Analysis of anti-RNA polymerase III antibody positive systemic sclerosis suggests altered GPATCH2L and CTNND2 expression in scleroderma renal crisis

Edward P. Stern[1, 2 *], Sandra G. Guerra[1 *], Harry Chinque[1], Vanessa Acquaah[1], David González-Serna[3], Markella Ponticos[1], Javier Martin[3], Voon H. Ong[1], Korsa Khan[1], Svetlana I. Nihtyanova[1], Mark Harber[2], Aine Burns[2], Maureen D Mayes[4], Shervin Assassi[4], Carmen Fonseca[1], Christopher P. Denton[1]

[*]Joint first authors

Affiliations
[1]UCL Centre for Rheumatology and Connective Tissue Diseases, London, U.K.
[2]UCL Centre for Nephrology, London, U.K.
[3]Instituto de Parasitologia y Biomedicina Lopez-Neyra, Consejo Superior de Investigaciones Científicas (CSIC), Granada, Spain
[4]University of Texas Houston – McGovern Medical School, Houston, TX, U.S.A.

Author for correspondence:
Christopher Denton PhD FRCP
Professor of Experimental Rheumatology
UCL Centre for Rheumatology and Connective Tissue Diseases
2nd Floor – UCL Medical School Building
Royal Free Campus
Rowland Hill Street
London NW3 2PF, U.K.

Tel +44 (0) 20 7794 0432
c.denton@ucl.ac.uk

Abstract

Objective

Scleroderma renal crisis (SRC) is a life-threatening complication of systemic sclerosis (SSc) strongly associated with anti RNA polymerase III antibody (ARA) autoantibodies. We explore genetic susceptibility and altered protein expression in renal biopsy specimens in ARA positive SRC.

Methods

ARA-positive patients (n=99) with at least 5 years' follow-up (49% with a history of SRC) were selected from a well-characterised SSc cohort (n=2254). Cases were genotyped using the Illumina Human Omni-express chip. Based on initial regression analysis, nine SNPs were chosen for validation in a separate cohort of 256 ARA+ patients (40 with SRC). Immunostaining of tissue sections from SRC or control kidney was used to quantify expression of candidate proteins based upon genetic analysis of the discovery cohort.

Results

Analysis of 641,489 SNPs suggested association of POU2F1 (rs2093658; $1.98x10^{-5}$), CTNND2 (rs1859082; $p=7.14 \times 10^{-5}$), HECW2 (rs16849716; $p=1.2 \times 10^{-4}$) and GPATCH2L (rs935332; $p=4.92 \times 10^{-5}$) with SRC. Furthermore, the validation cohort showed an association between rs935332 within the GPATCH2L region, with SRC (p=0.025). Immunostaining of renal biopsy sections showed increased tubular expression of GPATCH2L (p=0.026), and glomerular expression of CTNND2 (p=0.026) in SRC samples (n=8) compared with normal human kidney controls (n=8), despite absence of any genetic replication for the associated SNP.

Conclusions

Increased expression of two candidate proteins GPATCH2L and CTNND2 in SRC compared with control kidney suggests a potential role in pathogenesis of SRC. For GPATCH2L this may reflect genetic susceptibility in ARA positive SSc based upon 2 independent cohorts

Abstract 243 words

Keywords: scleroderma, renal, genetics, autoantibody, catenin, hypertension

Introduction

Systemic sclerosis (scleroderma; SSc) is a multisystem rheumatic disorder with a high case-specific mortality due to internal organ complications of the disease. The pattern and frequency of internal organ manifestations reflects the different antinuclear antibody reactivities that are a hallmark of SSc [1]. These include anti-centromere (ACA), anti-topoisomerase-1 (ATA) and anti-RNA polymerase III (ARA) antibodies. Frequency of major complications of SSc differs between these clinical and immunological subgroups. For example ATA associates with interstitial lung disease (ILD) [2] and ACA with pulmonary arterial hypertension (PAH) in certain subgroups [3]. A strong association has been observed between the presence of ARA and the occurrence of scleroderma renal crisis (SRC) [4,5].

SRC is a life-threatening complication of SSc characterised by accelerated phase hypertension and acute kidney injury [6]. SRC was almost universally fatal until the late 1970s, when its management was revolutionised by the introduction of Angiotensin Converting Enzyme inhibitors. Nevertheless, in the modern era 40-50% of SRC cases result in early death or renal replacement therapy and the five-year survival is as low as 50% in some series [7,8].

The different occurrence of SRC in specific subgroups of SSc, the observation that even within the highest risk groups only a minority of SSc cases develop SRC, and the evidence that it generally only occurs in early stage disease, are all consistent with a genetic predisposition to SRC that may be independent of the inherited risk of SSc itself. Data regarding the genetic contribution to SRC risk are limited; a single study reports the presence of MHC class I haplotypes, such as HLA-DRB1*04:07 and HLA-DRB1*13:04, as independent risk factors for developing SRC [9].

The present study has used a novel approach by looking at cases of SRC in the ANA group at highest risk, comparing cases that develop SRC with those that appear to be protected from development of SRC during long-term follow-up. This strategy

provides a platform of clinical and serological homogeneity in which we hypothesise that genetic susceptibility factors may be most relevant.

Materials and Methods

Patient demographics

For this study SRC patients (n=99) were selected from a tertiary UK referral centre for SSc. All patients met the 1980 American College of Rheumatology (ACR) or 2013 ACR/EULAR (European League Against Rheumatism) classification criteria for systemic sclerosis[10].

To define timing and frequency of SRC in contemporary SSc, we reviewed retrospective clinical and laboratory follow-up data for 2254 patients seen in the centre that included 134 SRC episodes (see Results below). Working on the assumption that ARA-positive patients who reached 60 months of follow-up without SRC were "SRC negative", a group of 99 ARA-positive patients with at least five years' follow-up data was assembled. The cohort was approximately evenly split for presence of SRC; 48 were SRC positive and 51 were SRC negative (see **Table 1**). All patients gave prior consent and the local Research Ethics Committee approved the study (NRES reference: 6398). Gender, ethnicity and other clinical features were typical of the whole cohort of diffuse cutaneous SSc cases. Only European ancestry patients were included for genetic analysis and 4 European ancestry ARA positive SRC patients in the overall cohort were excluded.

Genetic analysis

Genotyping of the Royal Free cohort was performed using the Illumina HumanOmniExpress bead array chip at the UCL genomics centre. All data underwent quality control checks for Hardy-Weinberg equilibrium and genotyping rate in PLINK v1.07 [11].

After filtering of single nucleotide polymorphisms (SNPs), a case-control logistic regression was performed in PLINK, comparing patients with and without SRC to determine genetic signature difference between the two groups of patients. Further statistical analysis was performed in R v3.4.1 [12]. There was no imputation of ungenotyped variants.

The top 9 autosomal SNPs with $p < 1.2 \times 10\text{-}4$ were selected for genetic validation in an independent population. In line with other recent studies using small samples the level of significance used to select those SNPS for replication in the second cohort was set well below conventional GWAS criteria as it was considered that SNPS may still be relevant in the context of the enriched cohort design of this work and the plan for additional functional validation of potential candidate proteins by immunostaining of SRC biopsy specimens. All subjects were of North European ancestry and were not related. In the validation cohort genotyping was undertaken using TaqMan SNP genotyping assays [13] using primers detailed in Supplementary **Table 1**.

Histological analysis

Histological validation was performed to investigate candidate protein expression in SRC biopsy specimens. Eight SRC biopsy samples were identified from Royal Free patients. None were related and they were independent of the samples used for the genetic analysis in this study. The renal biopsies were performed as part of routine clinical practice and so these have been undertaken 7 to 14 days after onset of SRC once the blood pressure has stabilised and clotting studies are normal to reduce the risk of biopsy. They provide information to confirm diagnosis and for risk stratification. These were compared with 8 normal kidney control samples (donated to the UCL Centre for Nephrology by the UK National Health Service Blood and Transplant). Kidney tissues were assessed with polyclonal Anti-CTNND2 and anti-GPATCH2L IgG antibodies (Abcam, Cambridge, MA). Distribution of CTNND2 and GPATCH2L staining in glomerular, tubular, interstitial and vascular compartments was scored blind by two of the investigators and scores were aggregated. In brief, there are two scores for each compartment—proportion of tissue with positive

staining (0-4) and intensity of staining (0-3). These were multiplied together to give a total score (0-12) for each of the tubules, vessels, glomeruli and interstitium. For localisation within the glomerulus with immunofluorescence, nuclei were counterstained with 4',6-diamidino-2-phenylindole (DAPI) and endothelial cells were stained with anti-Von Willebrand Factor (VWF) antibodies.

Results

Definition of serological risk for SRC

A cohort of 2254 SSc patients with information on antibody specificity was initially analysed. Of those 390 (17.3%) were male and 811 (36%) had the diffuse cutaneous subset of SSc. ARA was positive in 258 (11.5%) of the patients, ACA in 639 (28.4%), ATA in 508 (22.5%), U3RNP in 91 (4%) and PmScl in 102 (4.5%). Cumulative incidence of SRC for the cohort as a whole at 12, 24, 36 and 60 months was 3.3%, 4.3%, 4.8% and 5.6%.

Of all ARA positive subjects, 59 (22.9%) had developed SRC. Cumulative incidence of SRC at 12, 24, 36 and 60 months in this subgroup was 12.5%, 18%, 20.4% and 21.4%. **Figure 1** shows cumulative incidence of SRC by antibody subgroup. The high frequency of SRC in ARA positive cases together with a tendency for earlier occurrence compared with other specificities makes this an especially informative group for exploration of genetic susceptibility to SRC.

Genetic analysis

Ninty-nine ARA positive patients with follow-up of at least 5 years (48 with renal crisis and 51 without) were analysed. Characteristics are summarised in **Table 1** including an indication of the frequency of major organ-based complications and concurrent malignancy in the cohort. In addition to their autoantibody homogeneity, the two groups were similar in terms of age, gender and ethnicity. Limited

additional phenotypic data were available for the validation cohort that was selected to be similar in terms of ANA, ethnicity and with robust categorization by SRC status.

GWAS analysis in a UK discovery cohort

Genotype data underwent quality control checks for Hardy-Weinberg equilibrium and genotyping rate in PLINK (HWE $p < 0.001$, genotyping rate > 90%). 2309 SNPs were removed for missingness and 77122 failed MAF filters (MAF < 0.01).

After quality control, 641,489 SNPs were analyzed between the two groups. Genomic inflation factor lambda (λ) was 1.10357 and the Quantile–quantile plot performed. The presence of relatives and/or duplicates was assessed by computing identity-by-descent (IBD) estimation using PLINK.

Results of this Genome-Wide Association Study (GWAS) are illustrated in **Figure 2**. SNPs with a GWAS p value < 3 x 10-5 are annotated on the figure. In contrast with the majority of previous GWAS analyses in SSc, including the only previously documented association with SRC, there was no marked association with the MHC on chromosome 6.

Genetic validation in an independent USA cohort

We selected 9 autosomal SNPs from the top associations for a further validation analysis in 256 ARA-positive subjects. Characteristics of these individuals are summarised in **Table 1B**. There were derived from a cohort used and described in other studies of SSc genetics [14] and the serological testing is validated wherever possible with central testing using ELISA specific for anti-RNA polymerase III autoantibody [15]. The top 9 SNPs from the Royal Free cohort and the findings for each in the validation cohort are summarised in **Table 2.**

Of these 9 SNPs, only one showed replication of significant association in this second cohort, for the replicated SNP. This SNP (rs935332) is in the region of GPATCH2L (G patch domain containing 2-like) on chromosome 14. Polymorphisms within this gene region were previously seen to confer significant risk for diastolic hypertension in other robust GWAS analyses [16,17]. Whilst this replication reached nominal

significance with a nominal uncorrected p-value below 0.05, it should be noted that with 9SNPs tested an appropriate threshold for replication could be set at 0.05/9, which is 0.0055, which no SNP met. For this reason additional histological validation in renal biopsy specimens was sought for GPATCH2L, as outlined below.

Altered tissue expression of GPATCH2L and CTNND2 in scleroderma renal crisis

To explore whether genes associated with the SNP variants identified in the discovery cohort might show altered protein expression in SRC we performed immunostaining in renal biopsy specimens from SRC and in healthy renal tissue. First, we stained for GPATCH2L (G patch domain containing 2-like) protein to ask whether there was differential expression in SRC biopsies. There was significantly increased expression of GPATCH2L protein in SRC samples, localised mainly to the tubular and vascular endothelial structures. As described in methods, a total biopsy staining score (median, range) was calculated for control (1, 0-4)) and SRC (11, 9-21) which confirmed highly increased expression in SRC (p=0.0009). Using categorical analysis of the tubular staining that was the most striking difference between SRC and control biopsies there was positive staining (total score >1) in all 8 SRC biopsies but only 3/8 controls (p=0.026, Fisher exact test). **Figure 3** shows representative sections of SRC (3A-3D) together with IgG (3E) and normal healthy kidney control (3F) and tabulated quantification for each biopsy specimen is provided in **Supplementary Table 2**. Taken with the genetic association for ARA positive SRC in both discovery and validation cohorts in genetic analysis these results provide strong support for GPATCH2L being relevant to susceptibility and potentially implicated in SRC pathogenesis. This is notable given previous association of SNPS in this region with diastolic hypertension in other studies [17].

Although the validation genetic analysis only reached nominal significance for GPATCH2L, amongst the top associated SNPs from the discovery cohort that were tested, we considered that another candidate SNP, linked to CTNND2 warranted further evaluation in the renal biopsy samples, based upon mechanistic plausibility and genetic data from other studies in SSc [18]. As further evidence of shared

underlying aetiopathogenesis, ARA positivity has previously also been associated with high long term risk of PAH [19].

Blinded scoring of CTNND2 expression in SRC biopsy samples versus normal human kidney controls showed distinct glomerular staining, absent in all 8 control samples, in 5/8 SRC samples (Fisher's Exact p=0.026). Again, total biopsy staining score (median, range) was significantly increased in SRC (12, 1-17) compared with control biopsies (3, 0-8, p= 0.0135). Consensus scores for individual renal biopsy specimens are summarised in **Supplementary Table 2**. Brighfield staining of representative sections is shown in **Figure 4A** with appropriate controls **(4B, 4C)**. Immunofluorescence images showed anti-CTNND2 antibody staining in the glomeruli localising to the capillary loops. Double staining using the endothelial marker vWF showed collocation with CTNND2, strongly suggesting expression in endothelial cells, which would be consistent with microvascular perturbation in the glomeruli of SRC, associated with altered Wnt activity (**Figure 4H. 4I**).

Discussion

This study has used the high risk of SRC in cases of SSc with ARA together with the fundamental observation that SRC is very rare in established SSc more than 3 years from first non-Raynaud's phenomenon symptom, to powerfully enrich a rare disease population [20]. Our approach of using enriched cohorts with high susceptibility to a disease complication in GWAS has some similarity to a more conventional extreme phenotype design, which has also proven successful in rare diseases where careful phenotyping can overcome some of the limitations of a necessarily small sample size. Where an extreme phenotype study examines rare causal variants, our focus is on finding susceptibility alleles among functional genetic variants.

Previous studies have explored the genetic association for ARA and have identified MHC alleles associated with this autoantibody [21]. In addition, cohort studies have

explored associations of the major histocompatibility complex (MHC) with scleroderma renal crisis and identified associations in addition to those associated with ARA [4]. In our GWAS analysis it is notable that there is no Manhattan "peak" at chromosome 6, representing the loci associated with MHC. MHC associations are an almost universal finding in GWAS studies of complex autoimmune disease [22].

Our results using GWAS analysis for common variant SNPs in a discovery cohort from the UK and then testing candidate SNPs in a second independent US validation cohort identified only one SNP that was associated in both studies. Since GPATCH2L showed nominal replication in the USA cohort we went on to explore protein expression. A possible role in pathogenesis was supported by our results showing increased expression of the GPATCH2L protein in SRC. This protein is associated with altered RNA processing and could be a marker of cell perturbation that is known to occur in multiple cellular compartments in SSc and could be relevant to development of hallmark pathologies such as SRC. Overexpression is most clearly observed in renal tubular epithelial structures which are not generally regarded as a primary site of SRC pathology, but protein expression in the tubular epithelium plays a critical role in control of both intravascular volume and arterial tone, so it is certainly plausible that disruption of transcription in this region could be associated with disorders of arterial blood pressure. While no mechanism for the association between GPATCH2L and essential hypertension has yet been demonstrated, a single molecular basis explaining the link with SRC and essential hypertension is feasible and may be elucidated in future work, including exploration of any association with outcomes such as renal recovery or survival which may differ for ARA positive SRC compared with other antibody subsets [23].

As well as defining possible genetic susceptibility to SRC we expected that our enriched cohort design might also suggest possible pathogenetic relevance for genes linked with other SNPs associated with SRC risk in our discovery cohort. Based upon other recent published studies we were especially interested in the SNP rs1859082 (p=0.000029) on chromosome 5, which lies within the CTNND2 or delta 2 catenin (D2C) gene. It is notable that the same gene CTNND2 was associated with PAH in SSc in a recent genetic study using a combination of direct sequencing and association

analysis [18]. This association is mechanistically compelling as CTNND2 is an armadillo-related protein that regulates cell-to-cell adhesion via its interaction with the cadherins [24-27]. This gene is associated with Wnt regulation and has been associated with altered migration and adhesion in cancer cells [28]. A recent study demonstrated that HIF1alpha regulated CTNND2 expression and that this could be important in modulating Wnt signalling [29]. Hypoxia and HIF1alpha signalling have been previously implicated in SSc pathogenesis. The Wnt pathway is increasingly recognised to be associated with fibrosis and reported to be altered in SSc, potentially primed by upregulated TGFbeta signalling [30]. Demonstration of altered delta catenin expression in SRC is notable even if our genetic data do not support a general role in SRC susceptibility across cohorts. There is substantial evidence that δ-catenin, is not only involved in early development, cell-cell adhesion and cell motility in neuronal cells, but it also plays an important role in vascular endothelial cell motility and pathological angiogenesis [31] and in SSc fibrosis [32]. This could be highly relevant to pathogenesis of SRC. Regulation of Rho GTPases as downstream targets of TGFbeta in tissue repair is also supportive of the potential relevance in SSc and preclinical models of systemic sclerosis [33]. A recent clinical trial of a topical Wnt inhibitor in SSc suggested beneficial effects on tissue remodelling and adipogenesis [24]. E-cadherin and Wnt signalling influence the release and differentiation of circulating progenitor cells, including endothelial to mesenchymal transition [35,36]. It is plausible that variants in the CTNND2 gene have a role in the dysregulation of endothelial progenitor cells that is observed in scleroderma vasculopathy[37,38]. Both PAH and SRC are proliferative vasculopthic complications of SSc, albeit in different vascular beds and it is notable that the same common variant was associated with SSc in a study that explored genetic association for PAH in SSc [18].

Our immunofluorescence studies demonstrate CTNND2 in non-nucleated fragments within the glomerular capillary tuft. The typical histopathological finding of scleroderma renal crisis is thrombotic microangiopathy (TMA) with obstructive fibrin thrombi within the glomerular capillaries or arterioles, often containing fragmented red blood cells (RBCs) and platelets [39,40]. A recent in vitro model of TMA

demonstrated cell-to-cell adhesion between RBCs and endothelial cells as opposed to simple mechanical sequestration [41]. CTNND2 would be a potential mediator of such adhesion, and further investigation is justified into the role this protein plays in susceptibility to organ injury in other forms of thrombotic microangiopathy (TMA).

A strength of this study is the enriched phenotype method made possible by the very strong risk of SRC conferred by ARA and by the rarity of SRC more than 5 years after onset of SSc. This offers the possibility of identifying genetic risk of a specific organ complication within a rare disease group and is the first study of this kind to date. Using this method, we have provided candidates for further investigation in the genetic risk of scleroderma renal crisis and evidence that these candidates may be functionally relevant.

However, there are inevitable limitations. Firstly, none of the SNP associations identified in the discovery cohort reached conventional GWAS levels of significance [42], although this level may be considered too conservative for a small hypothesis generating study, such as the present analysis, with potential for independent genetic replication and corroborative histological analysis [43]. This is not surprising given the small sample size examined but makes replication approaches particularly important for interpretations of our results. It also means that our findings much be interpreted with caution, and that it is possible that other SNPs with association were not selected for genetic validation. These associations may emerge in future genetic analyses of SRC. Nevertheless, the novel method used to detect difference within a well-phenotyped subgroup may justify investigation of associations seen at a lower p value, as has been successfully reported in other recent SSc functional genomic studies where weaker genetic associations have been supported by robust histological and functional assays [44]. Second, there were differences in allele frequency between the two cohorts examined for some SNPs that may reflect differences in population genetic structure. Some investigators suggest that population structure analysis of cases and controls with GWAS data do not separate cases and controls as only a few loci of usually modest effect are different in frequency between them. Apparent differences in population structure may be

more likely when testing small numbers of matched cases and controls from a highly enriched disease subset.

It is also possible that bias may be introduced from requiring long term survival of at least 5 years for the non-SRC cohort to be outside the risk period for this complication. However, in other recent studies we have shown that overall survival in ARA positive cases is good and that most mortality arises from SRC and associated complications [45] so we consider that differences independent of SRC risk are unlikely to confound our results. Another limitation is that our findings may only be relevant to SRC associated with ARA. We have previously observed that SRC outcomes may differ according to ANA reactivity, with ARA positive patients having better overall survival and being more likely to discontinue dialysis than non-ARA cases [23]. Perhaps the most important mechanistic limitation is the application of the GWAS approach, using a common variant analysis platform for a disease area in which rare causal variants might be of prime important. In future studies an adaptive approach with extreme phenotyping could identify such variants.

At the time of case selection for the study there was no agreed expert consensus definition of SRC. This has now been proposed [6] and our definition for this study is largely in line with the suggested criteria. Future studies should use the new consensus definition to facilitate cross-study comparison. The overall frequency of SRC is the discovery cohort aligns with other published series from a larger series of patients managed in the same centre [45]

In conclusion, we have demonstrated genetic association with SRC in ARA positive cases of SSc of a common variant SNP previously associated with essential hypertension and demonstrated significantly altered expression of the GPATCH2L protein in SRC biopsies. Our approach has also identified strong overexpression of a CTNND2 in glomerular endothelium of SRC that is absent from healthy kidney samples, supporting a potential role for altered Wnt signalling in SRC pathogenesis.

List of abbreviations

SRC	scleroderma renal crisis
SNP	single nucleotide polymorphism
ARA	anti-RNA polymerase autoantibody
ACA	anti-centromere
ATA	anti-topoisomerase
ANA	anti-nuclear
SSc	systemic sclerosis
MHC	major histocompatibility complex
TGF	transforming growth factor
GWAS	genome wide association study

Declarations

Ethics approval and consent to participate

All patients gave prior consent for genetic analysis and histological examination and the Hampstead Research Ethics Committee approved the study. MREC Ref. 6398

Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

C Denton has received research grants from GlaxoSmithKline, CSF Behring, and Inventiva and consulting fees from Roche/Genentech, Actelion, GlaxoSmithKline, Sanofi Aventis, Inventiva, CSL Behring, Boehringer-Ingelheim, and Bayer. Other authors have no disclosures.

Funding

This work was supported in part by MRC grant MR/K015230/1 with additional support from a research grant from Scleroderma and Raynaud's UK (formerly The Scleroderma Society).

Authors' contributions

CD, CF and SG designed the study. All authors contributed to data collection, analysis and interpretations and were involved in writing the manuscript. ES and CD wrote the first draft and all authors read and approved the final submission.

Acknowledgements

Not applicable.

Authors' information

Not applicable.

References

1. Denton CP, Khanna D. Systemic sclerosis. Lancet 2017;390:1685-1699

2. Morelli S, Barbieri C, Sgreccia A, Ferrante L, Pittoni V, Conti F, et al. Relationship between cutaneous and pulmonary involvement in systemic sclerosis. J Rheumatol 1997;24:81–5.

3. Walker UA, Tyndall A, Czirjak L, Denton C, Farge-Bancel D, Kowal-Bielecka O, et al. Clinical risk assessment of organ manifestations in systemic sclerosis: a report from the EULAR Scleroderma Trials And Research group database. Ann Rheum Dis 2007;66:754–63.

4. Penn H, Howie a J, Kingdon EJ, Bunn CC, Stratton RJ, Black CM, et al. Scleroderma renal crisis: patient characteristics and long-term outcomes. QJM 2007;100:485–94.

5. Nguyen B, Assassi S, Arnett FC, Mayes MD. Association of RNA polymerase III antibodies with scleroderma renal crisis. J Rheumatol 2010;37:1068-69.

6. Butler E-A, Baron M, Fogo AB, Frech T, Ghossein C, Hachulla E, et al. Generation of a core set of items to develop classification criteria for Scleroderma renal crisis using consensus methodology. Arthritis Rheumatol 2019;71:964-971.

7. Lynch BM, Stern EP, Ong V, Harber M, Burns A, Denton CP. UK Scleroderma Study Group (UKSSG) guidelines on the diagnosis and management of scleroderma renal crisis. Clin Exp Rheumatol 2016;5:106–9.

8. Stern EP, Steen VD, Denton CP. Management of Renal Involvement in Scleroderma. Curr Treat Options Rheumatol 2015 8:1–13.

9. Nguyen B, Mayes MD, Arnett FC, del Junco D, Reveille JD, Gonzalez EB, et al. HLA-DRB1*0407 and *1304 are risk factors for scleroderma renal crisis. Arthritis Rheum 2011;63:530–4.

10. van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative. Arthritis Rheum. 2013;65:2737–47.

11. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet. 2007;81:559–75.

12. R Development Core Team. R: A Language and Environment for Statistical Computing . Vienna, Austria; 2017. Available from: https://www.r-project.org/

13. Shen GQ, Abdullah KG, Wang QK.The TaqMan method for SNP genotyping. Methods Mol Biol. 2009;578:293-306.

14. Mayes MD, Bossini-Castillo L, Gorlova O, Martin JE, Zhou X, Chen WV, et al. Immunochip analysis identifies multiple susceptibility loci for systemic sclerosis. Am J Hum Genet. 2014;94:47–61.

15. López-Isac E, Acosta-Herrera M, Kerick M, Assassi S, Satpathy AT, Granja J, Mumbach MR, Beretta L, Simeón CP, Carreira P, Ortego-Centeno N, Castellvi I, Bossini-Castillo L, Carmona FD, Orozco G, Hunzelmann N, Distler JHW, Franke A, Lunardi C, Moroncini G, Gabrielli A, de Vries-Bouwstra J, Wijmenga C, Koeleman BPC, Nordin A, Padyukov L, Hoffmann-Vold AM, Lie B; European Scleroderma Group†, Proudman S, Stevens W, Nikpour M; Australian Scleroderma Interest Group (ASIG), Vyse T, Herrick AL, Worthington J, Denton CP, Allanore Y, Brown MA, Radstake TRDJ, Fonseca C, Chang HY, Mayes MD, Martin J. GWAS for systemic sclerosis identifies multiple risk loci and highlights fibrotic and vasculopathy pathways. Nat Commun. 2019;10:4955.

16. Hoffmann TJ, Ehret GB, Nandakumar P, Ranatunga D, Schaefer C, Kwok PY, Iribarren C, Chakravarti A, Risch N Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. Nat Genet. 2017; 49:54-64.

17. Warren H. R., Evangelou E., Cabrera C. P., et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. Nature Genetics. 2017;49:403–415.

18. Gao L, Emond MJ, Louie T, Cheadle C, Berger AE, Rafaels N, et al. Identification of Rare Variants in *ATP8B4* as a Risk Factor for Systemic Sclerosis by Whole-Exome Sequencing. Arthritis Rheumatol 2016;68:191–200.

19. Nihtyanova SI, Schreiber BE, Ong VH, Rosenberg D, Moinzadeh P, Coghlan JG, et al. Prediction of pulmonary complications and long term survival in systemic sclerosis. Arthritis Rheumatol 2014;66:1625–35.

20. Steen VD. Autoantibodies in systemic sclerosis. Semin Arthritis Rheum 2005;35:35–42.

21. Kuwana M, Pandey JP, Silver RM, Kawakami Y, Kaburaki J. HLA class II alleles in systemic sclerosis patients with anti-RNA polymerase I/III antibody: associations with subunit reactivities. J Rheumatol 2003;30:2392–7.

22. Matzaraki V, Kumar V, Wijmenga C, Zhernakova A. The MHC locus and genetic susceptibility to autoimmune and infectious diseases. Genome Biol 2017;18:76.

23. Lynch B, Penn H, Harvey J, Burns A, Denton CP The Prognosis Of Scleroderma Renal Crisis In RNA-Polymerase III Antibody (ARA) Positive Compared To ARA Negative Patients [abstract]. Arthritis Rheumatol. 2013; 65 (suppl 10).

24. Carnahan RH, Rokas A, Gaucher EA, Reynolds AB. The molecular evolution of the p120-catenin subfamily and its functional associations. PLoS One. 2010;5:e15747.

25. McCrea PD, Gottardi CJ. Beyond β-catenin: prospects for a larger catenin network in the nucleus. Nat Rev Mol Cell Biol 2016 J;17:55–64.

26. Lu Q, Paredes M, Medina M, Zhou J, Cavallo R, Peifer M, et al. delta-catenin, an adhesive junction-associated protein which promotes cell scattering. J Cell

Biol 1999;14:519–32.

27. Lu Q, Aguilar BJ, Li M, Jiang Y, Chen Y-H. Genetic alterations of δ-catenin/NPRAP/Neurojungin (CTNND2): functional implications in complex human diseases. Hum Genet . 2016;135:1107–16.

28. Huang F, Chen J, Wang Z, Lan R, Fu L, Zhang L. δ-Catenin promotes tumorigenesis and metastasis of lung adenocarcinoma. Oncol Rep 2018;39:809-817.

29. Huang F, Chen J, Lan R, Wang Z, Chen R, Lin J, et al. Hypoxia induced δ-Catenin to enhance mice hepatocellular carcinoma progression via Wnt signaling. Exp Cell Res. 2019;374:94–103.

30. Gillespie J, Ross RL, Corinaldesi C, Esteves F, Derrett-Smith E, McDermott MF, et al. Transforming Growth Factor β Activation Primes Canonical Wnt Signaling Through Down-Regulation of Axin-2. Arthritis Rheumatol. 2018;70:932–42.

31. Ghose S, Min Y, Lin PC. δ-Catenin activates Rho GTPase, promotes lymphangiogenesis and growth of tumor metastases. PLoS One. 2015; 10:e0116338.

32. Bei Y, Hua-Huy T, Nicco C, Duong-Quy S, Le-Dong NN, Tiev KP, Chéreau C, Batteux F, Dinh-Xuan AT. RhoA/Rho-kinase activation promotes lung fibrosis in an animal model of systemic sclerosis. Exp Lung Res. 2016;42:44-55.

33. Akhmetshina A, Dees C, Pileckyte M, Szucs G, Spriewald BM, Zwerina J, Distler O, Schett G, Distler JH. Rho-associated kinases are crucial for myofibroblast differentiation and production of extracellular matrix in scleroderma fibroblasts. Arthritis Rheum. 2008;58:2553-64.

34. Lafyatis R, Mantero JC, Gordon J, Kishore N, Carns M, Dittrich H, et al. Inhibition of β-Catenin Signaling in the Skin Rescues Cutaneous Adipogenesis in Systemic Sclerosis: A Randomized, Double-Blind, Placebo-Controlled Trial of C-82. J Invest Dermatol. 2017;137:2473–83.

35. Liu X, McBride J, Zhou Y, Liu Z, Ma J-X. Regulation of endothelial progenitor cell release by Wnt signaling in bone marrow. Invest Ophthalmol Vis Sci . 2013;54:7386–94.

36. Heuberger J, Birchmeier W. Interplay of cadherin-mediated cell adhesion and canonical Wnt signaling. Cold Spring Harb Perspect Biol 2010;2:a002915.

37. Good RB, Gilbane AJ, Trinder SL, Denton CP, Coghlan G, Abraham DJ, et al. Endothelial to Mesenchymal Transition Contributes to Endothelial Dysfunction in Pulmonary Arterial Hypertension. Am J Pathol. 2015;185:1850–8.

38. Good RW, Trinder SL, Denton CP, Abraham DJ, Holmes AM. Dysfunctional Endothelial Progenitor Cells may Contribute to Vasculopathy in Systemic Sclerosis. Eur J Vasc Endovasc Surg 2014;47:694.

39. Lusco MA, Fogo AB, Najafian B, Alpers CE. AJKD Atlas of Renal Pathology: Thrombotic Microangiopathy.Am J Kidney Dis 2016;68:e33-e34.

40. Fisher ER, Rodnan GP. Pathologic observations concerning the kidney in progressive systemic sclerosis. AMA Arch Pathol 1958;65:29–39.

41. Smeets MWJ, Bierings R, Meems H, Mul FPJ, Geerts D, Vlaar APJ, et al. Platelet-independent adhesion of calcium-loaded erythrocytes to von Willebrand factor. Schulz C, editor. PLoS One 2017;12:e0173077.

42. Fadista J, Manning AK, Florez JC, Groop L. The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants. Eur J Hum Genet . 2016;24:1202–5.

43. Li, MX., Yeung, J.M.Y., Cherny, S.S. et al. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. Hum Genet.2012; 131: 747.

44. Dritsoula A, Papaioannou I, Guerra SG, Fonseca C, Martin J, Herrick AL, Abraham DJ, Denton CP, Ponticos M. Molecular Basis for Dysregulated

Activation of NKX2-5 in the Vascular Remodeling of Systemic Sclerosis. Arthritis Rheumatol. 2018;70:920-931.

45. Nihtyanova SI, Sari A, Harvey JC, Leslie A, Derrett-Smith EC, Fonseca C, Ong VH, Denton CP. Using autoantibodies and cutaneous subset to develop outcome-based disease classification in systemic sclerosis. Arthritis Rheumatol. 2019 Nov 4. doi: 10.1002/art.41153

Figure 1: Association of scleroderma renal crisis with autoantibody reactivity
Kaplan-Meier analysis of the cumulative incidence of scleroderma renal crisis according to circulating auto-antibody among 2254 patients in the Royal Free cohort. Number of individuals at risk of SRC is documented at 24-month intervals up to 10 years.

Figure 2: Common variant SNP association with scleroderma renal crisis in ARA positive systemic sclerosis

Manhattan plot of genome-wide association analysis for the occurrence of scleroderma renal crisis among ARA-positive patients in the UK cohort. X-axis shows chromosomal position. Y-axis shows the negative log value of the logistic regression p-value for each single nucleotide polymorphism. SNPs with a GWAS p value < 3×10^{-5} are annotated on the figure.

Figure 3: Expression of GPATCH2L protein in normal human kidney (NHK) and scleroderma renal crisis (SRC) core biopsy samples.

Immunostaining of renal crisis biopsy specimens showed consistently increased staining of GPATCH2L protein compared with health control kidney tissue suggesting that altered expression may result from genetic differences within the ARA positive cohort as supported by discovery and replication genetic analysis. Arrows highlight representative positive staining in glomerular (g), tubular (t) and vascular endothelial (e) structures. Panels A-D show different staining intensity and distribution in 4 representative SRC samples. Panel E is an IgG negative control SRC specimen and panel F shows only very low level of staining that is representative of the control healthy kidney sections. Detailed staining scores for individual biopsies are summarised in **Supplementary Table 1**.

Figure 4: Immunohistochemistry images of normal human kidney (control) and renal crisis core biopsy samples.

Immunoperoxidase demonstrates staining of anti-delta 2 catenin antibodies in renal crisis (A) versus normal tissue (B). Bottom right panel (C) shows IgG control staining of renal crisis kidney. To further delineate glomerular immunofluorescence images were used to demonstrate glomerular endothelial cells (VWF—green) as a background capillary tuft (D, E). Panels F and G show cell nuclei within the glomerulus are demonstrated in conjunction with delta 2 catenin (DAPI—blue + D2C—red). No collocation is seen in healthy kidney (H) but in renal crisis (I) D2C appears to collocate with the capillary endothelium distinct from the cell nuclei (VWF + D2C + DAPI). Detailed staining scores for individual biopsies are summarised in **Supplementary Table 1**.

Supplementary Figure 1

Plotted Q-Q (Quartile-Quartile) values for adjusted p-values (allelic model).

To provide additional confidence in our genetic association data we have plotted observed (y-axis) and expected (x-axis) distribution of association for our GWAS. The linearity of the plot suggests that there are not major confounders in the study population.

Table 1:

A. Clinical and demographic features of the discovery cohort (UK)

	SRC (n=48)	*No SRC during follow up (n=51)*
Male gender	8 (17%)	8 (16%)
Age at SSc onset	50.3 (22-70)	50.1 (20-76)
Age at SRC	50.6 (28-70)	
European ancestry	48 (100%)	48 (100%)
Diffuse skin involvement	37 (77%)	51 (100%)
Pulmonary arterial hypertension	3 (6%)	1 (2%)
Interstitial lung disease	12 (25%)	10 (20%)

B. Characteristics of the validation cohort (USA)

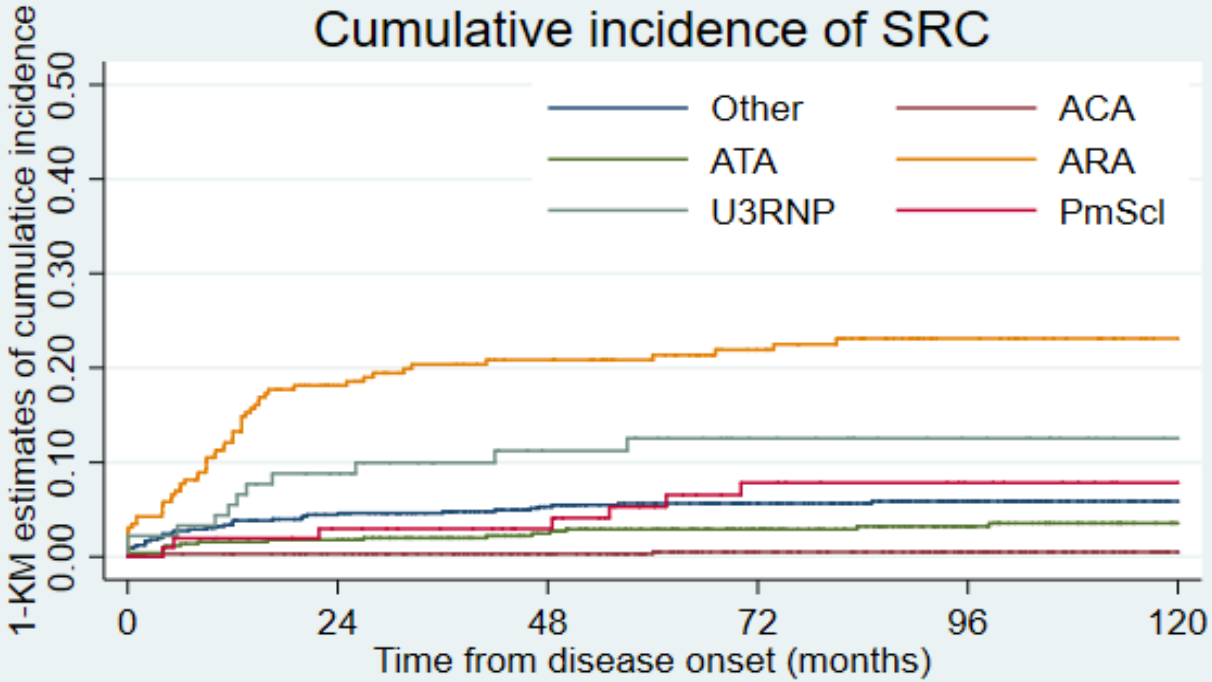
	SRC (n=40)	No SRC during follow up (n=216)
Male gender	7 (18%)	34 (16%)
Age at SRC	57 (44-82)	
European ancestry	40 (100%)	216 (100%)
Diffuse skin involvement	31 (78%)	164 (76%)

Table 2: Candidate SNPs (n=9) associated with SRC in the discovery cohort (UK) and their association with SRC in the validation cohort (USA)

				UK Cohort					USA Cohort			
Chromosome	**SNP**	**Gene**	**BP**	**A1**	**MAF**	**p UNADJ**	**GC**	**OR (CI 0.95)**	**A1**	**MAF**	**P value**	**OR (CI 0.95)**
1	rs2093658	POU2F1	167355192	G	0.490	7.37E-06	1.98E-05	0.24 (0.12-0.45)	G	0.359	0.787	1.07 (0.65-1.76)
2	rs16849716	HECW2	197195099	G	0.049	0.00012	0.0002512	6.11 (2.21-16.85)	G	0.1977	0.962	1.02 (0.56-1.84)
3	rs11708596	C3orf20	14689393	A	0.050	2.28E-05	5.53E-05	7.05(2.58-19.29)	A	0.137	0.223	0.60 (0.26-1.37)
3	rs7643629	IQCJ-SCHIP1	158778092	G	0.324	2.02E-05	4.96E-05	3.52 (1.95-6.35)	G	0.471	0.728	0.918(0.568-1.485)
3	rs2118096	Near to OTOL1	161525058	C	0.240	5.62E-05	0.0001257	3.455 (1.86-6.38)	C	0.368	0.598	0.87 (0.53-1.45)
4	rs10008833	near to EPHA5	66136223	G	0.133	2.05E-05	5.02E-05	4.437 (2.17-9.06)	G	0.168	0.078	1.66 (0.94-2.92)
5	rs1859082	CTNND2	11372743	G	0.272	5.58E-05	0.000125	3.599 (1.9-6.78)	G	0.423	0.9714	0.99 (0.61-1.62)
6	rs2327835	LOC101928354	14570894	A	0.080	7.14E-05	0.0001563	4.987 (2.14-11.57)	A	0.177	0.258	1.40 (0.78-2.50)
14	rs935332	GPATCH2L	76699590	C	0.059	4.92E-05	0.0001114	6.11 (2.38-15.67)	C	0.184	**0.025**	1.86 (1.08-3.20)

A1 Minor allele (in controls)
MAF Minor allele frequency in controls
GC Genomic-control corrected p-values
OR (CI 0.95) Minor allele odds ratio (CI 0.95)

Figure 1



Number at risk

Other	656	602	533	474	409	332
ACA	639	623	584	528	482	416
ATA	508	477	417	349	296	242
ARA	258	189	165	136	108	89
U3RNP	91	80	68	58	50	40
PmScl	102	95	86	71	65	58

Figure 2

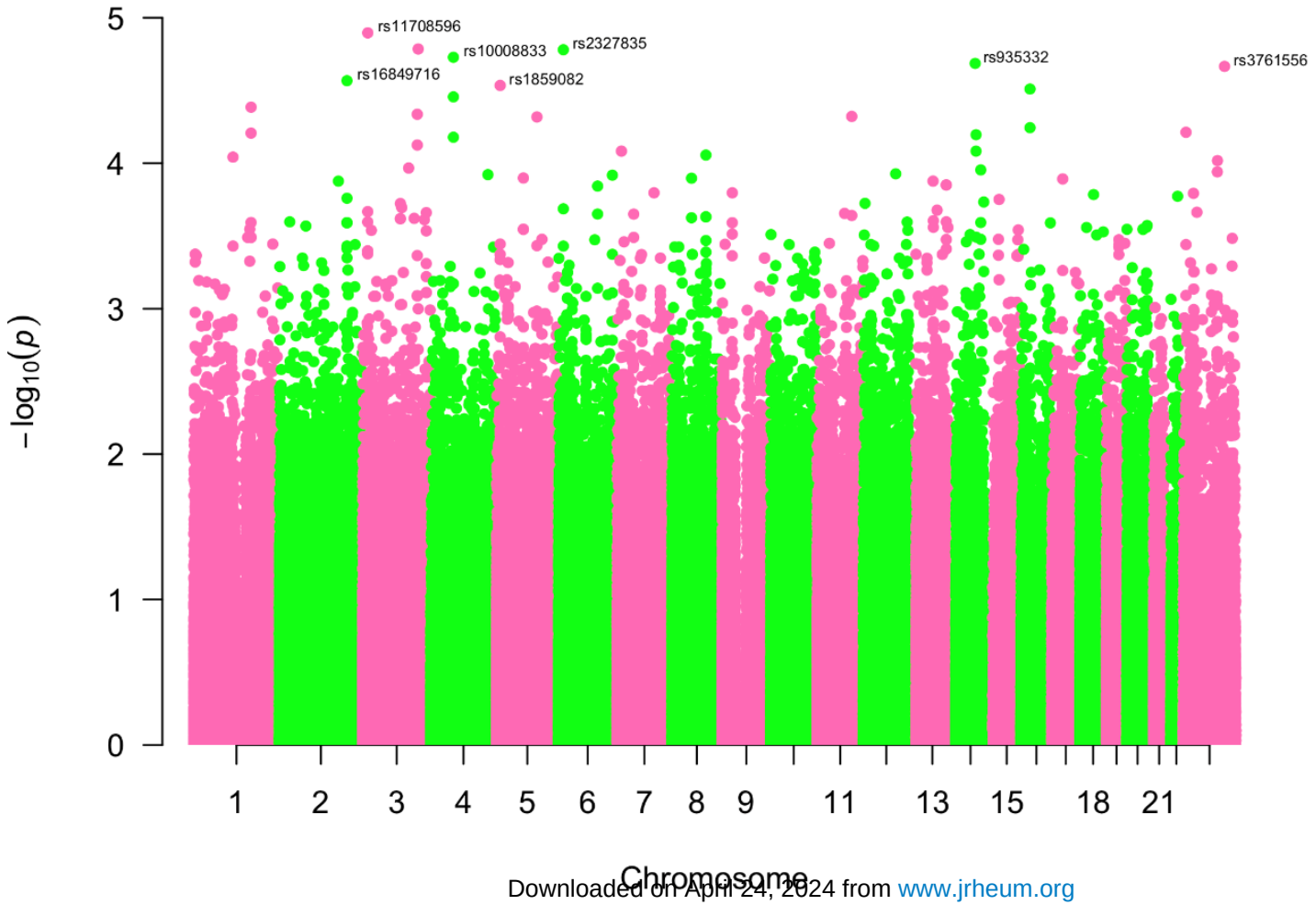


Figure 3

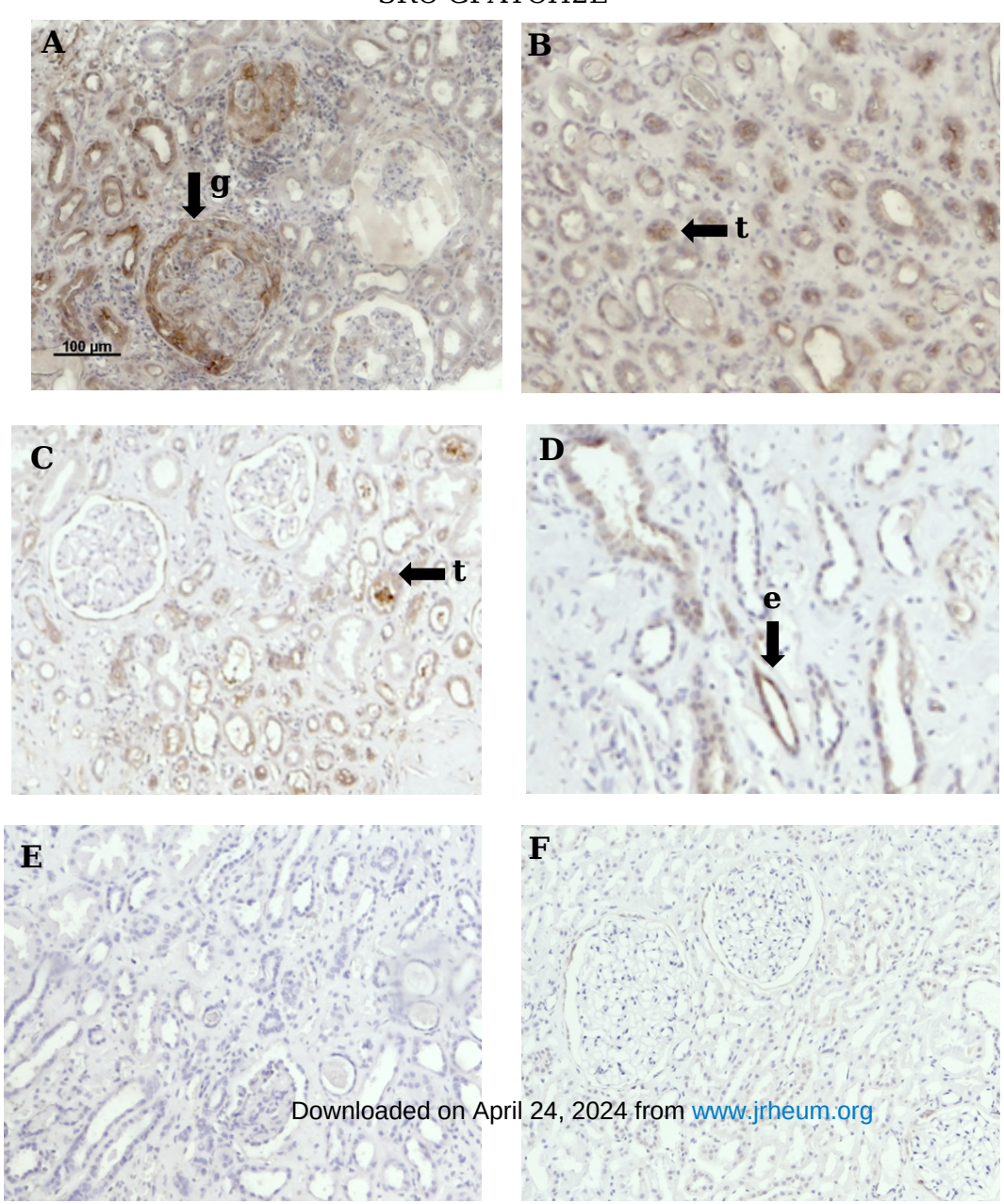


Figure 4