pseudogenes (NCF1B and NCF1C), nested polymerase chain reaction (PCR) was employed before TaqMan SNP genotyping assay for rs20180288010. An NCF1-specific fragment was amplified using KOD FX Neo (TOYOBO, Osaka, Japan) by targeting the GTGT sequence in the exon 2 of NCF1 (the primer sequences are shown in Supplementary Table 1). The PCR conditions consisted of initial denaturation at 94 °C for 2 min, followed by 35 cycles of denaturation at 98 °C for 10 s, annealing at 60 °C for 60 s, and elongation at 68 °C for 2 min. Next, agarose gel electrophoresis was performed to validate the amplification of PCR, and samples without PCR products were excluded. After exclusion, 1μL of the PCR product diluted 1:100 was subjected to TaqMan SNP Genotyping Assay.

TaqMan SNP genotyping assay. The genotypes of SNPs were determined by TaqMan SNP genotyping assays (ABI 7300, Applied Biosystems). For rs73366469 and rs80346167, the premade primer/probe sets were used (Assay ID: rs73366469: C_97234117_10 and rs80346167: C_100871497_10; Applied Biosystems), and for rs117026326 and rs201802880, the customized primer/probe sets were used (Applied Biosystems, the sequences were shown in Supplementary Table 1). For PCR, DNA samples were added to the reaction mixture containing TaqMan® Genotyping Master Mix (Applied Biosystems) and TaqMan probes. The PCR conditions consisted of initial denaturation at 95 °C for 10 min, followed by 40 cycles (for rs73366469, rs80346167 and rs117026326) or 25cycles (for rs201802880) of denaturation at 95 °C for 15 s, annealing at 60 °C for 60 s.

Statistical analysis. Association analysis was performed using logistic regression analysis using R software (https://journal.r-project.org) with adjustment for sex. The analysis was performed under the additive, dominant and recessive models (Table 1, Supplementary Tables 2 and 3), and because the Akaike's Information Criteria (AIC) was the lowest for all SNPs under the additive model in SLE, and almost equal under the three models in SSc and AAV (Supplementary Table 4), the additive model was selected for the association analysis throughout the study. P values for all case-control (Table 1) and case-case analyses (Table 3), 22 comparisons in total, were adjusted for multiple comparisons by controlling false discovery rate (FDR) using Benjamini-Hochberg method24. FDR-adjusted P (Q value) < 0.05 was considered significant. Statistical power was calculated by Power and Sample Size Calculation version 3.1.2, 2014 (http://biostat.mc.vanderbilt.edu/wiki/Main/PowerSampleSize)25. LD plot was constructed using Haplovie 4.2 software (https://www.broadinstitute.org/).

Data availability
Based on the “Act on the Protection of Personal Inormation” enforced in Japan and the conditions on which the informed consent was given, it is not permitted to disclose an individual's genotypes and clinical information. All publicly available data generated or analyzed during this study are included in this published article and its Supplementary Information.

Received: 30 August 2018; Accepted: 24 October 2019;
Published online: 08 November 2019

References
1. Parkes, M., Cortes, A., van Heel, D. A. & Brown, M. A. Genetic insights into common pathways and complex relationships among immune-mediated diseases. Nat Rev Genet 14, 661–673 (2013).
2. Li, Y. et al. A genome-wide association study in Han Chinese identifies a susceptibility locus for primary Sjögren's syndrome at 7q11.23. Nat Genet 45, 1361–1365 (2013).
3. Sun, C. et al. High-density genotyping of immune-related loci identifies new SLE risk variants in individuals with Asian ancestry. Nat Genet 48, 323–330 (2016).
4. Kim, K. et al. Association-heterogeneity mapping identifies an Asian-specific association of the GTF2I locus with rheumatoid arthritis. *Sci Rep* 6, 27566, https://doi.org/10.1038/srep27566 (2016).

5. Langefeld, C. D. et al. Transcranial mapping and genetic load in systemic lupus erythematosus. *Nat Commun* 8, 16021, https://doi.org/10.1038/ncomms16021 (2017).

6. Roy, A. L. Pathophysiolo... 467–475 (2018).

7. Olofsson, P. et al. Positional identification of NCF1 as a gene that regulates arthritis severity in rats. *Nat Genet* 33, 25–32 (2003).

8. Hultqvist, M. et al. Enhanced autoimmunity, arthritis, and esophageal dysmotility in mice with a reduced oxidative burst due to a mutation in the Ncf1 gene. *Proc Natl Acad Sci USA* 101, 12646–12651 (2004).

9. Kelkka, T. et al. Reactive oxygen species deficiency induces autoimmunity with type 1 interferon signature. *Antioxid Redox Signal* 21, 2231–2245 (2014).

10. Zhao, J. et al. A missense variant in NCF1 is associated with susceptibility to multiple autoimmune diseases. *Nat Genet* 49, 433–437 (2017).

11. Olsson, L. M. et al. A single nucleotide polymorphism in the NCF1 gene leading to reduced oxidative burst is associated with systemic lupus erythematosus. *Ann Rheum Dis* 76, 1607–1613 (2017).

12. Alharbi, S. et al. Epidemiology and survival of systemic sclerosis-systemic lupus erythematosus overlap syndrome. *J Rheumatol* 45, 1406–1410 (2018).

13. Hervier, B. et al. Systemic lupus erythematosus associated with ANCA-associated vasculitis: an overlapping syndrome? *Rheumatol Int* 32, 3285–3290 (2012).

14. Buniello, A. et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 47, D1005–D1012 (2019).

15. Bossini-Castillo, L., López-Isac, E. & Martín, J. Immunogenetics of systemic sclerosis: Defining heritability, functional variants and shared-autoimmunity pathways. *J Autoimmun* 64, 53–65 (2015).

16. Lyons, P. A. et al. Genetically distinct subsets within ANCA-associated vasculitis. *N Engl J Med* 367, 214–223 (2012).

17. Merkel, P. A. et al. Identification of for functional and expression polymorphisms associated with risk for antineutrophil cytoplasmic autoantibody-associated vasculitis. *Arthritis Rheumatol* 69, 1054–1066 (2017).

18. Zhong, J. et al. Association of NOX2 subunits genetic variants with autoimmune diseases. *Free Radic Biol Med* 125, 72–80 (2018).

19. Apel, F., Zychlinsky, A. & Kenny, E. F: The role of neutrophil extracellular traps in rheumatic diseases. *Nat Rev Rheumatol* 14, 467–475 (2018).

20. Hochberg, M. C. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 40, 1725 (1997).

21. Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). *Arthritis Rheum* 23, 581–590 (1980).

22. LeRoy, E. C. et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *J Rheumatol* 15, 202–205 (1988).

23. Watts, R. et al. Development and validation of a consensus methodology for the classification of the ANCA-associated vasculitides and polyarteritis nodosa for epidemiological studies. *Ann Rheum Dis* 66, 222–227 (2007).

24. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B* 57, 289–300 (1995).

25. Dupont, W. D. & Plummer, W. D. Jr. Power and sample size calculations. A review and computer program. *Control Clin Trials* 11, 116–128 (1990).
Takeda Pharmaceutical Co. Ltd., Mitsubishi-Tanabe Pharma Co., Astellas Pharma Inc., Bristol Myers Squibb K.K., Pfizer Japan Inc., UCB Japan Co. Ltd., and Abbvie Japan Co., Ltd., and has received research grant support from Takeda Pharmaceutical Co., Ltd, Astellas Pharma Inc., and Teijin Pharma Ltd. Dr. A. Hashimoto has received research funding from Chugai Pharmaceutical Co., Ltd. Dr. Matsui has received research funding from Chugai Pharmaceutical Co., Ltd. and Janssen Pharmaceutical K.K. Dr. Tamura has received research grant from Ayumi Pharmaceutical Co., Asahi Kasei Pharma Co., Chugai Pharmaceutical Co., Ltd., Takeda Pharmaceutical Co. Ltd., Mitsubishi-Tanabe Pharma Co., Astellas Pharma Inc., Bristol Myers Squibb K.K., Abbvie Japan Co., Ltd., and Eisai Co., Ltd., and speaker’s fee from Abbvie Japan Co., Ltd., Bristol Myers Squibb K.K., Janssen Pharmaceutical K.K., and Chugai Pharmaceutical Co. Dr. Makino is a consultant for AbbVie and Teijin, receives speaker honoraria from Boehringer-Ingelheim. Dr. Harigai has received research grants and/or honoraria from Abbott Japan Co., Ltd.; Astellas Pharma Inc.; Bristol-Myers Squibb K.K.; Chugai Pharmaceutical Co., Ltd.; Eisai Co., Ltd.; Janssen Pharmaceutical K.K.; Mitsubishi Tanabe Pharma Co.; Santen Pharmaceutical Co.; Ltd.; Takeda Pharmaceutical Co., Ltd.; Teijin Pharma, Ltd.; and Pfizer Japan Inc. Dr. Tsuchiya has received a research grant from Bristol-Myers Squibb, 2015 Japan College of Rheumatology Award from Japan College of Rheumatology and 2017 Novartis Rheumatology Award from Japan Rheumatism Association with research funding, and speaker’s honoraria from Ayumi Pharmaceutical Corporation.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-019-52920-0.

Correspondence and requests for materials should be addressed to A.K. or N.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019