UAB
Universitat Autònoma de Barcelona

Novel strategies for identification and characterisation çof autoantibodies in Systemic Sclerosis

Janire Perurena Prieto

UNIVERSITAT AUTÒNOMA DE BARCELONA

FACULTAT DE CIÈNCIES

DEPARTMENT DE BIOLOGIA CEL·LULAR, FISIOLOGIA I IMMUNOLOGIA

HOSPITAL UNIVERSITARI VALL D'HEBRON

SERVEI D'IMMUNOLOGIA

Novel strategies for identification and characterisation of autoantibodies in Systemic Sclerosis

PROGRAMA DE DOCTORAT EN IMMUNOLOGIA AVANÇADA

Janire Perurena Prieto

Barcelona, 2024

UAB
Universitat Autònoma de Barcelona

Vall d'Hebron
Barcelona Hospital Campus

Abbreviations

A

AAA: ATPase associated with diverse cellular activities

AC: anti-cell

ACE: angiotensin-converting enzyme

ACE-I: angiotensin-converting enzyme-inhibitors

ACR: American College of Rheumatology

AMA: anti-mitochondrial autoantibody

ANA: anti-nuclear autoantibody

ANCA: anti-cytoplasmic antibody

ARMC7: armadillo repeat-containing protein 7

B

BPS: branch point sequence

C

CB: Cajal bodies

CBX5: chromobox protein homolog 5

cDNA: complementary DNA

CENATAC: centrosomal AT-AC splicing factor

CENP-A: centromere protein A

CENP-B: centromere protein B

CENP-C: centromere protein C

CK: creatine kinase

CLF: centrilobular fibrosis

CLIA: chemiluminescence immunoassay

CRIPT: cysteine-rich PDZ-binding protein

CT: computed tomography

CTEPH: chronic thromboembolic pulmonary hypertension

CTLA4: cytotoxic T lymphocyte antigen 4

D

dcSSc: diffuse cutaneous SSc

DFC: dense fibrillar component

DLco: diffusing capacity of carbon monoxide

DM: dermatomyositis

DNA-PK: DNA-dependent protein kinase

DNA-PKcs: DNA-dependent protein kinase catalytic subunit

dsDNA: double stranded DNA

E

ECG: electrocardiogram

ECM: extracellular matrix

eIF: eukaryotic translation initiation factors

EIF2B: eukaryotic initiation factor 2B

ELISA: enzyme-linked immunosorbent assay

EMG: electromyography

EndoMT: endothelial-to-mesenchymal transition

ESRD: end stage renal disease

ET1: endothelin 1

ETS: external transcribed spacers

EULAR: European League Against Rheumatism

EUSTAR: European Scleroderma Trials and Research Group

F

FC: fibrillar centre

Fc: fragment crystallisable

FITC: fluorescein

FVC: forced vital capacity

G

GAR1: H/ACA ribonucleoprotein complex subunit 1

GAVE: gastric antral vascular ectasia

GC: granular component

GER: gastroesophageal reflux

GFR: glomerular filtration rate

GI: gastrointestinal

H

HEp-2: human epithelial type 2

HLA: human leukocyte antigen

HRCT: high resolution computed tomography

hUBF: human upstream binding factor

I

ICAP: international consensus on ANA patterns

IFI16: interferon gamma inducible protein 16

IFN: interferon

IIF: indirect immunofluorescence

IL: interleukin

ILD: interstitial lung disease

ILF: idiopathic lung fibrosis

IP: immunoprecipitation

ITS: internal transcribed spacers

IVTT: in vitro transcription and translation

L

lcSSc: limited cutaneous SSc

LSm: Sm-like

lSSc: limited SSc

M

M3R: muscarinic acetylcholine receptor subtype 3

MCTD: mixed connective tissue disease

miRNA: micro RNA

MNDA: myeloid cell nuclear differentiation antigen

mPAP: mean pulmonary arterial pressure

MRI: magnetic resonance imaging

mRNA: messenger RNA

MRP: mitochondrial RNA processing

MS: mass spectrometry

MTR3: myotubularin-related protein 3

MVP: major vault protein

MW: molecular weight

N

NAF1: nuclear assembly factor 1

ncRNA: non-coding RNA

NHEJ: non-homologous end-joinin

NHP2: H/ACA ribonucleoprotein complex subunit 2

NHP2L1: non-histone chromosome protein 2-like 1

NOP10: H/ACA ribonucleoprotein complex subunit 3

NOP56: nucleolar protein 56

NOP58: nucleolar protein 58

NOR: nucleolar organiser regions

NRHL: nodular regenerative hyperplasia of the liver

NSIP: nonspecific interstitial pneumonia

NVC: nailfold videocapillaroscopy

NVL: nuclear valosin-containing protein-like

O

OR: odds ratio

P

PAGE: polyacrylamide gel electrophoresis

PAH: pulmonary arterial hypertension

PAWP: pulmonary artery wedge pressure

PBC: primary biliary cirrhosis

PD-1: programmed cell death-1

pDC: plasmacytoendritic cell

PDCD7: programmed cell death protein 7

PD-L1: programmed cell death ligand-1

PFT: pulmonary function test

PH: pulmonary hypertension

PM: polymyositis

PMAT: particle-based multi-analyte technologi

Pol: polymerase

POT1: protection of telomeres protein 1

PPT: polypyrimidine tract

PVOD: pulmonary veno-occlusive disease

PVR: pulmonary vascular resistance

PYRIN1: pyrin and HIN domain-containing protein 1

R

RA: rheumatoid arthritis

RBM48: RNA-binding protein 48

rDNA: ribosomal DNA

RHC: right heart catheterisation

RIP-Seq: RNA immunoprecipitation sequencing

RNPC3: RNA-binding region-containing protein 3

ROS: reactive oxygen species

RP: Raynaud's phenomenon

rRNA: ribosomal RNA

RuvBL1: RuvB-like protein 1

RuvBL2: RuvB-like protein 2

S

SAM: S-adenosyl-L-methionine-dependent methyltransferase

SCNM1: sodium channel modifier 1

scRNA-seq: single-cell RNA sequencing

SDS: sodium dodecyl sulphate

SIBO: small intestine bacterial overgrowth

SINE: short interspersed repeated DNA-elements-encoded RNAs

SjS: Sjögren's syndrome

SL1: selective factor 1

SLE: systemic lupus erythematosus

SMA: spinal muscular atrophy

SMN: spinal motor neuron

snoRNA: small nucleolar RNA

snoRNP: small nucleolar ribonucleoprotein

snRNA: small nuclear RNA

SRC: scleroderma renal crisis

SRP: signal recognition protein

ss: splice site

SSA: Sjögren's syndrome antigen A

SSB: Sjögren's syndrome antigen B

SSc: systemic sclerosis

ssSSc: sine scleroderma SSc

SSU: small subunit

T

TCAB1: telomerase cajal body protein 1

TEP1: telomerase-associated protein 1

TERC: telomerase RNA component

TERF1: telomeric repeat-binding factor 1

TERF2: telomeric repeat-binding factor 2

TERT: telomerase reverse transcriptase

TFR: tendon friction rubs

TGF: transforming growth factor

Th: T helper

TXNL4B: thioredoxin-like protein 4B

TIN2: TERF1-interacting nuclear factor 2

TLR: toll like receptor

TMG: trimethylguanosine

TNF: tumour necrosis factor

TPP1: tripeptidyl-peptidase 1

Treg: T regulatory

tRNA: transfer RNA

U

U snRNA: uridine-rich small nuclear RNA

U snRNP: uridine-rich small nuclear ribonucleoprotein

UCTD: undifferentiated connective tissue disease

UIP: usual interstitial pneumonia

V

VPARP: vault poly(ADP-ribose)polymerase

W

WB: western blot

WGCNA: weighted correlation network analysis

Z

ZCRB1: zinc finger CCHC-type and RNA-binding motif containing protein 1

ZMAT5: zinc finger matrin-type protein 5

Contents

Figures

Tables

1. INTRODUCTION

1.1. History, terminology and definition

The earliest reports of scleroderma can be found in the notes of Hippocrates (460-370 BC), who described a specific "thickening of the skin" in some of his patients. However, the word "scleroderma" was first introduced into medical terminology in 1836 by Giovambattista Fantonetti, who used this word derived from the Greek terms "sklērós" (hard) and "derma" (skin) to describe his patient's skin lesions. In the last decades of the 19th century, many cases of patients with the same skin thickening were reported. Some of these patients presented severe internal organ involvement, and a systemic form of the disease was recognised [1, 2].

Scleroderma is used as an umbrella term to encompass both morphoea (localised scleroderma) and systemic scleroderma or systemic sclerosis (SSc). Morphoea refers to sclerotic skin disease limited to the skin and underlying connective tissues, while SSc is characterised by internal organ involvement in addition to skin lesions [2, 3]. SSc is a systemic autoimmune disease with a multifaceted aetiology characterised by inflammatory and fibrotic processes, principally affecting the skin, microcirculation, and internal organs [4].

1.2. Epidemiology, prevalence and incidence

It has been estimated that the overall prevalence of SSc ranges from 15.1 to 20.5 per 100,000 individuals, with an incidence of 1.1-1.9 per 100,000 person-years. Nevertheless, significant variability can be observed in prevalence and incidence reports of SSc. This variability can be attributed to the following factors:

- Difficulties related to the classification of patients due to the rarity of the disease, clinical heterogeneous manifestations, and application of different classification criteria. Indeed, classification criteria have significantly evolved over time, enabling the inclusion of an increasing number of patients in the studies, which, together with the current better overall survival of SSc patients, has resulted in recent studies reporting higher estimates than older ones [5].
- Significant impact of ethnicity leading to regional variations. Prevalence and incidence of SSc in Asia have been reported to be significantly lower than in other

regions of the world. On the other hand, it has been observed that SSc is more common in North and South America than in Europe, with prevalences of 25.9, 24.8 and 14.8 per 100,000 individuals, respectively. However, a north-to-south gradient where higher latitude indicates lower occurrence seems to be present in Europe. In fact, in a study conducted in northwestern Spain, a prevalence of 27.6 per 100,000 individuals was reported, more similar to the prevalence observed in America [5, 6].

- SSc has consistently been described as more common in women than men, and therefore, when stratified by sex, prevalence and incidence are significantly different. A women-to-men ratio of almost 5:1 is noticed in most studies, with a prevalence and incidence of 28.0 per 100,000 women and 2.3 per 100,000 women-years, and 6.0 per 100,000 men and 0.5 per 100,000 men-years [5].

1.3. Etiopathogenesis

Although rare, family history is one of the highest risk factors for SSc. An increased risk of developing disease of 15 to 19-fold and 13 to 15-fold has been described in siblings and first-degree relatives, respectively. Racial differences in disease prevalence and clinical manifestations further support the involvement of genetic factors in SSc pathogenesis. On the other hand, concordance for SSc is low in twins (4.7%) and similar in monozygotic and dizygotic twins (4.2% vs. 5.6%). However, concordance for anti-nuclear autoantibodies (ANAs) presence is significantly higher in the healthy monozygotic twin than in the healthy dizygotic twin of a SSc patient (95% vs. 60%). These data suggest that genetic factors are linked to autoimmunity, increasing the susceptibility to SSc, but are not enough to develop clinically definite SSc. In line with this idea, most of the susceptibility genes for SSc are human leukocyte antigen (HLA) haplotypes and non-HLA genes related to immunity and inflammation, which are shared by other systemic autoimmune diseases such as rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE). Therefore, it is widely believed that SSc develops due to a triggering event in an individual with a “permissive” genetic background. Interestingly, although SSc is characterised by vasculopathy and tissue fibrosis, few genes prominently involved in these processes have been identified so far as possible risk factors, suggesting that the aberrant immune response to a trigger could be the first physiopathological step on SSc [7–11].

Different environmental factors have been identified as possible triggers of SSc. Pathogens are considered one of the most important triggers of SSc; in particular, Human Cytomegalovirus, Human Herpesvirus-6, Parvovirus B19, *Toxoplasma gondii* and

Helicobacter pylori have been linked to the development of SSc. Moreover, it has been shown that gut and skin microbiomes are altered in these patients, making it possible for commensal bacteria to act as triggers. On the other hand, during the last decades, several chemical agents have been suggested as potential etiologic agents of SSc. Organic solvents such as epoxy resins, asbestos, welding fumes, pesticides, cocaine, hair dyes, heavy metals, silicone breast implants, particulate material, and silica have been linked to SSc development. Silica dust was one of the first proposed triggers for developing SSc, as it was associated with the so-called "Erasmus syndrome", characterised by a higher incidence in men with long-term exposure to silica dust and anti-Scl70 autoantibodies. Although not strictly environmental, the presence of tumours and precisely the immune response against them have also been suggested as possible triggers for SSc onset, especially in SSc patients with anti-RNApol III autoantibodies. Additionally, as in other autoimmune diseases, foetal-maternal and maternal-foetal microchimerisms have also been proposed as a triggering event for SSc. In these cases, low levels of foetal cells are harboured in maternal blood, or low levels of maternal cells are harboured in foetal blood. These cells can potentially cause host versus graft and graft versus host reactions, which could trigger autoimmunity. Finally, as with other autoimmune diseases, SSc development has been associated with vitamin D deficiency, which performs various immunomodulatory, anti-inflammatory, antioxidant and anti-fibrotic actions [7, 10, 12, 13].

However, genetics and environmental agents do not fully explain SSc heterogeneity. In this context, epigenetic modifications are considered a unique crossroad between genetics and environmental factors. Abnormalities in DNA methylation of genes associated with fibrosis and collagen production, increased interferon (IFN) production and altered IFN response, and impaired angiogenesis have been reported. Furthermore, like many other autoimmune diseases, SSc presents a striking female predominance, as previously discussed in section 1.2., and due to the high number of immune-related genes located on the X chromosome, it has also been suggested that this sex bias could also be caused by X-linked epigenetic modifications [9].

1.4. Physiopathology

Due to the complexity of the disease and the lack of an animal model that replicates human SSc faithfully, the pathophysiology of the disease has yet to be completely elucidated. Anyhow, it has been long known that it is a progressive self-amplifying condition in which the

clinical and pathologic manifestations of the disease are mainly produced by three distinct processes [7, 14]:

1) Innate and adaptive immune system abnormalities that lead to the production of autoantibodies, cytokines and cell-mediated autoimmunity.
2) Fibroproliferative vasculopathy of small vessels.
3) Fibroblast dysfunction that leads to excessive production of extracellular matrix (ECM) components that accumulate in skin, blood vessels, and internal organs.

1.4.1. Vasculopathy

As already discussed, since most SSc susceptibility genes are related to immunity, the aberrant activation of the immune system after exposure to certain environmental factors seems to be the first step of SSc development. Consistent with this notion, SSc-related autoantibodies are already present before the first clinical symptoms associated with SSc. Moreover, it has been demonstrated that fibrosis and vasculopathy can be reversed in SSc patients after autologous hematopoietic stem cell transplantation, which is thought to reset the immune system and eliminate the autoreactive immune cells [7, 8].

Nevertheless, histologically, the first manifestation of the disease presents as an alteration of the endothelium of small blood vessels (capillaries and arterioles). At early stages of the disease, swelling, vacuolisation, loss of membrane-bound storage vesicles and apoptosis of endothelial cells, and large gaps between them can be observed. The gaps between endothelial cells cause capillary dilation, with the related leakage of red blood cells that clinically may present as telangiectasias. Some other vascular structural alterations, such as capillary enlargement and capillary loss, can also be seen in later stages. Further morphologic changes in the vessel wall are also observed, including fibrosis, smooth muscle cell proliferation, intimal and pericyte proliferation, duplications of the basement membrane and accumulation of proteoglycans leading to arteriolar stenosis. Such histological changes can be detected widespread in internal organs (lungs, heart, kidneys and muscles) even without significant clinical symptoms [7, 8, 14].

Therefore, it has been suggested that endothelial cells of small blood vessels are the initial target of the aberrantly activated immune system. Actually, anti-endothelial cell autoantibodies and γδ T cells targeting endothelial cells have been described in SSc. However, the exact mechanism of initial vascular injury remains unknown, and endothelial cells could also be affected directly by environmental factors. Regardless of the initial trigger,

damaged endothelial cells are believed to undergo cell death and cell activation [7, 8, 14, 15].

Activated endothelial cells have been shown to overexpress cell adhesion molecules, chemokines, cytokines and growth factors. The activation of endothelial cells causes an enhanced interaction with circulating leukocytes, with an increased tethering, roiling, firm adhesion, extravasation and tissue infiltration of leukocytes. Moreover, it has been demonstrated that leukocytes from SSc patients also up-regulate cell adhesion molecules, targeting themselves to the endothelium (**Figure 1**). Consequently, lymphohistiocytic inflammatory infiltrates around affected blood vessels and tissues can be visualised by histology in lesional skin of SSc patients at early stages. Nevertheless, vascular changes in SSc are more typical of a vasculopathy than a vasculitic process, given the paucity of inflammation in the vessel wall [7, 8, 14, 15].

Activated endothelial cells also promote the interaction with platelets, while the apoptosis of endothelial cells and gaps between endothelial cells cause direct contact between subendothelial connective tissue and platelets. Both scenarios lead to the activation of platelets and consequent fibrin production and clot formation. Moreover, activated endothelial cells increase the production of vasoconstrictors and decrease the production of vasodilators, which also facilitates platelet aggregation and thrombosis. Furthermore, it has been demonstrated that SSc platelets show an intrinsic enhanced aggregation capacity to various triggers, as they express high levels of adhesion molecules, von Willebrand receptor and release higher quantities of vasoconstrictors, also contributing to the increased risk for thrombosis of SSc patients. Overall, the frequent and sustained alteration of the microvascular tone due to the altered production of vasoconstrictors and vasodilators is a noxious trigger to the endothelial barrier, leading to the opening of the endothelial junctions, inflammatory cells homing, increased microvessel permeability and continuous vascular leak, giving rise to microhemorrhages and local oedema. Moreover, this causes further direct contact of the subendothelial connective tissue with circulating blood elements, resulting also in vascular thrombosis. Additionally, the connective tissue reacts to this insult by promoting fibroblastic and no-fibroblastic stromal proliferation, contributing to the proliferative arteriopathy seen in SSc (**Figure 1**) [7, 8, 14].

Some of these observed alterations also result from the endothelial-to-mesenchymal transition (EndoMT) process. EndoMT is a trans-differentiation process by which endothelial cells lose their characteristic phenotype markers and morphology and acquire mesenchymal/myofibroblast features. EndoMT process seems to be mediated by

vasoconstrictor molecules and growth factors. Moreover, the pathologic endothelium promotes the switch of multipotent mesenchymal stem cells towards a profibrotic myofibroblast phenotype. Although in a physiological state, myofibroblasts are implicated in wound strengthening by extracellular collagen fibre deposition and wound contraction, in SSc, myofibroblasts potentially participate in fibro-proliferative vasculopathy and tissue fibrosis (**Figure 1**). In line with this, vasculopathy could potentially be a critical pathological step bridging the aberrant activation of the immune system and fibrosis [8, 14, 15].

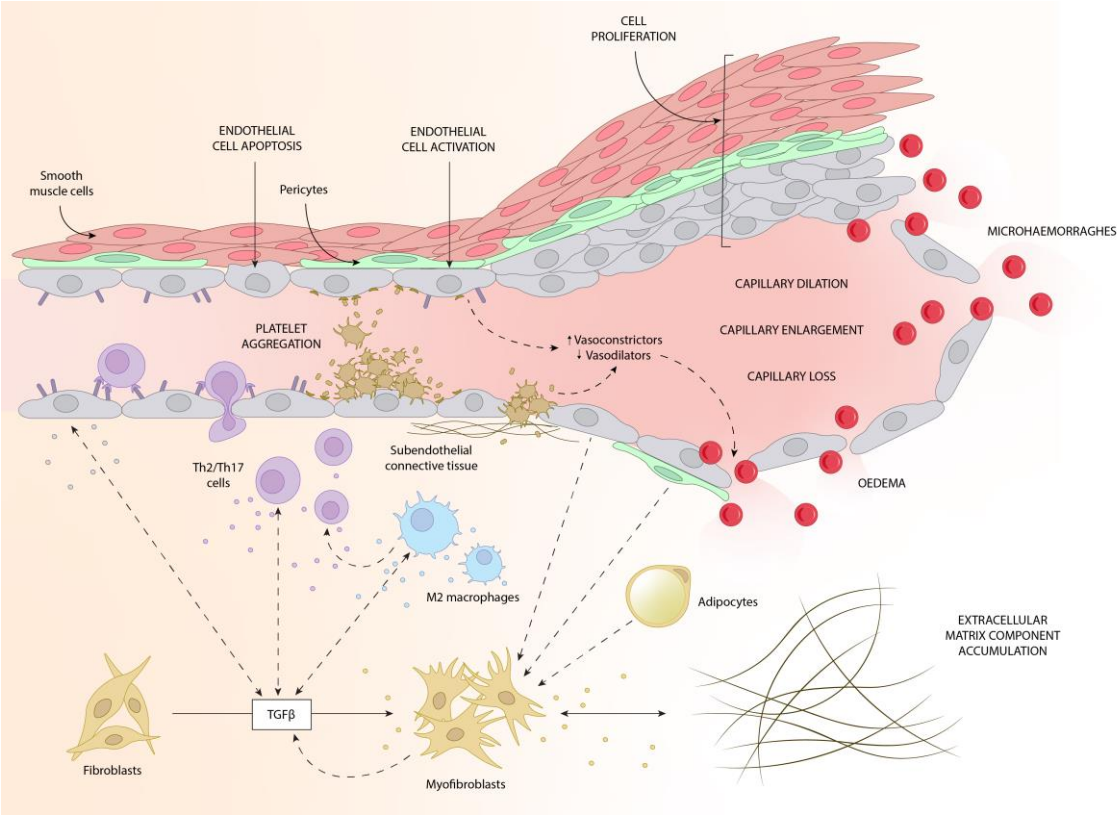


Figure 1. Summary of the physiopathology processes involved in the development of SSc. The initial trigger causes vascular injury, resulting in endothelial cell apoptosis and activation. Activated endothelial cells overexpress molecules that result in the extravasation of leukocytes and interaction with platelets. At the same time, apoptosis of endothelial cells leads to leakage of red blood cells and also exposes the subendothelial connective tissue, further promoting platelet aggregation. Activated endothelial cells and aggregated platelets produce high quantities of vasoconstrictors, resulting in luminal thrombosis which induces inflammation, and activation and proliferation of endothelial cells, pericytes and smooth muscle cells. Moreover, the altered vasoconstrictor production also results in the opening of endothelial junctions and capillary dilation, oedema and microhemorrhages, and further leukocyte homing. The observed lymphohistiocytic infiltrate is mainly composed of Th2 cells and M2 macrophages, which, together with endothelial cell production of ET1, results in an increased production of TGFβ. TGFβ contributes to the differentiation of fibroblasts to myofibroblasts and the transdifferentiation of other cell types, such as endothelial cells, pericytes and adipocytes, into myofibroblasts. Myofibroblasts produce high amounts of ECM, resulting in final tissue fibrosis, and also secrete TGFβ, generating a feed forward loop.

Finally, it has been demonstrated that the microvascular damage in SSc patients is paralleled by an abnormal capability for repairing the injured blood vessels. Blood vessels are generally repaired by angiogenesis, a physiological process through which new blood vessels are generated that is regulated by a balance between pro-angiogenic and anti-angiogenic factors. Tissue ischemia usually leads to the expression of angiogenic growth factors, which causes vasodilatation, proliferation, and migration of endothelial cells and stabilisation of the lumina to form new vessels. Both pro-angiogenic and anti-angiogenic factors are elevated in the plasma of SSc patients, presenting a dysregulated balance. At the same time, endothelial cells respond defectively to pro-angiogenic factors. Less frequently, vasculogenesis can also participate in the repair of injured blood vessels by the incorporation, differentiation, migration and proliferation of bone-marrow-derived progenitor cells to sites of endothelial injury. It has been demonstrated that SSc patients show dysfunction of bone marrow-derived progenitor cells and impaired recruitment to the damaged tissues [7, 8, 14].

1.4.2. Immune system

Altered innate and adaptive immune responses have been demonstrated to play a prominent role in early SSc pathophysiology. As previously discussed, inflammatory cell infiltration around small vessels on affected tissues can be detected by histology. These inflammatory cell infiltrates are mainly seen in the earlier stages of the disease. Still, once interstitial fibrosis sets in, in the later stages of the disease, these infiltrates are difficult to find [8, 14–16].

B cell infiltration is relatively limited in SSc-involved skin, where T cells, macrophages and mast cells are prominent. T cells were the first cells described to infiltrate SSc skin lesions. Furthermore, clonal expansions of T cells were demonstrated, suggesting a T cell response to a widely distributed and persistent antigen in lesional SSc skin sites. T cells are thought to be the principal mediators of the damaging immune responses seen in SSc. In fact, SSc patients show features of chronic graft-versus-host disease, which is mediated mainly by donor T cells. The infiltrating T cells exhibit an altered T helper 1 (Th1) and Th2 balance, skewed toward Th2 responses. While Th1 cells primarily secrete IFNγ, which produces anti-fibrotic effects, Th2 cells are characterised by secretion of interleukin 4 (IL4) and IL13, which have been long linked to fibrosis. The fibrotic changes caused by these two cytokines are mainly mediated by the induction of production of transforming growth factor β (TGFβ), fibroblast proliferation and differentiation to myofibroblasts and ECM component accumulation. Moreover, TGFβ induces IL13 synthesis, which can generate a feedforward

cycle of signalling events that activate the immune system cells and fibroblasts (**Figure 1**). Moreover, as in other autoimmune diseases, an altered Th17 and T regulatory (Treg) balance has also been reported, with Th17 predominance. It seems that Tregs are decreased in number and functional capacity, although there are controversial results, and some reports indicate that Treg number could increase in active disease phases. Moreover, Treg cells have been shown to transform into pathogenic effector circulating Th17 cells, as well as skin resident Th2 cells, that produce pro-inflammatory and pro-fibrotic cytokines, respectively [7, 8, 14–16].

On the other hand, in general, B cells are scarcely seen in SSc tissues, and germinal centres are not prominent. However, numerous lymphoid aggregates composed mainly of B cells are generally evident in the lungs of SSc patients with interstitial lung disease (ILD). Moreover, circulating B cells have been shown to be constitutively activated and produce several autoantibodies targeting nuclear, cytoplasmic, and extracellular autoantigens. Indeed, the presence of anti-nuclear autoantibodies, as discussed in section 1.7., is the hallmark of SSc, and they are observed even prior to the diagnosis of the disease. Although some pathogenic autoantibodies have been described, most autoantibodies are not pathogenic but serve as diagnostic and prognostic biomarkers. The strong association of non-pathogenic autoantibodies with specific clinical manifestations suggest that altered B cell phenotypes correlate with other anomalies that drive the progression of the disease through interaction with other immune and non-immune cells or by sharing the same genetic and epigenetic features. In line with this, it has been demonstrated that B cell depletion with rituximab improves skin and internal organ involvement of SSc, indicating that B cells are involved in the activation of vascular and fibrotic processes in addition to the activation of the immune system [8, 14–16].

Regarding the innate immune system, macrophages are also enriched in perivascular infiltrates since monocytes recruited from the bone marrow migrate into the tissues together with T cells. Once in the tissue, monocytes either differentiate into classically activated (M1) or alternatively activated (M2) macrophages. Generally, M1 macrophages are effector phagocytes that produce pro-inflammatory cytokines like tumour necrosis factor α (TNFα), IL6, and IL1, whereas M2 macrophages produce anti-inflammatory cytokines, mostly IL4, IL13 and IL10. In physiological conditions, during wound healing, or at the peak of the pro-fibrotic late immune response, M2 macrophages partially suppress M1 responses and promote ECM protein synthesis, including pro-fibrotic cytokine release, such as TGFβ. Moreover, M2 macrophages potentiate the anti-inflammatory response by inducing Th2

effector cell activities (Figure 1) [8, 14]. On the other hand, neutrophils and mast cells are also seen in tissue infiltrates. Both types of cells produce pro-fibrotic factors, including TGFβ. Additionally, neutrophils can increase the levels of TGFβ by producing reactive oxygen species (ROS), which trigger the release of TGFβ from its latent ECM-bound state, while mast cells release platelet-activating factors that cause platelets to aggregate, contributing to endothelial damage and the risk of thrombosis [15].

Finally, plasmacytoid dendritic cells (pDCs) have also been found to be enriched around small vessels of SSc lesional skin. pDCs produce large amounts of IFNα through the activation of Toll-Like Receptor 7 (TLR7) and TLR9 that can be stimulated by self-DNA. As endothelial damage provides self-DNA, pDCs produce excessive amounts of IFNα, inducing the senescence of endothelial cells and possibly causing a feedforward loop that promotes the progression of SSc through vascular injury and activation of the immune system. In line with this, clinical trials where recombinant IFNα was used as a treatment for SSc induced the development and progression of the disease, generally due to exacerbation of ILD. Furthermore, when IFNα has been used to treat other diseases, such as multiple sclerosis and chronic hepatitis C virus infection, the onset of typical SSc or SSc-like symptoms has been described [8].

1.4.3. Fibrosis

Fibrosis, characterised by progressive tissue accumulation of ECM components in the skin and multiple organs, is a prominent pathological finding and distinguishing hallmark of clinically overt SSc. In addition to the excessive amount of ECM components, fibrotic tissues are characterised by a high amount of myofibroblasts. These cells not only secrete ECM components and help in the wound contraction but also produce TGFβ and other pro-fibrotic mediators and can originate from activated resident fibroblasts, pericytes, bone marrow-derived fibrocytes, EndoMT, epithelial-mesenchymal transition and adipocyte-to-myofibroblast transdifferentiation (Figure 1). The induction to differentiate into myofibroblast can be considered a final consequence of the previous SSc-specific disease cascade, that is, the pro-fibrotic cellular milieu formed by growth factors, cytokines and other soluble mediators [8, 14, 15].

A key growth factor involved in the activation of fibroblasts and myofibroblast transdifferentiation is TGFβ, a potent inductor of ECM component production. TGFβ is elevated in the skin of patients with early active disease but weak or even undetectable in those with established fibrosis. In the early stages of the disease, TFGβ promotes

inflammation by recruiting leukocytes (regulation of adhesion molecules and generation of a chemokine gradient), activating leukocytes and inducing the production of pro-inflammatory cytokines and other mediators. It seems that in the fibrotic phase, although TGFβ expression is low, SSc fibroblasts are constitutively activated with a pro-fibrotic phenotype similar to that of normal fibroblasts treated with TGFβ, probably due to a self-activation system that relies at least partially on autocrine TGFβ signalling. Moreover, the increased tissue stiffness and reduced elasticity of fibrotic tissues result in mechanical stress mediated by integrins (focal adhesions) that also contributes to the activation of fibroblasts and myofibroblast transdifferentiation and further intensifies the progression of the fibrotic process. Moreover, integrin-mediated tension also promotes the release of active TGFβ from the ECM-bound latent form. Endothelin 1 (ET1) produced by activated endothelial cells also contributes to the activation of fibroblasts using the TGFβ machinery, and activated fibroblasts can also produce ET1 due in part to the action of TGFβ signalling, generating positive feedback. In addition, both molecules are pivotal players in the EndoMT process, inducing more myofibroblast generation (Figure 1) [8, 14, 15]. In addition, chronically activated/aggregated platelets in patients with SSc may also contribute to fibrosis by producing cytokines and growth factors that could induce the scleroderma fibroblast phenotype [7].

Furthermore, SSc fibroblasts present unique intrinsic characteristics that promote the pro-fibrotic status seen in SSc patients:

- Reduced sensitivity towards apoptosis, which, together with the increased rates of differentiation to these cells, leads to an increased number of myofibroblasts in SSc.
- Persistency in a continuous activated state.
- Possession of mechanisms for selectively responding to pro-fibrotic stimuli produced by T cells. Typically, collagen production of fibroblasts is inhibited by Th1 and Th2 cells through membrane-associated IFNγ and TNFα, respectively. SSc fibroblasts' increased collagen production is resistant to both Th1 and Th2 cell-mediated suppression, especially the one exerted by Th2 cells. It has been postulated that SSc fibroblasts produce an antagonist of TNFα receptors in a way that TNFα produced by Th2 cells stop to overcome the pro-fibrotic effect of IL4, also secreted by Th2 cells. Moreover, it has been shown that SSc fibroblasts induce the skewed balance of T cells to a Th2/Th17 phenotype, regulating the transdifferentiation of Tregs into Th2 cells and suppressing IFNγ expression [8, 15].

Briefly, chronic fibroblast activation mediated by different processes and feed forward loops leads to a failure to terminate the regular tissue repair due to the immune system

inflammatory induction, which finally leads to a disturbed ECM deposition, tissue fibrosis and ultimately to organ failure beyond repair [14, 15].

1.4.4. Transcriptomic studies

Multiple transcriptomic studies on different tissues and cells have shown a prominent type I IFN signature, which is defined as a high expression of a group of genes downstream after IFN I stimulation. In this line, it has been demonstrated that polymorphisms in IFN-regulatory factors confer an increased risk for developing SSc, and high levels of IFN I are evident in the blood and skin of a large percentage of SSc patients. However, this signature is not specific to SSc since it has been described in other autoimmune diseases such as SLE and indicates an aberrant immune system activation [14]. On the other hand, based on bulk-tissue transcriptomics of skin biopsies, four “intrinsic SSc subsets” have been identified: fibroproliferative, inflammatory, limited and normal-like. The presence of these subsets suggests that SSc is a heterogeneous disease in which similar but different physiopathological processes could cause the disease depending on the patient. Finally, single-cell RNA sequencing (scRNA-seq) studies have demonstrated the overrepresentation of myofibroblasts expressing transcription factors induced by activation of IFN I and TGFβ pathways. Regarding non-fibroblast skin cell populations, different cells of the innate immune system were found to be present in distinct subsets of the disease, once more indicating that SSc is a heterogeneous disease in which different physiopathological processes could participate [10].

1.5. Clinical characteristics

The aberrant activation of the immune system leading to a proliferative vasculopathy, hyper-coagulation status and excessive fibrosis, affects almost all the organ systems of the human body, including the skin, lung, heart, gastrointestinal tract, liver, kidney and musculoskeletal system, resulting in a very width range of clinical manifestations. In this section, the most frequent clinical manifestations presented by SSc patients are discussed.

1.5.1. Skin involvement

1.5.1.1. Skin fibrosis

Although a small subset (<5%) of patients with established SSc have no clinically detectable cutaneous thickening, classified as having SSc sine scleroderma (ssSSc), skin thickening is

the unifying feature of different SSc subtypes and all other types of scleroderma and scleroderma-like disorders (Figure 2). Nevertheless, wide heterogeneity in the extent of skin involvement exists between patients. Indeed, to date, the subgrouping of SSc patients still mainly relies on the skin fibrosis extension; the classification criteria proposed by *LeRoy et al.* in 1988 [4] proposed two big subsets of SSc, limited cutaneous SSc (lcSSc), in which skin fibrosis is limited to the distal limbs (hands, feet and forearms) and face, and diffuse cutaneous SSc (dcSSc), associated with skin changes that can also affect the trunk and proximal limbs in addition to the distal limbs and face (Figure 2). Moreover, skin involvement is also heterogeneous within the same individual patient over time, as patients may present flares or progressive skin involvement in later disease. In lcSSc, skin fibrosis develops slowly and shows little variability over time, while in dcSSc, there is a more rapid skin fibrosis progression, and posterior flares are more common. It is generally accepted that skin thickness tends to increase in early dcSSc and decrease in late dcSSc, being the time of peak involvement typically 12-18 months after the onset of skin thickening. Notably, the progression of skin fibrosis may be a marker of concurrent internal organ progression, and rapid skin progression is an independent predictor factor of early mortality. Consequently, the measurement of skin thickness is used as a surrogate for disease activity and severity, especially in patients with dcSSc. This has led to the development of different standardised methods for evaluating skin thickness progression, the most used being the modified Rodnan skin score [3, 17–19].

Three phases can be distinguished in the skin thickening process: oedematous, fibrotic/indurative and atrophic. The oedematous phase is considered reversible, but skin thickening becomes irreversible once the atrophy has settled. Usually, the skin fibrosis in SSc starts in distal fingers and toes and progresses proximally. Although the three phases can overlap, in general, at the start, the fingers of the hands become puffy, “puffy fingers”, because of microvascular changes and inflammation, in a symmetrical way. Over time, excessive collagen deposition gives rise to skin thickening and tightening, referred to as sclerodactyly. Skin tightening and thickening are observed distal to the metacarpophalangeal and metatarsophalangeal joints but proximal to the proximal interphalangeal joints. Over time, atrophy occurs, with hide-bound skin (bound to subcutaneous tissues), leading to spindle-shaped fingers surrounded by contracted skin. The limited movement of the skin over joints results in friction rubs, joint contractures and deformities of both large and small joints in some patients, leading to significant disability due to pain, weakness and reduced mobility (see section 1.5.7.1.1.1.) [3, 17].

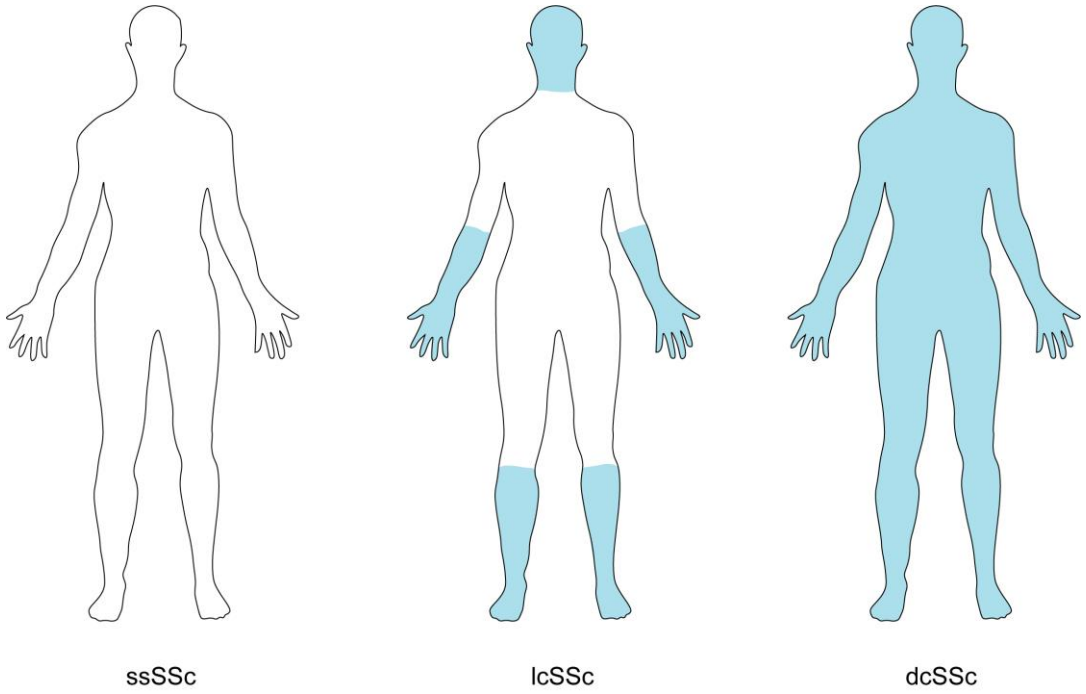


Figure 2. Skin extension of the SSc subsets proposed by LeRoy criteria. The classification criteria of *LeRoy et al.* (1988) proposed two main subsets of SSc, the limited cutaneous SSc (lcSSc) subset, in which skin fibrosis is limited to the distal limbs (hands, feet and forearms) and face, and diffuse cutaneous SSc (dcSSc), associated with skin changes that can also affect the trunk and proximal limbs in addition to the distal limbs and face. There is no clinically detectable cutaneous thickening in less than 5% of established SSc cases, classified as having SSc sine scleroderma (ssSSc).

Skin fibrosis of the fingers is so characteristic of SSc that according to the revised set of criteria adopted by the American College of Rheumatology (ACR) and the European League Against Rheumatism (EULAR) of 2013 for the classification of SSc, patients with skin thickening sparing the fingers are not classified as having SSc (see section 1.6.). In addition, these classification criteria include skin thickening of the fingers of both hands that extends proximal to the metacarpophalangeal joints as a sufficient criterion for classification of SSc to present. However, as mentioned, skin changes can be seen in skin from other parts of the body and have direct clinically significant consequences. For example, skin fibrosis around the mouth can lead to radial perioral furrowing, microstomia, and microcheilia, resulting in impaired speech and mastication, while thoracic fibrosis can cause restricted lung disease. On the other hand, tightening of the skin on the face can cause a pointed, beaked nose, impaired eye opening, and decreased facial expression (amimia). These skin lesions, in addition to physiological problems, produce a significant psychosocial impact [3, 17, 20].

Regarding the physiopathology of skin fibrosis, several studies have demonstrated the upregulation of disease-associated molecules in the epidermis of SSc-involved skin, such as ET1 and TGFβ, suggesting that SSc keratinocytes likely contribute to the activation of dermal fibroblasts and progression of fibrosis. On the other hand, skin fibrosis in SSc is accompanied by loss of the subcutaneous adipose tissue. Moreover, subcutaneous adipose mesenchymal stem cells and mature adipocytes have both been involved in the transdifferentiation into fibroblast-like cells, contributing to skin fibrosis. In fact, a significant proportion of activated myofibroblasts in SSc-involved skin appear to arise from adipocytes adjacent to the deep dermis, suggesting that loss of subcutaneous adipose tissue is a fundamental process for skin fibrosis [8, 15].

1.5.1.2. Cutaneous vasculopathic changes

1.5.1.2.1. Nailfold capillary changes

Vascular system abnormalities can be observed even in the non-lesional skin of SSc patients and in the majority of cases, this is the first sign of SSc. The gold standard method for assessing these abnormalities is the high-resolution nailfold videocapillaroscopy (NVC) with 200-fold magnification. Capillaries typically run perpendicular to the skin surface, but at the nailfold, capillaries run parallel to the skin surface, so their structure can be observed without an invasive technique. The following capillary characteristics are usually evaluated when assessing a NVC image: capillary density, capillary morphology, capillary dimension (width of the apical limb of the capillary), and presence of haemorrhages. In healthy individuals, nailfold capillaries are arranged in orderly “hairpin” rows, while in SSc, giant capillaries (apical diameter ≥50μm), capillaries with “abnormal” shapes, capillary loss, and microhaemorrhages can be seen (Figure 3). Besides the “scleroderma patterns”, various “non-scleroderma patterns” that occur in healthy individuals or other autoimmune diseases are also seen. For example, a decreased capillary density, enlarged capillaries (apical diameter 20-50μm), not typical capillary morphology and microhaemorrhages, alone or in combination, can be observed in these “non-scleroderma patterns”. Regarding atypical shapes of capillaries, to distinguish “scleroderma patterns” from “non-scleroderma patterns”, it is considered that capillaries with a “hairpin” shape, once or twice crossing shape or tortuous shape (afferent and efferent limb undulate but do not cross) are “normal”, on the condition that the tip of the capillary is convex. In contrast, all other shapes are defined as being “abnormal” and SSc-associated [3, 21].

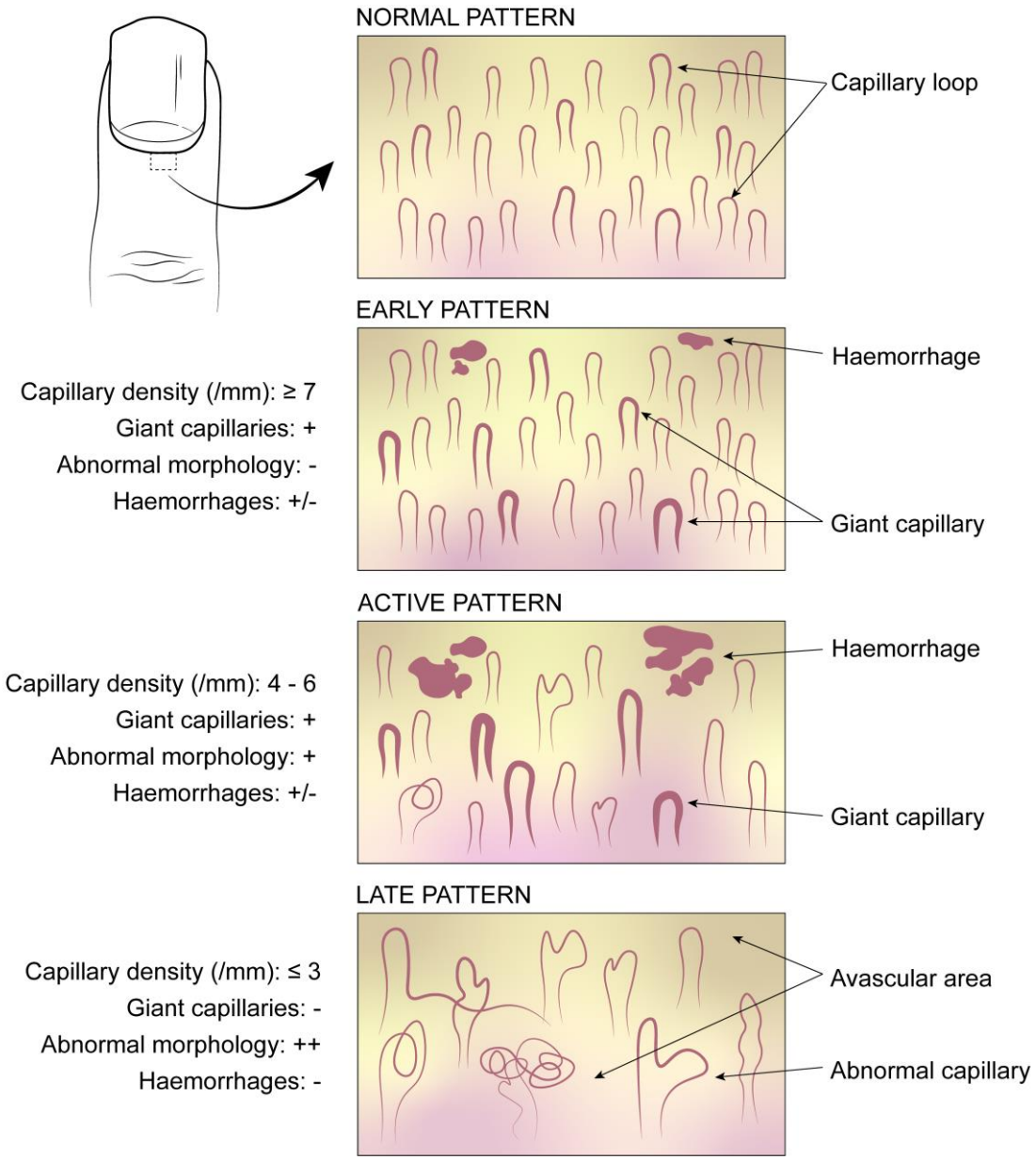


Figure 3. Graphic illustration of NVC patterns observed in SSc patients based on Cutolo's classification. Capillary density, capillary morphology, capillary dimension (width of the apical limb of the capillary), and presence of haemorrhages are evaluated by NVC.

Cutolo et al. subclassified the NVC "scleroderma pattern" into "early", "active" and "late" scleroderma patterns (Figure 3). In the "early" pattern, a small number of giant capillaries are seen alone or in combination with microhaemorrhages but with preserved capillary morphology and without evidence of capillary loss. The active "pattern" is characterised by numerous giant capillaries, sometimes with microhaemorrhages, but always in combination with abnormal capillary shapes and lowered capillary density. Finally, the "late" pattern is defined by a profound loss of capillaries, abnormally shaped capillaries and near absence of giant capillaries and microhaemorrhages. Notably, the three SSc-specific NVC patterns

have been found to correlate with the duration of both Raynaud's phenomenon (RP) and SSc and possibly reflect both the evolution of the disease process and disease severity [3, 21, 22]. In recognition of the diagnostic significance and insight into disease activity provided by nailfold capillary abnormalities, this feature was included as an item in the revised 2013 ACR/EULAR criteria for the classification of SSc (see section 1.6.) [20].

1.5.1.2.2. Raynaud's Phenomenon

RP is a medical condition in which the spasm of small arteries causes episodes of reduced blood flow to end arterioles. A triphasic colour change characterises aberrant digital perfusion: initial pallor (vasospasm), followed by blue/purple cyanosis (deoxygenation of sequestered blood), and then erythema (post-ischemic reactive hyperperfusion). Indeed, at least two colour changes are usually required for the diagnosis. Although cold is considered a trigger, in SSc, these symptoms are often present throughout most of the year, not only in cold seasons. The impaired digital perfusion is initially functional and transient but causes numbness, tingling, and pain. Moreover, repeated bouts can finally cause ischemia-reperfusion injuries, eventually resulting in digital ulcers and, in more severe cases, autoamputation of phalanges or entire fingers. RP affects mainly the fingers but can also affect the toes, ears, lips, and nipples. In addition, a similar phenomenon to RP can also occur in internal organs, with subsequent pathological changes such as impaired renal perfusion and heart damage [3, 14, 15, 18].

Regarding the physiopathology of this clinical manifestation, evidence suggests that RP in SSc results from the subjacent vasculopathy involving all layers of the peripheral blood vessels, partly caused by the dysfunction of the endothelium. Furthermore, it has been shown that there is also a decreased release of vasodilatory neuropeptides from sensory nerves and an increased capacity of vascular smooth muscles to contract, further triggering vasospasm with little provocation [14].

More than 90% of patients with SSc experience RP, which is often the first manifestation of SSc. Time from onset of RP to development of non-RP symptoms frequently varies according to SSc subtype, from months in dcSSc to years in lcSSc. As RP can be the unique evident clinical symptom in the early stages of SSc, it is crucial to distinguish primary RP, which is not associated with any other clinical characteristic, from secondary RP due to SSc. On this basis, LeRoy and Medsger's 2001 revised criteria for the classification of early scleroderma proposed the presence of SSc-specific autoantibodies and scleroderma-type changes on

NVC as the most significant risk factors for developing SSc in patients with RP (see section 1.6.) [3, 14, 23].

1.5.1.2.3. Ulcers

Microvascular pathology, together with RP, gives rise to ischemic ulcers in distal areas of fingers and toes (acral necrosis) in approximately 20-50% of SSc patients, bluntly referred to as “rat bite” ulcers. In the majority of patients, digital ulcers are observed early in the disease (first five years since the onset of disease), especially in dcSSc. Digital ulcers represent a significant burden for SSc patients, causing both functional impairment and lower quality of life. Moreover, due to prolonged time of healing, digital ulcers may be complicated by chronic soft tissue or bone infections and require aggressive medical or surgical treatment, including amputation. On the other hand, non-ischemic ulcers over areas of calcinosis or related to trauma, which typically occur over extensor surfaces of the small joints of the hands that are under increased skin tension due to sclerodactyly can also be seen in SSc patients [3, 14, 24].

1.5.1.2.4. Cutaneous telangiectasias

Telangiectasias are dilated, noninflammatory, superficial postcapillary venules that blanch under applied pressure and are a marker of microvascular abnormalities. Matted or square-shaped telangiectasias are characteristic of SSc and are typically found on the face, lips, oral mucosa, chest, and palms, although they may also be distributed over the upper trunk and lower extremities. Telangiectasias are common in both lcSSc and dcSSc, occurring in 40-70% of SSc patients. Indeed, more than a third of patients present more than ten lesions on the hands or face, and 25% of all observed telangiectasias are larger than 5mm in diameter [3, 24].

1.5.1.3. Pruritus

Approximately 40% of patients with SSc experience pruritus, most commonly noted on the head, back, and dorsal surface of the hands and extremities. This clinical manifestation typically presents in the early inflammatory phase of the disease. Although its pathogenesis has still not been totally elucidated, it has been hypothesised that it may develop due to inflammatory irritation of nerves and fibrotic nerve-ending entrapment. Pruritus is exacerbated by xerosis, greatly impacting the quality of life [3, 24].

1.5.1.4. Dyspigmentation

Dyspigmentation is a common finding in patients with SSc, most commonly affecting the supraclavicular and suprascapular regions and also the scalp, forehead, neck, dorsal hands and extensor forearms. Thermovascular effects, photo-exposure, and friction or pressure have been suggested as contributors to dyspigmentation in SSc [24].

Different dyspigmentation patterns can be detected in SSc patients:

- Diffuse generalized hyperpigmentation that is accentuated by photoexposure.
- Focal vitiligo-like hypopigmentation with perifollicular hyperpigmentation over trauma-prone sites and overlying sclerosis ("salt-and-pepper" dyspigmentation).
- "Streaky" hyperpigmentation overlying superficial cutaneous blood vessels in sites of relative hypopigmentation.
- Reticulate hyperpigmentation.

Although the precise aetiology is unclear, fibroblasts may directly induce melanocyte proliferation as well as melanin distribution and degradation. Cutaneous hyperpigmentation is more commonly observed in dcSSc than lcSSc, with more than 50% of patients presenting this manifestation. Moreover, "salt-and-pepper" dyspigmentation has been associated with more severe prognosis and dcSSc [24].

1.5.1.5. Calcinosis

It has been reported that 20-50% of patients with SSc present intradermal or subcutaneous deposition of insoluble calcified material, also called calcinosis cutis. These calcinotic nodules are composed of calcium hydroxyapatite deposits, surrounded by inflammatory cells, and are generated by a dystrophic calcification process. Dystrophic calcification occurs in the presence of normal calcium and phosphate metabolism and takes place in presumably damaged or devitalized tissues. In this line, dystrophic calcification is noted most often in subcutaneous tissues secondary to trauma or infection and also described in several autoimmune disorders. Although average physiologic tissue concentrations of calcium and phosphate are close to their saturation in healthy individuals, tissue calcification is unusual due to endogenous calcification inhibitors. Therefore, dystrophic calcification occurs in tissues that have been altered in some way to promote calcification, which may include tissue structural damage, hypovascularity and hypoxia, age-related tissue changes, and genetically determined predispositions favouring calcification. Nevertheless, the etiopathogenesis of dystrophic calcinosis is still poorly understood, and the loss of

calcification inhibitors or the appearance of calcification promoters could be involved [25, 26].

In the case of SSc, ischemia may play an essential role in the pathogenesis of calcinosis, through the induction of increased production of oxidative stress products. Indeed, an association between calcinosis and microvascular disease manifestations, such as digital ulcers/pitting scars, acro-osteolysis and late NVC pattern, as well as with macrovascular disease, specifically ulnar artery occlusion, even in the absence of digital ulcers, has been reported. Also, it has been demonstrated that there is reduced perfusion in the superficial skin layers involving calcinotic areas when compared with non-calcinotic areas. On the other hand, dysregulated bone metabolism, including vitamin D deficiency and secondary hyperparathyroidism, osteoporosis, and increased vertebral fractures, have also been associated with calcinosis in SSc, suggesting systemic alterations in the ECM and mineral metabolism. Other mechanisms may include recurrent trauma, as calcification nodules are usually localized at sites of recurrent microtrauma and pressure sites such as the hands, particularly the fingers, forearms, elbows, knees, feet and hips [25–28].

Clinical features of calcinosis cutis may vary significantly among patients with SSc, and the size and location of the deposits determine the associated morbidity. Calcinosis in SSc is usually limited, involving a relatively localised area with small deposits in the skin and subcutaneous tissues, especially over the extensor aspects of the joints and fingertips. In the majority of cases, the deposits remain simply as bothersome subcutaneous lumps, which may rarely cause recurrent local inflammation due to the release of hydroxyapatite crystals, and patients remain asymptomatic. However, calcification of the digits may be associated with nerve pain, and as the most frequently involved sites are the hands, most commonly the thumbs or index of the dominant hand, calcinosis can represent a significant cause of functional dysfunction and a high burden of disability. In addition, ulceration at the site of lesions can occur due to spontaneous extrusion of the calcinotic nodule through the skin or the release of hydroxyapatite crystals into the surrounding tissue. In these cases, a strict follow-up is required to lower the risk of complications such as infections and fistulation, which are not very frequent. Extensive diffuse calcifications (calcinosis cutis universalis) accumulating in the deep dermis, subcutaneous tissue, fascia, and muscles, particularly in regions and areas of repeated trauma, is rare in SSc. In this case, calcinotic nodules can induce muscle atrophy and joint contractures. Finally, “tumoral” or “pseudotumoral” calcinosis, consisting of multiple, large, and often symmetrical calcified masses, may occur in about 3% of SSc patients, more commonly in those with a severe vascular phenotype.

These masses are difficult to treat and may be complicated by ulceration, infections, and nerve compression [25–27].

Traditionally, calcinosis in SSc has been linked to positive anti-centromere autoantibodies and lcSSc, but it has not been confirmed in recent reports. Conversely, a consistent association has been found between calcinosis and anti-PM/Scl autoantibody positivity. Another proposed risk factor is advanced age, as patients with calcinosis cutis are significantly older than those without calcinosis cutis. Nevertheless, this could be explained by the fact that calcinosis is a late complication of the disease, usually presenting more than ten years after the onset of the disease. Finally, calcinosis cutis in SSc patients has recently been linked with proton pump inhibitors used for gastroesophageal reflux treatment [25–27].

1.5.2. Lung involvement

Lung involvement in SSc can be caused by extrapulmonary features, such as respiratory muscle weakness or skin fibrosis of the chest wall, or by intrapulmonary manifestations. The two most important intrapulmonary complications in SSc are ILD and pulmonary arterial hypertension (PAH). Other processes involving the lung parenchyma, such as organising pneumonia, and other forms of pulmonary vasculopathy, such as thromboembolic disease and pulmonary veno-occlusive disease (PVOD), can also develop [29].

1.5.2.1. Assessment of lung involvement

Systematic approaches for screening and assessing patients with SSc for lung involvement are associated with improved long-term survival compared with historic cohorts without regular screening. Therefore, a comprehensive assessment at initial diagnosis followed by a regular, systematic assessment of cardio-respiratory function is recommended for all patients with SSc. This strategy is vital for differential diagnosis, management of lung disease and monitoring the treatment response. The investigation of lung involvement in SSc relies mainly on clinical examination, pulmonary function tests (PFTs), imaging methodologies such as chest radiography, echocardiography or computed tomography (CT) and right heart catheterisation (RHC) [29].

PFTs include spirometry, measures of gas exchange determined by the diffusing capacity for carbon monoxide (DLco) and measures of lung volumes such as total lung capacity (see Annexe 1). PFT often reveals restriction in patients with SSc due to ILD or PAH but also due to extra-pulmonary manifestations (hidebound chest and myopathy). Some patients have an additional obstructive component due to small airway disease. Additionally, in the case

of intrapulmonary manifestations, gas exchange is also reduced. Specifically, patients with ILD usually show a decrease in forced vital capacity (FVC) in proportion to DLco, resulting in a volume-adjusted DLco. In contrast, a disproportionate reduction in DLco suggests the presence of PAH [30].

However, in addition to PFTs, other evaluations are required for the investigation of lung involvement in SSc, as they have been shown to not be sensitive enough at initial screening. This lack of sensitivity is caused in part by wide ranges of normality and high intraindividual variation due to extrapulmonary factors such as oral or thoracic fibrosis, myopathy, fatigue, and cachexia. In fact, CT imaging is currently considered the best tool for detecting ILD, and RHC is required to diagnose PAH definitively. Nevertheless, PFTs should still be performed at baseline and during longitudinal follow-up, as they have repeatedly been shown to be independent predictors of poor outcomes, and dynamic monitoring over a long period can more accurately predict the course of disease [30].

1.5.2.2. Interstitial lung disease

Up to 30-40% of SSc patients develop clinically significant ILD, presenting commonly with dyspnoea (initially only on exertion and eventually at rest), non-productive cough, and overwhelming fatigue. Prevalence of ILD increases during the course of the disease, but its onset is most often within the five years of the first non-RP symptom and seldom more than 15 years after diagnosis of SSc. Nevertheless, a considerable proportion of patients with SSc-ILD are asymptomatic, especially at the initial stages of the disease, and don't present the typical velcro-like crackles on auscultation nor abnormalities on PFTs. It should also be noted that mobility limitation caused by non-pulmonary manifestations of SSc may lead to a lack of recognition of dyspnoea in the early stages of clinically significant ILD.

Given the morbidity and mortality associated with ILD and the inability to reverse the fibrotic process once established, interest in the role of screening for the disease has been considerable. Protocols involving high-resolution computed tomography (HRCT) have been more successful than those using only PFTs, as HRCT has shown a high sensitivity for identifying patients who could develop clinically significant ILD. In fact, interstitial abnormalities are evident by HRCT in up to 80% of patients, a prevalence similar to the one demonstrated by autopsy (up to 90%). Moreover, the absence of ILD on an initial HRCT is a favourable sign, with up to 90% of these individuals remaining disease-free at follow-up. However, as HRCT is associated with increased radiation exposure in a population of primarily middle-aged or young women who could experience the side effects of ionising

radiation, HRCT should be exclusively done at baseline and repeated only when there is a clinically meaningful decline in PFTs or new symptoms that could be attributed to SSc-ILD [30].

Although historically, ILD has been subclassified based on histological patterns observed in lung biopsy samples, high correlation with CT findings has resulted in the use of CT for subclassification of the disease. In fact, nowadays, a lung biopsy is not required to diagnose SSc-ILD and has no bearing on the treatment strategy. The most common subtype in patients with SSc is nonspecific interstitial pneumonia (NSIP), seen in more than 80% of SSc patients by CT. NSIP can be further subdivided into reversible cellular NSIP and irreversible fibrotic NSIP based on the degree of inflammation present in the lung. A minority of patients with SSc present usual interstitial pneumonia (UIP), the defining pattern of idiopathic lung fibrosis (ILF). In ILF, the UIP pattern is associated with much faster progression and worse survival than the NSIP pattern; however, this does not apply to SSc-ILD. In fact, the outcome of SSc-ILD is much better than ILF, with a similar mortality between patients with a NSIP and an UIP pattern. Some patients with SSc can also have other forms of ILD. For example, fibroelastosis in association with pleural abnormalities resembling idiopathic pleuroparenchymal fibroelastosis can also be detected in a minority of SSc cases, a rare form of ILD with a poorer prognosis than other forms of SSc-ILD. Finally, patients with overlapping forms of SSc can exhibit different ILD patterns associated with the overlapping disease. For example, inflammatory myopathies are associated with organising pneumonia, and SLE is associated with shrinking lung syndrome or lymphocytic interstitial pneumonia. In addition to these pulmonary patterns, other common findings of CT in SSc patients include a dilated, patulous, or fluid-filled oesophagus, changes associated with chronic aspiration (see section 1.5.5.), pulmonary artery enlargement, and right ventricular dilation or hypertrophy (see section 1.5.4.), which indicates pulmonary hypertension (PH) (see section 1.5.2.3.) [29, 30].

Although SSc-ILD was traditionally considered slowly progressive with an expected median survival of 15 years, evidence suggests that it might have vastly variable progression rates, with the exact phenotype determined early in the course of the disease. Male sex, active smoking, older age at presentation, presence of arthritis, digital ulcers, PH, progressive skin fibrosis, renal disease, and myocardial fibrosis are also associated with an aggressive form of ILD and early mortality. Furthermore, early development of SSc-ILD, in less than three years after diagnosis, could also be associated with an aggressive clinical course of this disease and is increasingly being recognised [30].

CT is not only useful for diagnosis and classification of the disease but can also be used to determine the extent of the disease in the majority of cases. Patients can be stratified into groups based on whether they have mild or extensive fibrosis (<20% lung involvement or >20% lung involvement, respectively) by HRCT. Extensive fibrosis is associated with a dramatic increase in the odds of mortality. Patients for whom the extent of fibrosis is indeterminable by CT imaging are instead stratified into these groups based on FVC measurements (>70% for mild fibrosis and <70% for extensive fibrosis). This staging system based on PFTs has been shown to be associated with symptoms of dyspnoea and impaired exercise capacity and can also predict an increased risk of considerable future decline and decreased survival. Furthermore, in patients with extensive lung fibrosis, serial PFTs can be used to monitor disease progression and predict mortality. For example, a drop in FVC of >10% at 12 months is a strong predictor of mortality. Additionally, a 5-9% decline in FVC alongside with a 15% fall in DLco at 12 months is also predictive of disease-associated mortality. However, short-term changes (<12 months) should be interpreted with caution because they might not be as reliable as longer trends. Moreover, when assessing a decline in spirometry, factors such as worsening myopathy or increased truncal skin thickness should also be considered [29, 30].

Regarding the physiopathology of this clinical manifestation, the initial event has been proposed to be an injury to the alveolar epithelium or vasculature, or both. Moreover, it has been suggested that aspiration or micro-aspiration could be involved in the initial epithelial injury and damage, as a link between lung fibrosis and oesophageal reflux severity has been found (see section 1.5.5.). Histologically, microvessels with abnormal structures, as well as the thickening of the alveolar septa with a large number of myofibroblasts, are observed in the early stages. The subsequent excessive ECMs deposition results in the replacement of normal pulmonary architecture by scarring. The fibrosis progresses until it extensively damages the vital structures of the lung. Finally, the lung transforms into a contracted fibrous organ lacking alveoli and vasculature, leading to restrictive pulmonary physiology and impaired gas transfer [8, 29–31].

Although, as previously mentioned, overall clinically significant ILD prevalence is 30-40%, it varies greatly depending on the studied cutaneous subtype of the disease. Up to half of patients with dcSSc eventually develop severe ILD, while the frequency of ILD in patients with lcSSc is approximately half that of those with dcSSc. Moreover, autoantibodies are good biomarkers for predicting patients with a higher risk of developing ILD, specifically anti-Scl70, as approximately two-thirds of patients positive for this autoantibody ultimately develop

moderate to severe ILD. Other less common autoantibodies, such as anti-U11/U12 snRNP and anti-Th/To, have also been associated with ILD. In contrast, the presence of anti-centromere or anti-RNApol III autoantibodies seems negatively associated with severe ILD. On the other hand, demographic and clinical factors associated with the presence of SSc-ILD include male sex, African American race, NVC abnormalities, digital ulcers, longer disease duration, and PH. Additionally, significant oesophageal dysmotility and gastroesophageal reflux are associated with the presence of ILD and lower lung function, and therefore, robust evaluation for the presence of these gastrointestinal manifestations is recommended in patients with suspected or confirmed SSc-ILD [29–31].

1.5.2.3. Pulmonary arterial hypertension

PH is characterised by a chronic and progressive increase in the pressure of the pulmonary vascular system, which ultimately leads to a high right ventricular afterload with subsequent dysfunction and right-sided heart failure. PH can be caused by different pathophysiological mechanisms but is defined by the presence of a mean pulmonary arterial pressure (mPAP)≥20mmHg evaluated by resting RHC. PH can be subclassified as pre- and post-capillary; pre-capillary PH is characterised by a pulmonary artery wedge pressure (PAWP)≤15mmHg and a pulmonary vascular resistance (PVR)≥2WU, while post-capillary PH is defined by a PAWP>15mmHg with normal PVR≤2WU. In the case of presenting PAWP>15mmHg and high PVR≥2WU, a combined pre- and post-capillary PH is diagnosed (see Annexe 2) [32, 33].

The World Health Organization (WHO) classification stratifies PH into five categories. Group 1: PAH characterised by abnormalities of pulmonary arterioles; Group 2: PH associated with left heart disease; Group 3: PH due to lung disease and hypoxia; Group 4: PH due to pulmonary artery obstructions (chronic thromboembolic pulmonary hypertension or CTEPH); Group 5: PH with unclear and multifactorial mechanisms, which do not fit into the other four categories. From a hemodynamic standpoint, patients in Group 1, Group 3 and Group 4 present a pre-capillary form of PH, as all the conditions included in these groups originate from problems of the pulmonary arteries, while those in Group 2 may present both an isolated post-capillary PH or a combined pre- and post-capillary PH, given by an increase in PVR [8, 34].

In the context of SSc, PH aetiology can be highly heterogeneous. PH in SSc may be caused by a primary vasculopathy and the remodelling of the small- and medium-sized pulmonary arteries, resulting in vasoconstriction, fibrosis, intraluminal micro-thrombosis and,

consequently, in a progressive increase in the PVR (Group 1). Nevertheless, PH in SSc can also be caused by left heart dysfunction (Group 2), due to ILD (Group 3), or a consequence of the increased risk of pulmonary thromboembolic disease (Group 4), with a higher risk of the last group in patients with anti-phospholipid autoantibodies. Furthermore, SSc patients are at risk of developing PVOD, a rare and yet underdiagnosed form of PH classified as Group 1 and characterised by the obstructive intimal fibrosis of the small veins and venules of the pulmonary circulation, a poor response to pulmonary vasodilators, and an extremely poor prognosis. Moreover, the overlap of different forms of PH can occur within the same SSc patient, and it has been shown that PH aetiology within the same patient can change during the disease's natural history [29, 33, 35].

The incidence rate of PH of any form is approximately 1-2% per year in patients with SSc. Therefore, the frequency of PH varies across different reports depending on the duration of follow-up [29]. However, a meta-analysis of five European studies showed that the global prevalence of PH is 7% in SSc patients. The distribution of those patients showed that almost 80% had pre-capillary PH, where nearly 2/3 belonged to Group 1 as PAH (51% of total patients), and 1/3 belonged to Group 3 as PH secondary to ILD (26% of total patients). In comparison, 21% had post-capillary PH due to left-sided heart disease, and 2% had PH due to PVOD. The prevalence of CTEPH was not assessed. The DETECT study showed a similar proportion between SSc-PAH and other types of PH in SSc [33].

Overall, the presence of PH, despite the specific aetiology, significantly affects the prognosis of SSc patients. Indeed, a meta-analysis showed that, among SSc patients with PH, the global 3-year survival rate was 52%. Given the increased risk of developing PH in SSc patients and its strong impact on the mortality of those patients, it is recommended to refer patients for early and periodic screening, as early treatment could improve the long-term outcome. Moreover, as SSc-PH patients tend to be asymptomatic in the early course of the disease or present with unspecific symptoms, such as dyspnoea on exertion and fatigue, it is of utmost importance to implement those screening methods for detecting this severe manifestation [33, 35, 36]. RHC is the gold standard test for PH definitive diagnosis, as it represents the only approach to directly measure the pressures of pulmonary vascular circulation (mPAP and PAWP), as well as cardiac output, both of which are necessary to calculate PVR (see Annexe 2). However, as RHC is an invasive procedure, in the past years, approaches such as the DETECT algorithm to assess the risk of PH and evaluate which patients should undergo RHC have been developed. These algorithms generally integrate

PFTs, serum biomarkers, electrocardiogram (ECG) and echocardiography parameters, and clinical manifestations [29].

Once the suspect of PH is addressed by RHC, a careful phenotyping of PH in SSc needs to be properly and promptly addressed in order to provide the most appropriate treatment, which is very different for each specific underlying condition. RHC alone can only distinguish the pre- and post-capillary components, so other approaches are needed to differentiate between the different PH groups. In fact, for the diagnosis of PAH, other causes of PH should be ruled out, which can be challenging in SSc. HRCT is extremely useful for this purpose, as PH associated with ILD and PVOD can be identified. On the other hand, a ventilation/perfusion scan can be used to exclude the possible presence of CTEPH [33, 36].

Regarding the physiopathology of PH, different groups originate due to distinct causes and processes, as previously stated. In the case of PAH, vascular injury, including pulmonary endothelial cell injury and endothelial dysfunction, might be a key component in its development, as vascular injury can lead to an aberrant fibroproliferative or aberrant repair process that also results in an obliterative pulmonary vasculopathy, which histologically is characterised by the fibrotic occlusion of pulmonary arterioles [8, 29]. The differences in the aetiologic processes are also reflected by the distribution of distinct PH groups in specific SSc subsets. SSc-PAH is more commonly found in patients with the limited cutaneous form of the disease and often has been associated with the presence of anti-centromere autoantibodies. However, no such association has been confirmed in the last reports. Moreover, SSc-PAH has also been associated with extensive telangiectasias, longer disease duration, older age (>60 years old) and female gender. On the other hand, obviously, SSc-ILD is more common in patients with the diffuse cutaneous form of the disease with anti-Scl70, anti-U11/U12 snRNP and anti-Th/To autoantibodies [36].

1.5.3. Renal involvement

Kidney disease in SSc is common, as demonstrated by autopsy studies, in which renal histological involvement due to blood vessel abnormalities is present in 60-80% of patients. Nevertheless, renal involvement in SSc remains subclinical until the late stages of the disease, with up to 50% of asymptomatic patients presenting biomarkers indicative of renal dysfunction, such as proteinuria, increased creatinine concentration, or hypertension. It seems that renal involvement in SSc is primarily characterised by vascular damage and glomerular hypofiltration. Moreover, there is an association between renal dysfunction and PH, reflecting that PH and right heart failure could contribute to renal dysfunction through

fluid retention and neuroendocrine activation. Although in the majority of cases, kidney involvement is mild, in a minority of patients, it can be very severe and a cause of death if untreated. Scleroderma renal crisis (SRC) is the most serious complication and the most studied specific form of renal involvement in SSc. Beyond SRC, some patients may present anti-neutrophil cytoplasmic antibody (ANCA)-associated vasculitis, isolated reduced glomerular filtration rate (GFR) and high intrarenal arterial stiffness [37, 38].

1.5.3.1. SRC

SRC typically presents with an acute onset of moderate to marked hypertension and oliguric kidney failure. On presentation, 90% of patients with SRC consistently have blood pressure levels exceeding 150/90mmHg and decreased renal function (≥30% reduction in estimated GFR). Hypertension is frequently accompanied by manifestations of malignant hypertension, such as hypertensive retinopathy, encephalopathy, seizures, fever and general malaise, and it can cause congestive acute heart failure, pulmonary oedema, and renal failure. The onset of kidney failure is acute and usually in the absence of significant previous kidney involvement. Due to kidney failure, there is a substantial and rapid increase in serum creatinine concentration. High renin levels and endothelial cell perturbation markers (soluble adhesion molecules) have also been observed in serum. The urine sediment is normal in most cases, and although mild proteinuria may be present, it usually is less than 1g/day. However, in about half of the cases, haematuria and granular cast may be visible on the urinalysis, as SRC may concur with thrombotic microangiopathy. In these cases, laboratory findings of haemolytic anaemia and thrombocytopenia may also be present, as elevated lactate dehydrogenase, low or absent haptoglobin, reticulocytosis and presence of schistocytes together with a negative direct coomb test [8, 37, 38].

Although the pathogenesis of SRC is not yet fully elucidated, it is hypothesised that both endothelial dysfunction and vascular damage result in vascular thrombosis, vasospasm and narrowing of renal arterioles. In contrast to PAH, characterised by slowly progressive vasculopathy, the vascular changes in renal scleroderma usually develop rapidly due to higher values of systemic blood pressure compared to pulmonary pressure. Decreased renal perfusion leads to juxtaglomerular apparatus hyperplasia and renin secretion, causing vasoconstriction and renal ischemia that results in accelerated hypertension and progressive renal injury. The critical rise in blood pressure not only causes further damage to kidney blood vessels but also initiates a feedforward cycle, eventually leading to malignant hypertension and acute renal failure (**Figure 4**) [37, 38].

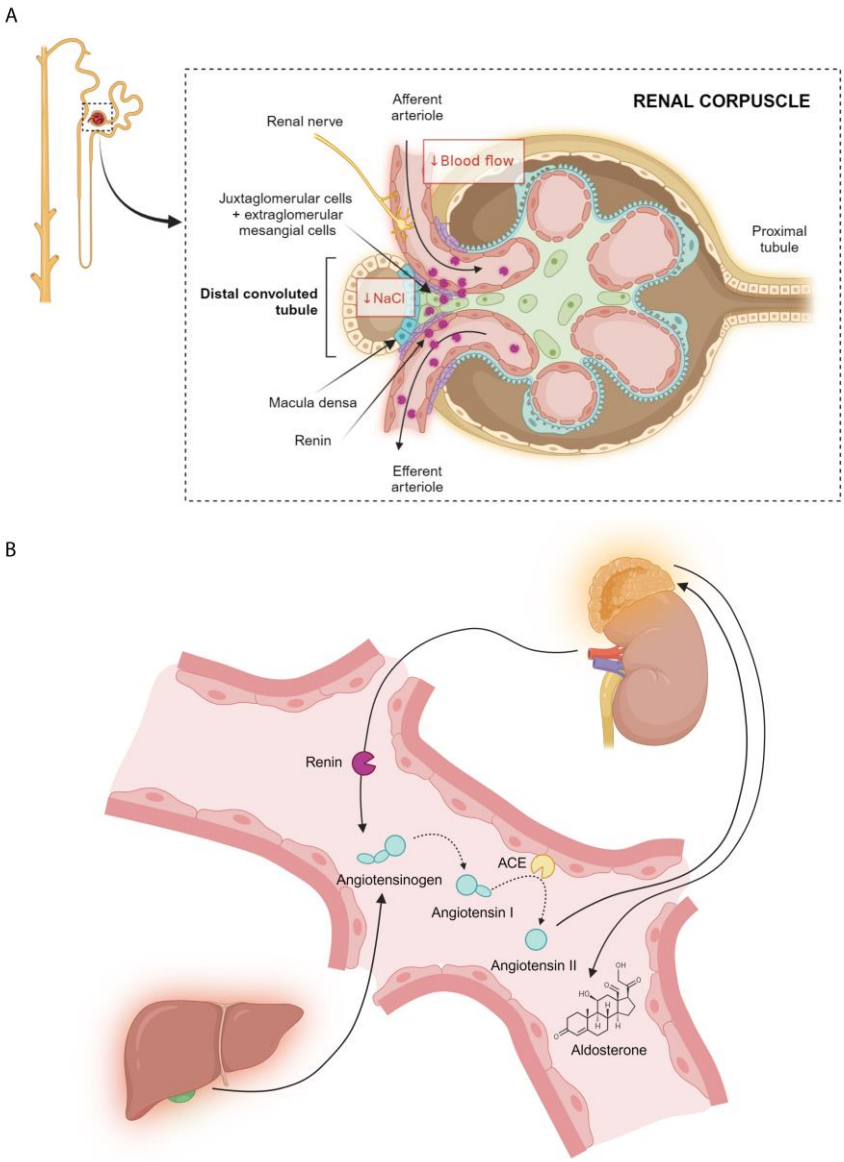


Figure 4. Graphical summary of the renin-angiotensin system. (A) Renin is mainly synthesised by extraglomerular mesangial cells and juxtaglomerular cells. Constitutively, renin is secreted as prorenin, a precursor, but in response to some stimuli, these cells secrete the active form of renin. These stimuli include a decrease in arterial blood pressure detected by baroreceptors; a decrease in sodium chloride on the distal tubules of nephrons that stimulate macula densa cells, which in turn decrease resistance for blood flow in the afferent arterioles and increase renin release; and direct activation by the sympathetic nervous system of juxtaglomerular cells. (B) Renin converts angiotensinogen released by the liver to angiotensin I, which is subsequently converted to angiotensin II by the angiotensin-converting enzyme (ACE) found on the surface of vascular endothelial cells. Angiotensin II is a potent vasoconstrictor that increases blood pressure and stimulates aldosterone secretion from the adrenal cortex. Aldosterone causes increased sodium reabsorption in the renal tubules and, therefore, water reabsorption into the blood. Overall, the response to aldosterone increases the volume of extracellular fluid in the body and consequently also increases blood pressure. *Created with BioRender.com*

Histologically, the early changes reflect a proliferative arteriopathy, with fibrous thickening of the adventitia and hypertrophic juxtaglomerular apparatus. Glomeruli may be normal or

show areas of fibrin deposit, and mesangiolysis can also be observed in the case of thrombotic microangiopathy. The tubules usually show changes due to ischemic injury, with tubular degeneration and necrosis in acute stages and tubular atrophy and interstitial fibrosis proportional to vascular injury in chronic disease. Nevertheless, fibrosis is usually restricted to perivascular regions, as opposed to other organs. On the other hand, immunofluorescence studies are non-specific and electron-dense deposits are not seen [8, 38].

In approximately 10% of patients, SRC occurs in the absence of hypertension (normotensive SRC). In some of these cases, relative hypertension may be present, a significant increase in blood pressure compared to the patient's baseline values, which still remains within the normal range. The absence of hypertension may delay SRC diagnosis and treatment, leading to disease progression. In fact, normotensive SRC is associated with a worse prognosis, higher mortality rate, and earlier need for renal replacement therapy, attributed in part to subclinical renal injury in the setting of delayed diagnosis that results in thrombotic microangiopathy and severe glomerular disease. Therefore, any change in blood pressure or the appearance of a kidney dysfunction sign in patients with SSc requires close monitoring and additional tests to exclude normotensive SRC [8, 37].

SRC may have a severe clinical presentation and represents a medical emergency since, without intervention, permanent kidney failure or even death may occur. The prognosis of SRC largely depends on the rapid control of malignant hypertension and improvement of the ongoing renal ischemia, for which angiotensin-converting enzyme inhibitors (ACE-I) have been shown to be highly effective. In fact, before the use of ACE-Is as a treatment for SRC, patients with this manifestation had a 1-year survival rate of 15%, while currently, patients treated with ACE-I present a 1-year survival rate of 76%. Moreover, if diagnosis of SRC is delayed or if ACE-Is are not used aggressively, irreversible kidney damage and other complications such as hypertension, retinopathy, pulmonary oedema, and death are more likely to occur [37, 38].

Unfortunately, despite the improvement in survival and short-term prognosis of patients with SRC due to the use of ACE-Is, it remains a life-threatening complication characterised by a high rate of progression to end-stage renal disease (ESRD) requiring dialysis. In fact, regardless of the use of ACE-Is, dialysis is still needed in 23% of SRC patients, with half of these patients requiring permanent dialysis. In these cases, renal transplantation offers superior survival compared with long-term dialysis but is not always feasible due to the severe multiorgan involvement. In part, it has been suggested that this significant long-term

mortality may be caused because of a delay in diagnosis. For this reason, it is of utmost importance to implement preventive measures such as regular blood pressure monitoring and renal function assessment (including GFR, serum creatinine and renin concentration and urine analysis) for the early detection of SRC. Nevertheless, to the present day, there is no evidence to support the use of ACE-I prophylactically. On the other hand, despite the lack of clinical studies, ACE-Is are often used in normotensive SRC patients, but they seem to be less effective in this group of patients.

In conclusion, while SRC was a predominant cause of death in previous decades, the use of ACE-I has resulted in a decline in its mortality. However, various studies still suggest a strong association between this clinical manifestation and worse prognosis, and it is considered a predictor of early mortality in SSc patients [37, 38].

SRC occurs in 10-15% of patients with dcSSc and only rarely (1-2%) in lcSSc and develops typically within the first 3-5 years, simultaneously with the worsening of the skin involvement. Several risk factors for SRC that identify the population that should be closely followed up have been identified. The strongest risk factors are the presence of diffuse skin involvement, the development of proteinuria and or hypertension, and the use of certain drugs, such as glucocorticoids and cyclosporine. Anaemia, extensive joint contractures, African American race, tendon friction rubs, digital pitting scars, myalgia and myopathy, and cardiac involvement have also been reported to be risk factors for SRC. Regarding the association of SRC with SSc-specific autoantibodies, patients positive for anti-RNApol III and anti-Scl70 present an increased risk of developing SRC, while anti-centromere autoantibody positivity is considered protective. More specifically, it has been shown that patients with anti-RNApol III autoantibodies develop SRC very early in the course of the disease (first 18 months). In contrast, anti-Scl70 positive patients typically develop SRC later (first four years). In addition, all anti-RNApol III positive patients develop hypertensive renal crisis, while 40% of patients positive for anti-Scl70 autoantibodies present with normotensive SRC. In line with this, normotensive SRC is more common in patients with cardiac involvement, indicating that this group could present a different clinical phenotype [37, 38].

1.5.3.2. ANCA-associated vasculitis and SSc

ANCA-associated vasculitis can be found in up to 9% of SSc patients. It has been postulated that SSc vasculopathy exacerbates the interaction of ANCA with the endothelium and neutrophil activation in the glomerulus. Most cases of ANCA-associated vasculitis are described as normotensive renal failure related to crescentic glomerulonephritis. In contrast

to classic SRC, this renal manifestation is more common in patients with lcSSc rather than dcSSc, and the process has a subacute presentation with progressive renal failure, mild hypertension and proteinuria that typically occurs after several years of SSc diagnosis. Despite these differences, the SRC and ANCA-associated vasculitis can only reliably be distinguished by histopathological examination, in which ANCA-associated vasculitis typically shows mononuclear cell infiltration and vessel wall destruction [37].

1.5.3.3. Isolated reduced GFR in SSc

Many patients with SSc demonstrate less severe complications associated with a decreased GFR but without clinically apparent renal disease. It has been estimated that one-third of patients with SSc have a GFR <90mL/min, with around 20% of all patients presenting a GFR <60mL/min. These alterations are usually not progressive, and kidney dysfunction tends to decline at a rate similar to that of the general population. However, in some cases, progression to ESRD may be accelerated by the presence of arterial hypertension and moderate proteinuria. Moreover, despite the slow progression of kidney dysfunction, there is a strong association between kidney involvement and outcomes in SSc, with a three-fold increased risk of mortality from PH if renal insufficiency is present [37, 38].

1.5.4. Cardiac involvement

Cardiac involvement is common in SSc, with almost all SSc patients having post-mortem pathologic findings. However, this clinical manifestation often remains clinically silent until presenting in the later stages of the disease, as it has been estimated that only 10-30% of cases with cardiac alterations are symptomatic. Nevertheless, it is important to note that almost a third of deaths related to SSc are attributed to cardiac causes. Moreover, overt cardiac involvement in SSc is associated with a mortality rate of up to 70% over five years. Furthermore, independent of mortality, cardiac involvement in SSc indicates aggressive systemic disease [8, 39–41].

SSc can cause pathology in all aspects of the heart, including the pericardium, myocardium, conduction system, vasculature, and, less commonly, the valves. However, it is crucial to consider that cardiac manifestations in SSc can also be caused secondary to PHA, ILD or SRC [8, 39–41].

1.5.4.1. Pericardial disease

Pericardial involvement is demonstrated in 30-70% of patients by autopsy or echocardiography. Nevertheless, pericardial involvement is often limited to asymptomatic small effusions that rarely cause cardiac tamponade, with only 5-20% of patients presenting symptomatic pericardial disease and a higher prevalence in lcSSc than in dcSSc (30% vs. 16%). It has been hypothesised that pericardial effusions in SSc may be secondary (transudative) in the setting of right-heart failure, as recent studies have demonstrated that effusions in patients with PAH are more common. Although pericardial effusions are generally asymptomatic or present with mild symptoms, with disease progression, patients might present with chest pain, chest tightness, dyspnoea and fever. Moreover, pericardial effusion can be a sign of impending SCR, and thus, renal function should be carefully monitored in patients with this manifestation. In some rare cases, large pericardial effusions can develop, even before skin thickening and diagnosis of SSc. Thus, SSc should be considered in patients with pericardial effusions of unknown aetiology. Other less common pericardial manifestations in SSc include pericardial inflammation, fibrinous pericarditis, fibrous pericarditis, pericardial adhesions, cardiac tamponade, and constrictive pericarditis [39–41].

1.5.4.2. Myocardial diseases

Both autopsy and magnetic resonance imaging (MRI) studies in SSc patients reveal focal myocardial lesions, varying from contraction band necrosis to regional fibrotic scarring, unrelated to any associated obstructive coronary artery disease. The pathophysiological mechanism underlying myocardial disease in SSc begins with microvascular disease with endothelial damage, which leads to ischemia-reperfusion injury, subsequent necrosis, and eventual fibrosis. Supporting the critical role of microvasculature, but not coronary arteries, fibrotic changes in the myocardium are often patchy and distributed over both ventricles, independent of coronary artery supply territories. This notion is supported by evidence that the absence of previous treatment with calcium channel blockers is an independent factor associated with myocardial disease [8, 39, 40].

Clinically, myocardial fibrosis and stiffness lead more commonly to diastolic dysfunction of ventricles (restrictive cardiomyopathy) than to systolic dysfunction. Diastolic dysfunction reflects impaired ventricular filling and causes heart failure with preserved ejection fraction, which is highly prevalent in SSc. Diastolic dysfunction of ventricles provokes congestive

heart failure in 20-25% of SSc patients and presents with symptoms of marked dyspnoea. Moreover, impaired ventricular filling may eventually originate upstream effects, such as atrial enlargement and associated dysrhythmias, pulmonary venous congestion and oedema, and ventricular systolic dysfunction, that could cause heart failure with reduced ejection fraction. Right ventricular dysfunction in SSc may be the result of left heart failure with or without preserved ejection, primary abnormalities of the right ventricle, or secondary to PAH or ILD. Similarly, right ventricle dysfunction and pulmonary haemodynamics also affect left ventricle dysfunction. Finally, it has been shown that myocardial involvement is more common in dcSSc, though patients with lcSSc and ssSSc may also present significant myocardial disease [8, 39, 40].

1.5.4.3. Cardiac conduction disease and arrhythmias

Arrhythmias are the most frequent cardiac complications of SSc, with 25-75% of SSc patients presenting abnormal ECG findings due to different patterns of arrhythmias. Arrhythmias can be caused by autonomic cardiac neuropathy, myocardial fibrosis in the conduction system, microvascular injury, or a combination of all these processes. With disease progression, fatigue, palpitations, syncope, and dizziness are the typical clinical manifestations of patients with arrhythmias. In addition, more severe manifestations can also be caused by arrhythmias, and it has been estimated that up to 6% of SSc-related deaths are attributed to this pathology [39, 40].

1.5.5. Gastrointestinal tract involvement

Gastrointestinal (GI) involvement is the most frequent internal organ system manifestation in SSc, as 90% of patients with SSc present different forms of GI affectation. Moreover, while GI involvement can occur at any time in the disease course, it is the presenting manifestation of the disease in 10% of cases. Although GI manifestations can affect any part of the GI tract, they disproportionately impact the upper tract, with evidence of oesophagus involvement in up to 90% of SSc patients by endoscopy. Oesophageal manifestations in SSc are associated with a hypotensive lower oesophageal sphincter and reduced or absent peristalsis, which can lead to gastroesophageal reflux (GER) disease with or without oesophagitis, stricture formation, Barrett's oesophagus (a pre-malignant condition caused by chronic acid exposure), adenocarcinoma and oesophageal dysmotility. These oesophageal manifestations are often associated with progressive motility dysphagia that generally presents later in the course of the disease, compared to GER, which frequently

presents at the onset. Nevertheless, many patients are asymptomatic (up to 80%) despite presenting oesophageal involvement by endoscopy. Patients who present with typical symptoms of GER should be treated empirically with proton pump inhibitors or H_2-receptor antagonists. However, mechanical obstruction and dysmotility evaluation should be performed if empiric therapy fails. In fact, when there is a concern for dysmotility, manometry has been shown to have higher sensitivity and is recommended when the initial assessment is unrevealing [17, 42].

On the other hand, oesophageal disease has been identified as a risk factor for the development of ILD. In fact, it has been shown that GER and regurgitation positively correlate with the degree of pulmonary fibrosis. Although the causal association has not been determined, microaspiration of gastric contents into the lungs has been implicated as a trigger of pulmonary parenchymal lesions, suggesting that GER therapy could improve pulmonary symptoms and PFTs. As previously explained in section 1.5.2.2., NSIP is the predominant histologic pattern in SSc. Nevertheless, a novel histologic pattern called centrilobular fibrosis (CLF), distinct from NSIP, has been described in SSc and associated with GER. This pattern is characterised by a predominant bronchocentric distribution of the lesions and the presence of intraluminal basophilic content and foreign bodies inside the bronchi, occasionally with multinucleated giant cell reaction. Overall, only 21% of SSc lung biopsies present isolated CLF, but this pattern is also present in 84% of patients with a predominant NSIP pattern [42, 43].

Although the most frequent GI organ involved in SSc is the oesophagus, the anorectum and small bowel are also frequently affected. Small bowel dysmotility may be seen in up to 60-80% of SSc patients. Studies have shown reduced amplitude and frequency of the migrating motor complex in SSc, resulting in delayed transit through the small bowel, associated with small intestinal bacterial overgrowth (SIBO), and more severe manifestations such as pseudo-obstruction. SIBO, which affects 40% of patients with SSc, often presents with postprandial distension, nausea, diarrhoea, and excessive flatulence. Moreover, it is associated with malabsorption, leading to deficiencies of vitamin B12, iron, and fat-soluble vitamins. On the other hand, the anorectum is affected in 50-70% of patients with SSc, with more than 20% of patients developing faecal incontinence. It has been suggested that there is an atrophy of the internal anal sphincter, which is a smooth muscle similar to the internal oesophageal sphincter. This atrophy may be secondary to neural dysfunction or primary smooth muscle dysfunction. Rectal prolapse can also occur, manifesting with a bulging sensation in the anus and chronic stool leakage [42, 43].

Stomach involvement in SSc includes gastric antral vascular ectasia (GAVE) and gastroparesis. Most patients with GAVE present with iron-deficiency anaemia; however, in some rare cases, severe bleeding can occur, and it can be the presenting manifestation of SSc. The pathogenesis of GAVE has been proposed to be similar to that of the immune-mediated development of telangiectasis. Indeed, histologically, GAVE is characterised by mucosal capillary dilatations containing fibrin thrombi, fibromuscular hyperplasia and reactive foveolar epithelial changes. On the other hand, gastroparesis leads to delayed gastric emptying, which contributes to GER, and presents with early satiety, postprandial nausea, distention, abdominal pain and vomiting with subsequent malnutrition. As these symptoms can be seen in up to 80% of patients with SSc and are not specific, evaluation for gastric dysmotility is generally required prior to initiation of therapy [42–44].

Colonic involvement is also present in a subset of SSc patients and presents with diarrhoea or constipation. Constipation is often prominent early in the course and is associated with colonic dysmotility, although it can also be the result of intestinal pseudo-obstruction, which affects 4%-10% of patients. Pseudo-obstruction presents with the inability to move intestinal luminal contests forward but in the absence of a mechanical true obstructive process. This condition is painful, often recurrent and, at times, life-threatening. On the other hand, diarrhoea frequently occurs in prolonged disease and is multifactorial due to coexistent SIBO, fibrosis of lymphatic drainage and impaired reabsorption. Additionally, steatorrhea and malabsorption may also be markers of overlap with primary biliary cirrhosis (PBC) [17].

In addition, numerous deleterious effects on the oropharynx are seen in SSc, including reduced interincisal distance and changes of the mandible, which, together with sicca syndrome (see section 1.5.8.), contributes to interference with mastication and oral hygiene, leading to increased periodontal carious disease. Furthermore, approximately 25% of patients with SSc may develop oropharyngeal deglutition dysfunction, which is related to tongue-palate incompetence, inadequate pharyngeal contraction, and laryngeal and epiglottic dysfunction. Oropharyngeal dysfunction is associated with the duration of the disease, occurring several years after the onset of RP and has been associated with oesophageal and pulmonary disease [42].

The pathophysiology of GI tract involvement in SSc is not completely understood, but a four-stage process has been proposed. The first proposed event is an early vascular lesion, resulting in capillary loss and arteriolar stenosis, which may induce hypoxia and a subclinical enteropathy that manifests as mild changes in intestinal permeability, transport, and absorption. The second stage is based on neural dysfunction and is responsible for the

earliest symptomatic lesions, which can present as hypomotility or dysmotility. This dysfunction produces abnormalities in smooth muscle function before smooth muscle contractility is impaired by atrophy and collagen deposition. The next stage is smooth muscle atrophy, which may be partially reversible with prokinetic drugs. Superimposed on existing neural dysfunction, atrophied muscle is capable of responding but only weakly. The end-stage lesion is muscle fibrosis, which leads to an altered peristaltic activity with multiple secondary problems. At this final stage, pharmacologic restoration of function is not possible [8, 45].

Although severe GI manifestations are mostly related to fibrosis, early neural disease without fibrosis can also lead to clinical manifestations, as previously mentioned. It has been demonstrated that this early neural disease is caused not only by the initial hypoxia but also by autoantibodies directed against the myenteric neurons, specifically against the muscarinic acetylcholine receptor subtype 3 (M3R). These autoantibodies have been shown to be pathogenic, with their titres having a positive correlation with the severity of GI symptoms. However, the autoimmune aetiology does not explain the entire spectrum of GI disease in SSc. In line with this, as previously described, significant heterogeneity exists among patients, with some presenting with a predominance of upper GI dysfunction and others presenting with a predominance of lower GI dysfunction. Whereas the M3R has been shown to be responsible for lower oesophageal sphincter activity, other receptors, primarily the muscarinic acetylcholine receptor subtype 2, are responsible for oesophageal motility, suggesting that different biological mechanisms may be necessary in explaining the phenotypic differences [42, 43].

Another essential component of GI-specific pathology is gut microbiota, which affects the development and function of the immune system and seems to play a role in autoimmune diseases through microbiota-related immune dysfunction. In SSc, several cohorts have demonstrated decreased levels of commensal bacteria and increased levels of pathogenic bacteria compared to healthy individuals. Nevertheless, it is unclear whether microbiota changes precipitate and perpetuate the SSc-associated immune system or result from SSc itself and related therapies [8].

Overall, GI manifestations have a profound impact on quality of life and are a cause of depression in SSc patients. More importantly, severe GI involvement, such as pseudo-obstruction and malabsorption, which affects 8% of SSc patients, presents a very poor prognosis. Indeed, these severe GI manifestations may result in dependence on enteral or total parenteral nutrition and are associated with recurrent hospitalisations, leading to high

mortality, with only 15% of such patients surviving after nine years. Moreover, malabsorption alone is an independent predictor of mortality, with a 50% mortality rate at 8.5 years, and most patients with severe malnutrition who need total parenteral nutrition die within two years from the consequences of their disease [17, 42, 43]. There is limited information about risk factors associated with severe GI involvement, but possible risk factors include the presence of anti-U3 snoRNP, anti-U11/U12 snRNP, and anti-M3R autoantibodies. Regarding clinical and demographical characteristics of patients with GI involvement, it has been demonstrated in both dcSSc and lcSSc subsets equally, but is more prevalent in African American SSc patients [43].

1.5.6. Hepatic involvement

SSc is rarely associated with severe liver complications. PBC is the most common liver disease in SSc patients, with a reported prevalence of 2-22% that increases when anti-mitochondrial (AMA) and anti-gp100 autoantibodies are employed for diagnosis. Moreover, the development of PBC in SSc patients is highly associated with anti-centromere autoantibody. Temporally, the onset of PBC may be first, concurrent, or follow SSc disease onset, but in general, the prognosis of SSc-PBC is better than that of PBC alone, with a slower progression to end-stage liver disease and usually not requiring any treatment. In fact, PBC in SSc is generally clinically silent despite the elevation of cholestatic enzymes and the presence of AMA and IgM hyperglobulinemia. Histologically, portal track fibrosis, in the absence of any other abnormalities, can be detected in the liver of SSc patients with PBC. However, portal hypertension is quite rare in SSc patients. It has been proposed that SSc and PBC may share pathogenetic pathways that give rise to chronic inflammation, leading to duct destruction and a fibrogenic response, promoting a fibroproliferative response to cholestatic injury. However, the liver is relatively protected from extensive fibrosis compared to other organs, such as the skin, GI tract, lungs, and heart. So far, the underlying mechanisms remain unknown, but the unique vascular system of the liver may modify the common pathologic cascade of SSc, especially the process of bridging vasculopathy and tissue fibrosis [8, 42, 44, 46].

On the other hand, SSc has been suggested to be associated with nodular regenerative hyperplasia of the liver (NRHL), which is histologically defined by diffuse micronodular transformation without fibrous septa. NRHL is thought to develop as a result of microvascular alterations due to endothelial cell damage. Patients with NRHL may remain asymptomatic, but at least 50% of reported cases present portal hypertension and related symptoms,

including splenomegaly, ascites and oesophageal or gastric varices. In these cases, transaminases might be normal or slightly elevated, whereas cholestatic measures are often more significantly increased. Supporting the contribution of vasculopathy to the development of NRHL, SSc patients with this liver involvement are highly susceptible to vascular complications, such as digital ulcers, PAH and SRC [8].

1.5.7. Musculoskeletal involvement

1.5.7.1. Muscle involvement

The prevalence of muscle involvement in SSc varies widely, ranging from 5% to 96%, due to the lack of diagnostic consensus criteria and the inclusion or exclusion of SSc-myositis overlap syndromes in studied cohorts. Indeed, there is no unanimity as to whether an inflammatory myopathy in SSc should rather be considered as a disease symptom or as SSc-myositis overlap since SSc is the most common connective tissue disease associated with inflammatory myopathies and accounts for 30-40% of patients with myositis overlap. Nevertheless, usually, patients with SSc diagnosis and myopathy are considered as presenting SSc-myositis overlap only when they also satisfy the published diagnostic criteria for polymyositis (PM) or dermatomyositis (DM), which has been estimated to occur in around 10% of SSc patients [47–49].

As myopathy is relatively frequent in SSc patients and may be an early disease manifestation, all patients should be screened for muscle involvement by clinical evaluation. SSc patients with myopathy usually present symmetric proximal limb weakness, although distal weakness may also be detected. Muscle weakness evaluation by physical examination may be challenging as it is difficult to distinguish it from the limitation of movement due to skin sclerosis, articular changes in proximity to the assessed muscles or fibrosis of underlying tissues. Moreover, muscle weakness in SSc can also be secondary to fatigue, pulmonary and heart involvement and atrophy due to weight loss or due to side-effects of drugs. Therefore, additional testing is indicated when the history or physical exam suggests the possibility of proximal muscle weakness, including laboratory testing of muscle enzymes and muscle biopsy. Less commonly, other muscles distinct from the muscles of the limbs might also be affected, such as head extensor muscles. There is no data about the involvement of respiratory muscles in SSc; however, in patients with SSc-myositis overlap syndrome, respiratory muscles may also be affected [47, 49].

Laboratory testing for evaluation of muscle involvement includes creatine kinase (CK) and aldolase levels, as elevation of one or both are characteristic of an underlying myopathic process. Moreover, it has been shown that CK levels correlate negatively with muscle strength. However, a normal value does not exclude inflammatory myopathy, as it has been demonstrated that only 82% of patients with biopsy proven myositis had increased CK, and only 76% had increased aldolase. On the other hand, electromyography (EMG) is not very useful for the evaluation of muscle involvement, as it displays pathologic findings in the vast majority of SSc patients regardless of clinical muscle involvement, laboratory features, or disease duration, with features similar to those of patients with polymyositis. In line with this, currently, muscle biopsy is the most sensitive and recommended test for diagnosing muscle involvement. Nevertheless, the histological findings of muscle biopsies in SSc patients with myopathy are heterogeneous and non-specific, as they are indistinguishable from patients with PM or DM. These findings include interstitial fibrosis in the perimysium and epimysium, intimal proliferation of endomysial and perimysial vessels and other vascular abnormalities, perivascular infiltrates, atrophy, myofiber necrosis and regeneration of variable degree. In addition, overexpression of HLA I, complement deposits on vascular walls and predominance of CD4+ T cells, similar to what is observed in DM, or absence of complement deposits with a predominance of CD8+ cells, like in PM, are observed in SSc patients [47–49].

Overall, it has been estimated that myopathy is more common in the dcSSc subset than in the lcSSc subset. Moreover, it seems that patients with musculoskeletal disease present a worse prognosis, as this manifestation has been associated with an increased risk of myocardial involvement. However, several autoantibodies have been associated with myopathy, even in the lcSSc subset. Specifically, anti-PM/Scl and anti-Ku positive patients are commonly classified as presenting SSc-myopathy syndrome, while anti-U3 snRNP autoantibodies are associated with isolated myopathy in SSc. On the contrary, anti-centromere autoantibody presence seems protective for developing any muscle involvement in SSc [47, 49].

1.5.7.2. Skeletal involvement

Skeletal involvement in SSc can be divided into articular and non-articular.

1.5.7.2.1. Articular involvement

The frequency of joint involvement in SSc is inconsistent between studies, partly because of the difficulties of its evaluation by physical examination and partly because of the lack of

consensus on assessment techniques. It seems that arthralgia and joint contractures are the most common manifestations, followed by arthritis, synovitis and tendon friction rubs (TFR). Joint involvement can be the initial manifestation of SSc, and its onset is very heterogeneous, as it can be acute or insidious, with an intermittent, chronic remittent, slowly progressive or rapidly progressive course, and present in monoarticular, oligoarticular, or polyarticular pattern. Despite this heterogeneity, in SSc the overall involvement of the hands is more prominent and frequent than the feet, but feet involvement should also be considered [47].

Joint pain is widespread in patients with SSc, with some studies reporting arthralgias in almost all patients. Additionally, it can be the presenting symptom in one-fifth of patients. Radiographic studies demonstrate a wide range of joint pathology, from inflammatory arthritis to osteoarthritis, but the most common presentation is joint space narrowing. Patients who fulfil the classification criteria for both SSc and RA are considered to present SSc-RA overlap syndrome. However, since SSc itself presents with significant articular damage and similar changes are noted in joints of the hands of RA and SSc patients, the determination of SSc-RA overlap is often difficult. In addition, rheumatoid factor and anti–cyclic citrullinated peptide antibodies have not shown a clear association with SSc-RA overlap syndrome, although they may be more prevalent in this subset of patients [47, 50].

Tenosynovitis, the inflammation of the tendon sheath that results in damage to the tendon and surrounding structures, is seen in up to one-third of patients with SSc. Although some patients experience inflammatory tenosynovitis similar to RA, patients with SSc are more likely to have fibrotic tendon sheath thickening, termed “sclerosing tenosynovitis”. On the other hand, TFRs are characterised by a leathery crepitus felt above the tendons but do not necessarily mean inflammation of the tendon sheath. TFRs were initially assumed to be caused by fibrinous deposits in tendon and tendon sheaths. Still, newer imaging studies suggest that the infiltrates are localised in the deep soft tissues and septae that divide fat into lobules, leading to the thickening of tendons and retinacula (layer of connective tissue around tendons). While tenosynovitis mainly affects extensor tendons, TFR occurs in both the extensor and flexor tendons of the fingers and wrists, and tendons over the elbows, knees and ankles. Although TFRs do not require surgical intervention, they are correlated with increased functional disability, fingertip ulcers, gangrene and amputations, and joint contractures. In addition, tenosynovitis and TFRs are more common in dcSSc and are associated with worse disease severity, increased renal, cardiac, pulmonary, and gastrointestinal involvement and decreased survival rates [47, 50].

Finally, claw hand deformity with proximal interphalangeal joint flexion contractures is present in up to 30% of patients with SSc and is often progressive. Joint contractures result from sclerosis of the skin, volar plates (a fibrocartilaginous structure that reinforces the joint capsules of metacarpophalangeal and interphalangeal joints), and flexor tendon sheaths with concomitant flexor tendon shortening. With time, central slip attenuation and lateral band volar subluxation ensue (**Figure 5**). Progressive proximal interphalangeal flexion contracture causes tension, ischemia, and atrophy of the dorsal skin, leading to ulcerations and infections. Surgical procedures aimed to improve the range of motion of the proximal interphalangeal joint can be performed. Overall, joint contractures are more common in dcSSc but can occur in both subtypes [50].

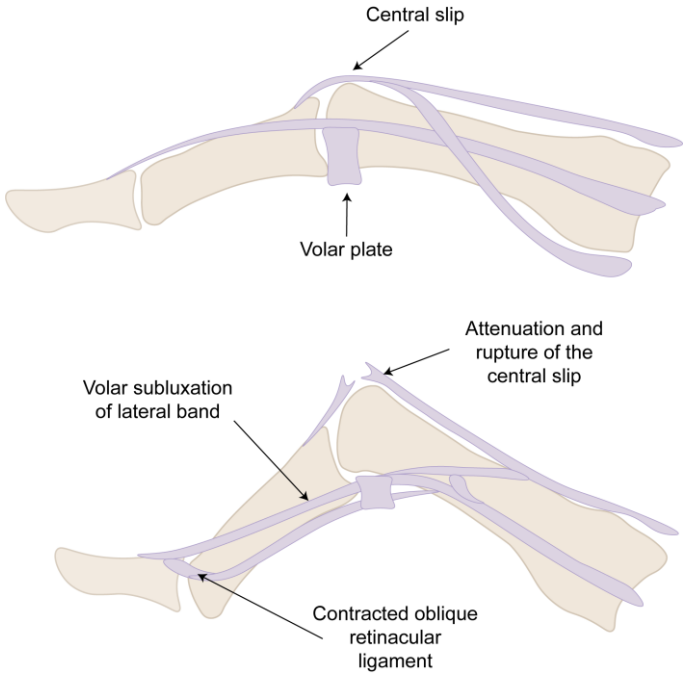


Figure 5. Proximal interphalangeal joint flexion contractures. Sclerosis of the lateral band and volar polar plates results in proximal interphalangeal joint flexion contracture, leading to a central slip attenuation that progresses until tearing. Contraction of the oblique retinacular ligament also takes place.

1.5.7.2.2. Non-articular involvement

The primary forms of non-articular involvement in SSc are generalised and localised osteoporosis, digital tuft resorption, and osteolysis in other body regions. Many studies have established an increased risk of bone loss and fracture in individuals with chronic inflammatory conditions. Still, occult malabsorption or malnutrition, major disability, and the use of corticosteroid therapy can also contribute to bone disease [47].

1.5.8. Sicca syndrome

The exact prevalence of sicca syndrome in SSc varies according to inclusion criteria and definition of sicca symptoms. Reported subjective xerostomia varies from 50-80%, while 60-70% of patients present sicca syndrome when evaluated by objective functional tests such as Schirmer's test or salivary flow rate. Nevertheless, studies exploring the correlation between sicca symptoms and histological findings have demonstrated a lack of concordance between subjective sicca symptoms and histological sialadenitis typically seen in Sjögren's syndrome (SjS), while fibrotic lesions on biopsies of the minor salivary glands have been shown to better correlate. Fibrotic lesions are defined as fibrous tissue with secreting fibroblasts surrounding the acini and located around capillaries and excretory ducts. These findings suggest that sclerosis of eccrine glands causes xerostomia and reduced tear secretion resulting in sicca syndrome in SSc patients.

Reduction of salivary flow can lead to several complications, including difficulties in eating dry food, speaking without taking frequent sips of water, increased risk of oral infections and a high rate of caries. On the other hand, patients with xerophthalmia may report burning or soreness of the eyes and increased sensitivity to light. Sicca symptoms equally affect patients with lcSSc and dcSSc but present a lower prevalence in ssSSc patients. Regarding autoantibody status, sicca syndrome is more frequent in SSc patients positive for anti-centromere autoantibody [51].

Although sicca symptoms can be detected in a high number of SSc patients, only 10-20% of SSc patients are considered to present an overlap with SjS syndrome. Studies comparing SSc-SjS overlapping patients with SjS patients showed that this subset of patients had similar characteristics regarding sicca manifestations and the presence of anti-Ro60 and anti-La autoantibodies. On the other hand, it has been suggested that SSc patients with SjS overlap syndrome tend to have a milder disease phenotype, with less lung involvement, but with a specific subgroup characterised by anti-centromere autoantibody positivity, lcSSc and higher risk of non-Hodgkin's lymphoma compared with the rest of SSc patients. Additionally, a specific subset of patients with SjS and positivity for anti-centromere autoantibody but without SSc has also been recently explored. Whether this population of patients with sicca symptoms corresponds to a specific subgroup of SjS or a pre-scleroderma condition overlapping with SjS is still to be determined. These data may suggest that an immunological continuum may exist between SjS and SSc patients with anti-centromere autoantibody [51, 52].

1.5.9. Malignancy

Compared to the general population matched for sex and age, a higher risk for cancer, particularly for lung, breast, and cutaneous melanoma, has been reported in SSc patients. Common genetic and environmental risk factors may be involved in developing both diseases. Indeed, some SSc pathophysiological mechanisms, such as immune and vascular dysregulation and exuberant fibrogenesis, are also involved in cancer development. Moreover, cancers in SSc seem to be more frequent in sites affected by exaggerated fibrosis or high disease activity and damage, suggesting that a persistent inflammation leading to fibrosis may underlie in places where the malignancy develops. For example, data suggests that there is an increased risk of lung cancer in patients with ILD, skin cancer in thickened areas, oesophagus cancer in patients with severe reflux and Barrett's oesophagus, liver cancer in patients presenting overlap with PBC and thyroid cancer if there is concomitant autoimmune thyroiditis. Nevertheless, the link between SSc and oncogenesis remains unknown, and the relationship between cancer and SSc is likely complex and bidirectional. In fact, a biphasic association between SSc and cancer is observed, with a first peak of cancer incidence around the time of scleroderma onset and a second peak after 6-8 years of disease of duration, which suggests distinct pathogenic mechanisms [53, 54].

When cancer occurs within the first five years of SSc onset, it is hypothesised that malignant cells or mutated autoantigens may serve as immunogens, leading to an autoimmune response in a genetically susceptible host, indicating that SSc could be a paraneoplastic disease. In this line, some reports indicate that synchronous cancer treatment and remission result in dramatic improvements in SSc. Anyhow, it should also be pointed out that when cancer is diagnosed a few months prior to SSc onset, SSc may arise as a consequence of anti-cancer therapies, including chemotherapy, radiation therapy, and immunotherapy. Indeed, it has been reported that the use of immune checkpoint inhibitors targeting cytotoxic T lymphocyte antigen 4 (CTLA-4) and programmed cell death-1 (PD-1) and its ligand (PD-L1) is associated with the development of immune-related adverse effects and disease exacerbations in patients with pre-existing autoimmune diseases, including features of SSc [53–55].

In the second peak of incidence, the trigger could be linked to chronic inflammation and mesenchymal dysfunction, which are characteristic of SSc. Moreover, this second peak in cancer incidence could also be related to the use of immunosuppresants. For example, it has been described that exposure to cyclophosphamide, an alkylating agent, increases the

risk of hematologic and bladder cancers, in particular with higher cumulative doses. It is also important to note that patients with SSc may have a high cumulative exposure to ionising radiation from medical tests over time, potentially increasing the risk of cancer development. In fact, it has been described that radiation therapy may trigger both cutaneous and pulmonary fibrosis, although most reports describe localised scleroderma or exaggerated fibrosis developing in patients with a previous diagnosis of SSc [53, 54].

Identifying risk factors to recognise patients more likely to develop cancer is essential for establishing different cancer screening follow-up recommendations. In SSc patients, the association of cancer and traditional cancer risk factors, such as smoking and immunosuppressive drugs, remains controversial. Regarding demographic characteristics, it seems that there is a higher incidence of cancer in men than in women and patients with an older age of onset of disease. As previously discussed, some clinical characteristics, such as lung fibrosis or Barrett's oesophagus and order conditions, are also associated with specific cancers. On the other hand, some SSc-specific autoantibodies have been demonstrated to be associated with an increased risk of cancer and are used to identify patients in which a more strict cancer screening follow-up should be carried out [53, 55].

Malignancy frequency is higher in anti-RNApol III positive patients than in those with anti-Scl70 or anti-centromere autoantibodies (14.2-20.9% vs. 6.3-13.6% vs. 6.8-16% respectively). Moreover, patients with anti-RNApol III autoantibodies had a significantly shorter cancer-SSc interval than those with anti-Scl70 or anti-centromere (medians of -1.2 years, +13.4 years and +11.1 years, respectively). Indeed, when compared to other SSc patients, anti-RNApol III autoantibody positive patients presented nearly a 5-7 times increased risk of developing synchronous cancer. Even when compared to patients matched for sex, cutaneous phenotype, age of SSc onset, and disease duration but negative for anti-RNApol III, cancer was significantly more common in the anti-RNApol III positive group (17.7% versus 9.0%), particularly concerning cancers diagnosed within two years of SSc onset (9.0% vs. 2.5%). Breast cancer has been reported to be the most frequent cancer found in anti-RNApol III positive patients, although other malignancies are also found in this group of patients [53, 54].

Other SSc-specific autoantibodies have also been associated with cancer, although in a less robust way. For example, some reports indicate that anti-Scl70 positive patients have a higher risk of lung cancer related also with ILD, as previously discussed. More interestingly, an increased risk of developing synchronous breast cancer and melanoma is found in patients with lcSSc negative for anti-centromere, anti-Scl70, and RNApol III autoantibodies.

This is a heterogeneous population that likely consists of different subsets of patients, and probably present known or unknown autoantibodies that could also predict the development of malignancy. In fact, patients positive for anti-PM/Scl and anti-U11/U12 snRNP autoantibodies have been reported to have an increased risk for developing cancer, although it has not been confirmed. In any case, older age at SSc onset is considered a risk factor for cancer development and for its occurrence in close temporal relationship with SSc in all these groups of patients, particularly in patients with anti-aScl70 autoantibodies and the subset of patients negative for anti-centromere, anti-Scl70 and anti-RNA pol III autoantibodies. On the other hand, the presence of some autoantibodies seen to be associated with a lower risk of developing malignancy, such as anti-centromere and anti-Th/To autoantibodies [53, 54, 56].

Altogether, all these data suggest that clinical and instrumental examinations indicated for age- and sex-matched general population have to be performed in SSc patients. Although no universally accepted malignancy-screening protocols exist for SSc, it is generally recommended to apply more intensive screening methods when risk factors are present, including older age at disease onset, diffuse disease type, family history of cancer, concomitant precancerous conditions such as Barrett's oesophagus or PBC, globus sensation or unexplained dysphagia that could indicate head or neck cancers, significant exposure to cyclophosphamide, cytopenias, monoclonal gammopathies or presence of certain autoantibodies. Screening at baseline and regular follow-up are also recommended, which should be scheduled considering the biphasic relationship between cancer and SSc. Clinical examination for malignancy screening is based on evaluating new or changing skin lesions, lymphadenopathy or suspicious symptoms such as weight loss or unexplained fever. In addition, breast and prostate evaluations and ultrasound examinations must be considered with all the other non-invasive tests. However, additional studies are underway to define the optimal approach to cancer screening in these high-risk subsets that maximises cancer detection while minimising the harms of over-screening [17, 53, 54].

1.5.10. Mortality

Systemic sclerosis is one of the autoimmune systemic diseases with worse prognosis. Several studies have been performed since the 1960s, reflecting a higher mortality rate between 1 to 7-fold compared with the general population and a standardised mortality ratio of 2-4. After the introduction of new therapies in the last decades, there have been changes in the pattern of death, with a significant reduction in the number of deaths related to kidney

involvement because of the use of ACE-I (see section 1.5.3.1.). Nowadays, ILD, PAH and cardiovascular diseases, followed by malignancies and infections, are the leading causes of death in SSc patients. Furthermore, although survival studies have shown improvement over the years, SSc and, more specifically, dcSSc, still carry the highest case-based morbidity and mortality among all rheumatic diseases. While survival of lcSSc patients has been estimated to be 90% at 5 years and 80% at 10 years, dcSSc patients present a survival of 70% at 5 years and 55% at 10 years [53, 57–60].

1.6. Classification

The wide spectrum of clinical presentations, as discussed in the previous section, has considerably impacted the classification criteria of SSc in the past and is still a challenge for daily clinical management as well as for the proper selection of patients in clinical trials. Indeed, as internal organ involvement is closely linked to prognosis, recent attempts for disease subgrouping have emerged in order to fully capture the disease heterogeneity [61].

The first set of classification criteria was endorsed by the ACR in 1980. Based on these criteria, for definite SSc classification of a patient, one major criterion (proximal cutaneous sclerosis) or at least two minor criteria are required (**Table 1**) [62].

Table 1. The 1980 ACR classification criteria for SSc.

Major criterion	Minor criteria
Proximal cutaneous sclerosis*	Sclerodactyly
	Digital pitting scars of fingertips or loss of substance of the distal finger pads
	Bibasal pulmonary fibrosis in the absence of proximal scleroderma

*Proximal cutaneous sclerosis is defined as tightness, thickening, and nonpitting induration of the skin proximal to the metacarpophalangeal or metatarsophalangeal joints, affecting other parts of the extremities as face, neck or trunk, usually bilateral, symmetrical and almost always including sclerodactyly.

In 2013, a revised set of criteria was adopted by the EULAR and ACR, with the objective of including patients with early and more advanced SSc that were excluded by the previous set of criteria. As patients with recent disease onset may be more susceptible to benefit from early therapeutic interventions, these updated criteria included clinical features classically present in early or very-early SSc, such as puffy fingers, SSc-specific autoantibodies and abnormal NVC (see section 1.5.1.2.1.). Current clinical trials generally use these classification criteria, in which a patient requires a total score of 9 or more to be classified as having definite SSc (**Table 2**) [20].

Table 2. The 2013 ACR/EULAR classification criteria for SSc.

Items	Sub-Items	Score
Skin thickening of the fingers of both hands extending proximal to the metacarpophalangeal joints		9
Skin thickening of the fingers (only counts the highest score)	Puffy fingers Sclerodactily*	2 4
Finger tip lesions (only counts the highest score)	Digital tip ulcers Pitting scars	2 3
Telangiectasia		2
Abnormal nailfold capillaries		2
PAH and/or ILD (2 points maximum)	PAH ILD	2 2
RP		3
SSc-specific autoantibodies (anti-centromere, -Scl70, -RNApol III)		3

*Distal to metacarpophalangeal joints but proximal to proximal interphalangeal joints

These two classification criteria did not aim to highlight specific subgroups of SSc patients with distinct prognoses. To date, the subgrouping of SSc is mainly based on the classification criteria proposed by *LeRoy et al.* in 1988 (**Table 3**) and revised in 2001 (**Table 4**).

Table 3. The 1988 LeRoy classification criteria for SSc.

dcSSc	lcSSc
Onset of RP within 1 year of onset skin changes (puffy or hidebound)	RP for years (occasionally decades)
Early and significant incidence of ILD, oliguric renal failure, diffuse GI disease and myocardial involvement	Significant late incidence of PH, with or without ILD, skin calcifications and telangiectasia
Absence of anti-centromere autoantibody and presence of anti-Scl70 autoantibody	High incidence of anti-centromere autoantibody
Nailfold capillary dilation and capillary destruction	Dilated nailfold capillary loops, usually without capillary dropout
Presence of TFR	

Table 4. The 2001 LeRoy and Medsger revised classification criteria for SSc.

lSSc
RP (objective documentation) plus SSc-type NVC pattern or presence of SSc-specific autoantibodies*
RP (subjective documentation) plus both SSc-type NVC pattern and SSc-specific autoantibodies*

*Anti-centromere, -Scl70, -U3 snoRNP, -PM/Scl or -RNApol III autoantibody at a titre of >1:100.

Original LeRoy criteria proposed two SSc subsets relying solely on the skin fibrosis extension: lcSSc, in which skin fibrosis is limited to the distal limbs (hands, feet and forearms) and face, and dcSSc, associated with early skin changes affecting also the trunk and proximal limbs (**Figure 2**). Although this subclassification highlighted NVC abnormalities, the presence and

severity of organ involvement, and specific autoantibodies associated with these specific cutaneous subgroups, they were not included in the definition of these subsets per se [4].

These classification criteria were revised in 2001 (LeRoy and Medsger criteria) for the classification of early SSc. As previously mentioned, early diagnosis is important as it can imply the initiation of disease-modifying treatments that have the potential to improve patient outcomes. As the earliest sign of SSc is normally RP, this classification criteria are based on the presence of this manifestation. Anyhow, RP is non-specific for SSc and can have several alternative causes, and for that reason, criteria for detecting patients with RP who are at greatest risk of the development of SSc were proposed. LeRoy and Medsger 2001 criteria considered SSc-specific autoantibodies and scleroderma-type changes on NVC together with RP to diagnose a patient as having pre-scleroderma, as it was demonstrated that a definite diagnosis of SSc occurred in 65.9% and 72.7% of patients at 5 and 10 years, respectively, in those with RP who had both SSc-specific autoantibodies and a "scleroderma pattern" on NVC (see section 1.5.1.2.1.). Pre-scleroderma was defined as limited SSc (lSSc), which differed from lcSSc by the absence of cutaneous involvement, and differed from ssSSc, which was previously defined as visceral manifestations of the disease in the absence of cutaneous changes [23]. Based on these criteria, lcSSc and dcSSc classification is considered when criteria for lSSc are present together with distal or proximal cutaneous changes, respectively.

These classification criteria have been widely used in clinical trials, as it is well accepted that dcSSc has a higher mortality rate and earlier visceral involvement than lcSSc. However, nowadays, it is recognised that there is a lack of strict parallelism between the degree of skin extension and severity of the disease, and recent studies with large sample sizes have highlighted that significant heterogeneity exists within these subsets. Characterising the heterogeneity and including this multifaceted diversity within new comprehensive subgroups that capture and classify the entire phenome of SSc have become a new important challenge. Various studies have shown that autoantibodies could predict organ damage and survival better than the cutaneous classification alone. A study from the European Scleroderma Trials and Research Group (EUSTAR) used unsupervised machine learning analysis to distinguish and characterise homogeneous groups of SSc patients without any prior assumption. In this case, primary cluster analysis resulted in two subsets that only partially overlapped with classical cutaneous classification. Moreover, an exploratory analysis further yielded 6 homogeneous groups of patients that largely differed regarding clinical features, autoantibody profiles and mortality. Most importantly, the presence of

organ damage and serological profiles markedly impacted survival regardless of cutaneous involvement, highlighting that skin involvement does not fully represent the whole clinical spectrum and associated prognosis [63]. Moreover, recent studies have proposed the combination of autoantibody specificity and extent of skin involvement for defining SSc subsets, demonstrating that this classification can predict clinical presentation and outcomes of SSc patients better than classical classification defined only by cutaneous extension [64].

1.7. ANAs in SSc

Anti-nuclear autoantibodies (ANAs) are detectable by indirect immunofluorescence (IIF) in 90-95% of patients with SSc and target molecules involved in vital processes such as transcription, splicing, and cell division. ANAs constitute a diagnostically important feature of the immunological abnormalities on SSc, and as already discussed, SSc-specific autoantibodies are included in some of the present classification criteria for SSc (see section 1.6.). SSc-specific ANAs are rarely observed in other connective tissue diseases and are even less frequent in non-immune disorders or in healthy individuals. Moreover, these autoantibodies are normally mutually exclusive and have been linked with distinct clinical subsets of SSc (Figure 6). Therefore, detecting a particular autoantibody can help in predicting the possible organ involvement and prognosis of the patient, which may have an impact on monitoring and treatment. This is especially important in patients with very early SSc, as SSc-specific autoantibodies are usually detected at disease onset or even before and do not change from one to another during the course of the disease, being almost always mutually exclusive. On the other hand, despite this high association with a specific disease subset, it has been demonstrated that anti-nuclear autoantibodies do not play a significant role in the pathogenesis of the disease and are considered only biomarkers of the disease and specific organ involvement [65].

In addition to SSc-specific autoantibodies, patients with SSc can also present autoantibodies that are also detected in other connective tissue diseases, such as anti-Ro60 and anti-IFI16 autoantibodies. These autoantibodies are denominated SSc-associated autoantibodies, and although their utility for diagnosis is limited due to their presence in other diseases, they have been associated with specific clinical subsets or manifestations in SSc patients. Usually, when referring to both SSc-specific and SSc-associated autoantibodies, the term SSc-related autoantibodies is used [66, 67].

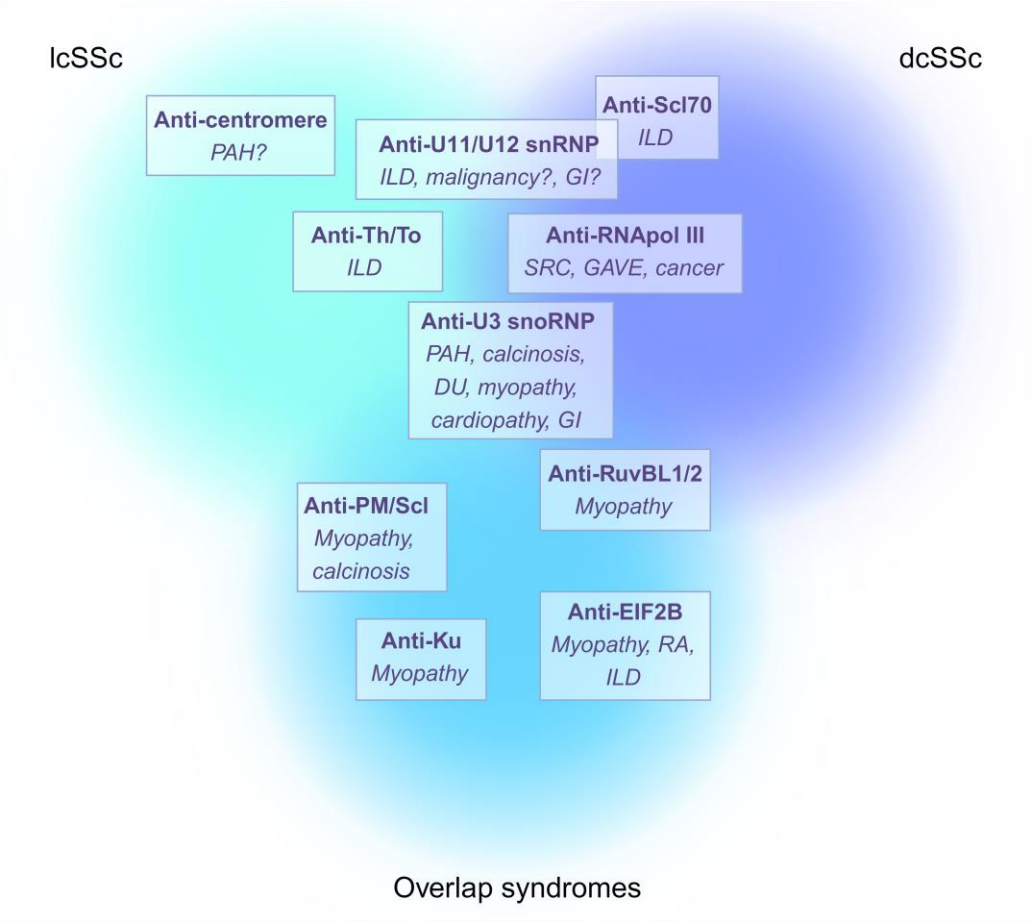


Figure 6. SSc-specific autoantibody association with disease subsets and clinical manifestations. SSc-specific autoantibodies are mutually exclusive and correlate with specific clinical phenotypes. *Created with BioRender.com*

Target proteins of SS-specific autoantibodies are often substrates for proteases that are associated with stress and cell death. Indeed, many autoantigens are substrates for caspases, which carry out selective protein degradation during apoptosis [68–73]. In line with this, it has been demonstrated that caspase cleavage sites seem to occur more commonly among proteins that become autoantigens than among cellular proteins as a whole. Moreover, an even stronger correlation was noted between autoantigenic potential and the presence of cleavage sites for Granzyme B, a protease that is produced by cytotoxic T lymphocytes and natural killer cells [74]. These observations have led to the hypothesis that peptides generated from products cleaved by caspases and Granzyme B contain cryptic epitopes capable of stimulating autoreactive T-cell responses [69, 72]. As Granzyme B is secreted by cytotoxic T lymphocytes and natural killer cells, which operate only in the periphery and not in the thymus, there should be no opportunity for negative selection of autoreactive T lymphocytes that are specific for Granzyme B-dependent immunocryptic epitopes [75].

On the other hand, autoantibodies recognising different constituents of the same macromolecular complex are common in systemic autoimmune diseases [76, 77]. “Linked sets” of autoantibodies may be a consequence of B cell surface immunoglobulin-mediated uptake and processing of antigenic particles. In this scenario, T cells responsive to a peptide derived from one component of a supramolecular complex may provide intermolecular and intra-structural help to B cells producing antibodies to other components [78–80]. Another factor that may influence the generation of linked autoantibody sets is the strength of interactions between components of an antigenic particle. In the second case, stabilising antibodies may play a role in the pathogenesis of linked autoantibody sets [81].

1.7.1. ANA detection methods

ANAs can be detected by three main types of laboratory methods: IIF, assays based on attaching specific autoantigens in solid phases, and assays that test sera reactivity against protein extracts in native conditions, such as immunodiffusion and immunoprecipitation (IP).

1.7.1.1. Indirect immunofluorescence for ANA detection

IIF using Human Epithelial type 2 (HEp-2) cells as a substrate is currently recommended as the gold-standard method for ANA screening. In this assay, fixed and permeabilised HEp-2 cells are incubated with patient serum; if an autoantibody that recognises an autoantigen is present in the serum, the autoantibody will get attached to the HEp-2 cell substrate. Subsequently, the substrate is incubated with a fluorochrome-conjugated antibody specific for human IgG's fragment crystallisable (Fc) portion. The autoantigens recognised by autoantibodies may be located in the nuclear or cytoplasmatic compartments using a fluorescence microscope. Due to the differential distribution of autoantigens in HEp-2 cells, different ANA patterns can be associated with distinct autoantibodies. In this line, HEp-2 cells are considered the best substrate as the initial identification of many different autoantibodies is facilitated because of the relatively big nucleus and the diversity of native humans’ antigens distributed over all cellular compartments (Figure 7) [65, 82].

Although ANA evaluation on HEp-2 cells is a relatively good screening test, a second assay is usually needed to detect the specific autoantigen recognised by as different autoantigens have the same subcellular localization and result in an identical IIF pattern. Another drawback of IIF is the subjective interpretation of the results, which in the past was exacerbated by a lack of a consensus to inform positive reports. The International Consensus on ANA Patterns (ICAP) initiative, which started in 2014, aimed to establish a

consensus on the nomenclature of the distinct ANA patterns recognised by IIF on HEp-2 cells. This effort has resulted in a classification that distinguishes, at present, 30 different ANA patterns enumerated as anti-cell (AC) 0 to 29, each with a specific staining (**Table 5**) [65, 82, 83].

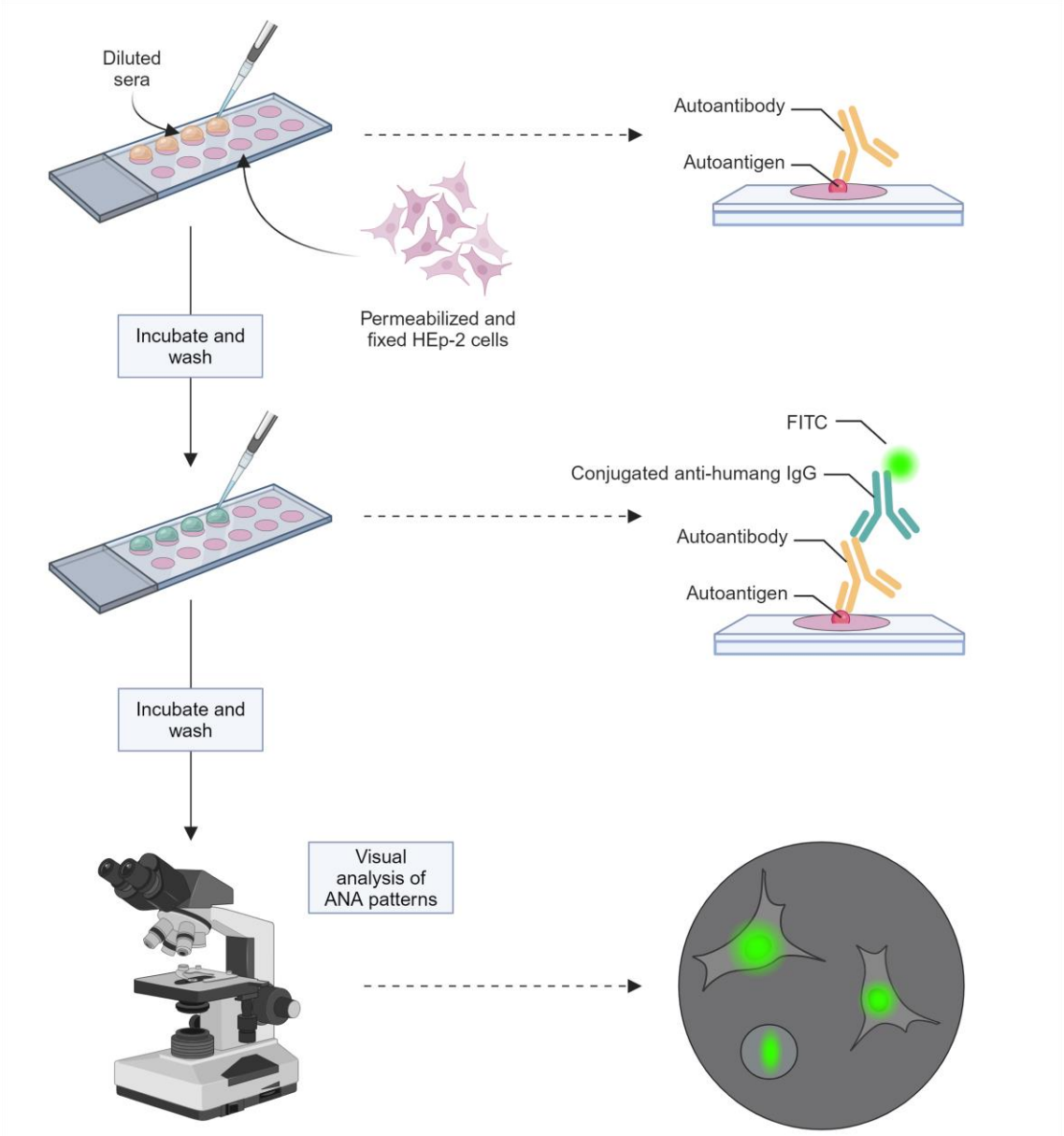


Figure 7. IIF for ANA detection. Permeabilised and fixed HEp-2 cells are incubated with diluted patient sera. Autoantibodies of the sera will bind autoantigens present in HEp-2 cells with high affinity. After a washing step, the slides are incubated with anti-human IgG antibody conjugated with fluorescein (FITC) fluorochrome. Using a fluorescence microscope, the distribution of the fluorescence and, hence, of the autoantibody and recognised autoantigen is visualized and analysed. *Created with BioRender.com*.

1.7.1.2. Methods based on solid-phase assays for ANA detection

Enzyme-linked immunosorbent assay (ELISA), chemiluminescence immunoassay (CLIA) and immunoblotting are based on the coating of a solid surface, as a plate well, bead or nitrocellulose strip, respectively, with a known autoantigen. Patient serum is incubated with the solid phase. If an autoantibody that recognises the specific autoantigen is present in the serum, the autoantibody will get attached to the solid phase. After incubation with an enzyme-conjugated antibody specific for the Fc portion of human IgG, a substrate for the enzyme is added, which is converted to a product that can be detected as colour (ELISA or Immunoblot) or light (CLIA) (Figure 8). More recently developed solid phase assays for the detection of ANAs include multiplex assays in which multiple autoantigens are coated in a solid surface, and autoantibody reactivity is confirmed by using a fluorochrome-conjugated antibody specific for the Fc portion of human IgG [65, 82].

There is a plethora of assays based on these methodologies. These assays have made it possible to detect autoantibodies with high throughput, automated, cost-effective, simple technologies, which is completely necessary for diagnostic routine laboratories. However, the lack of autoantigen standardisation has led to great variability in reported values, which is especially problematic for detecting autoantibodies targeting macromolecular complexes formed by different protein components. Moreover, it has been reported that cryptic epitopes that are not accessible in intact complex structures of native proteins became exposed when attached to a solid phase and may be recognised by autoantibodies that do not react against that protein *in vivo*, resulting in false positive results. On the contrary, some epitopes may get occult, resulting in false negatives, as have also been reported when using solid phase assays. In addition, unlike T cells, which recognise short continuous peptides associated with HLA molecules, B cells and antibodies recognise primarily discontinuous or conformational epitopes. Therefore, detecting autoantibodies by assays in which denatured proteins are used as an autoantigen source may be of limited value, resulting in false positives and false negatives due to the disappearance of native epitopes and the creation of aberrant epitopes. On the other hand, these assays are normally based on *in vitro* production and purification of the autoantigens, which can lead to contamination of the autoantigen of interest with other proteins or compounds. Indeed, in some cases, it has been reported that some solid phase assays detected reactivity against other components of the purified product and not the autoantigen of interest [65, 82, 84–89].

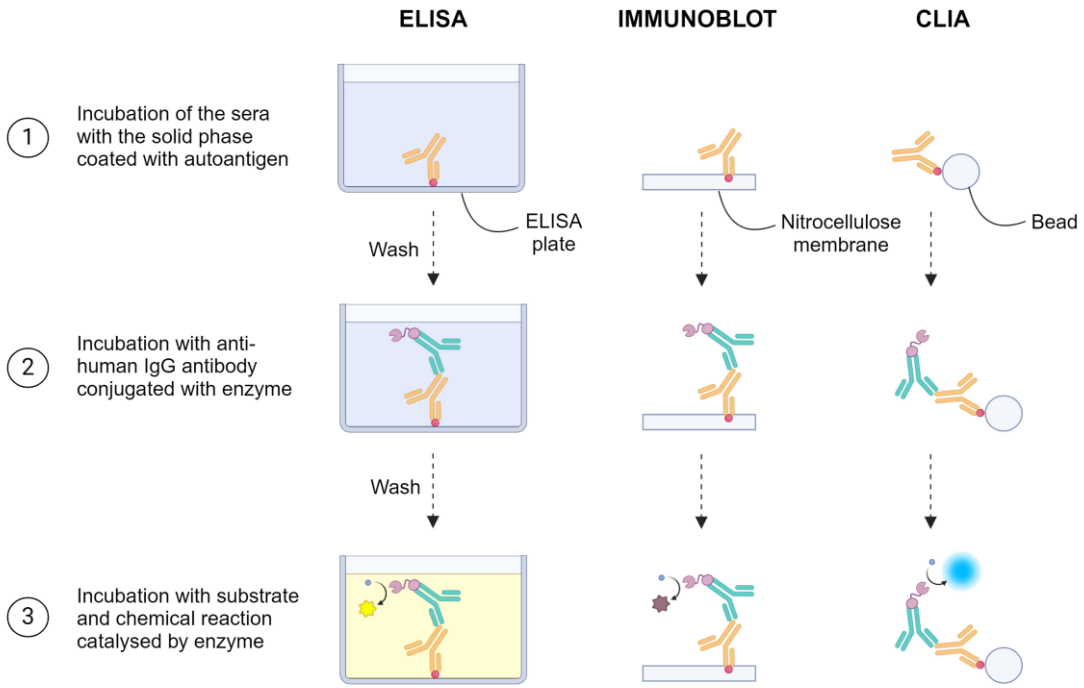


Figure 8. Graphical summary of solid phase methodologies for ANA detection. Autoantigens coated in ELISA plates, nitrocellulose strips and beads are incubated with patient sera. Autoantibodies of the sera will bind autoantigens present in the solid phase with high affinity. After a washing step, the solid phase is incubated with an anti-human IgG antibody conjugated with an enzyme. Finally, a substrate for the enzyme is added, which is converted to a product that can be detected as colour (ELISA or Immunoblot) or light (CLIA). *Created with BioRender.com*

1.7.1.3. Double Immunodiffusion

Immunodiffusion assays are based on precipitation reactions between antibodies and antigens at an appropriate ratio. In double immunodiffusion (Ouchterlony), two wells are carved into an agar surface so that they are only a few millimetres apart. Patient serum is added to one well and the protein extract (mainly calf or rabbit thymus or bovine spleen extracts) on the other, forming two concentration gradients that begin to overlap, such that, at the point where the concentrations of autoantibody and antigen reach the zone of equivalence, a visible precipitin line forms. A variation of the double immunodiffusion technique has been particularly useful for ascertaining whether two antigens share identical epitopes and for quickly assessing the purity of either antigen or antibody. Although, in this case, the native structure of autoantigens is maintained, and therefore, false positives and false negatives are rare, double immunodiffusion is a time-consuming assay by which it is impossible to test many samples simultaneously. Moreover, the specificity of the autoantibody is not properly determined, and positive samples always must be compared with positive controls [90–92].

Table 5. ANA patterns classification defined by ICAP.

AC number	Name	Description of observed staining by IIF using HEp-2 cells as substrate	Antigen association
AC-0	Negative	No specific staining detected.	None
AC-1	Nuclear homogeneous	Homogeneous and regular fluorescence across all nucleoplasm. The nucleoli may be stained or not. Mitotic cells* have the chromatin mass intensely stained in a homogeneous hyaline fashion.	dsDNA, nucleosomes, histones
AC-2	Nuclear dense fine speckled	Speckled pattern distributed throughout the interphase nucleus with characteristic heterogeneity in the size, brightness and distribution of the speckles. The metaphase plate depicts strong speckled pattern with some coarse speckles standing out.	DFS70
AC-3	Centromere	Discrete coarse speckles (40-80/cell) scattered in interphase cells and aligned at the chromatin mass on mitotic cells	CENP-B, CENP-A
AC-4	Nuclear fine speckled	Fine tiny speckles across all nucleoplasm. The nucleoli may be stained or not. In mitotic cells* the chromatin mass is not stained.	Ro-60, La, Ku, RuvBL1/2, Mi-2
AC-5	Nuclear large/coarse speckled	Coarse speckles across all nucleoplasm. The nucleoli may be stained or not. In mitotic cells* the chromatin mass is not stained.	U1 snRNP, Sm
AC-6	Multiple discrete nuclear dots	Countable discrete nuclear speckles (6-20/cell)	Sp100, PML, NXP-2
AC-7	Few discrete nuclear dots	Countable discrete nuclear speckles (1-6/cell)	SMN, coilin
AC-8	Homogeneous nucleolar	Diffuse fluorescence of the entire nucleolus, while the metaphase plate shows no staining.	PM/Scl, Th/To, nucleolin
AC-9	Clumpy nucleolar	Irregular staining of the nucleoli and Cajal bodies with a peri-chromosomal staining at the metaphase plates	Fibrillarin
AC-10	Punctate nucleolar	Densely distributed but distinct grains seen in the nucleoli of interphase cells. In metaphase cells, up to 5 bright pairs of the NOR can be seen within the chromatin body.	NOR90, RNApol I
AC-11	Smooth nuclear envelope	Homogeneous staining of the nucleus with greater intensity at its outer rim and no staining at the metaphase and anaphase chromatin plates. There is a peculiar accentuation of the fluorescence at the points were adjacent cells touch each other.	Lamin A, B and C, and lamin-associated proteins
AC-12	Punctate nuclear envelope	Nuclear envelope reveals a punctate staining in interphase cells, with accentuation of fluorescence at the points where adjacent cells touch each other. No staining of the metaphase and anaphase chromatin plates	gp210
AC-13	PCNA-like	Pleomorphic speckled nucleoplasmic staining, with variability in size and brightness of the speckles.	PCNA
AC-14	CENP-F-like	Nuclear speckled pattern with striking variability in intensity. The centromeres are positive only in prometaphase and metaphase, revealing multiple aligned small and faint dots.	CENP-F
AC-15	Cytoplasmatic fibrillar linear	This pattern is characterized by staining of cytoskeletal fibres spanning the long axis of the cells, sometimes with small, discontinuous granular deposits.	Actin, non-muscle myosin
AC-16	Cytoplasmatic fibrillar filamentous	Staining of microtubules and intermediate filaments spreading from the nuclear rim.	Vimentin, cytokeratins, tropomyosin

Table 5 (continued)

AC number	Name	Description of observed staining by IIF using HEp-2 cells as substrate	Antigen association
AC-17	Cytoplasmatic fibrillar segmental	Staining of short segments and periodic dense bodies along the stress fibres	α-actinin, vinculin
AC-18	Cytoplasmatic discrete dots	Staining of discrete dots that correspond to GW bodies in the cytoplasm of interphase cells, with high numbers in late S2/G2 cells	GW182, Ago2
AC-19	Cytoplasmatic dense fine speckled	Almost homogeneous staining of the cytoplasm, with cloudy appearance	PL7, PL12, EJ, OJ, KS, Zo, Ha, ribosomal P protein
AC-20	Cytoplasmatic fine speckled	Scattered small speckles in the cytoplasm mostly with homogeneous or dense fine speckled background	Jo-1
AC-21	Cytoplasmatic reticular	Coarse granular filamentous staining extending throughout the cytoplasm	Mitochondrial antigens
AC-22	Golgi	Discontinuous speckled or granular perinuclear ribbon-like staining with polar distribution in the cytoplasm	Golgi antigens
AC-23	Rods and rings	Distinct rod and ring structures in the cytoplasm of interphase cells.	IMPDH2
AC-24	Centrosome	One or 2 speckles in the cytoplasm and at the poles of mitotic spindles	Centriole antigens
AC-25	Spindle fibres	The spindle fibbers between the poles are stained in mitotic cells.	HsEg5
AC-26	NuMA-like	Nuclear speckled staining with additional staining of the spindle fibbers between the poles.	NuMA
AC-27	Intercellular bridge	Staining of the intercellular bridge that connect daughter cells by the end of cell division, but before cell separation	None
AC-28	Mitotic chromosomal	Punctate coloration of chromosomes in prometaphase and metaphase with no staining of interphase cells	Modified histone H3, MCA-1
AC-29	Topoisomerase I-like	Prominent fine speckled nuclear staining in interphase cells together with consistent strong fine speckled or homogeneous staining of condensed chromatin in mitotic presenting also strong staining of nucleolar organizing region. Variable nucleolar staining and weak cytoplasmic staining in interphase cells can also be observed.	DNA topoisomerase I

*Mitotic cells: cells in metaphase, anaphase and telophase. *https://www.anapatterns.org/index.php consulted on 19th August 2024.*

1.7.1.4. Immunoprecipitation

Protein and RNA IP are still considered the gold-standard techniques for detecting the specificity of the majority of ANAs due to the limited sensitivity and specificity of solid phase assays and the limited capacity of double immunodiffusion for detecting specific autoantibodies [93–95]. Both techniques rely on incubating patients' sera with protein extracts obtained from human cell lines subjected to radioactive metabolic labelling. Autoantibodies present in sera are attached to beads by the Fc region, while the antigen-binding fragment (Fab) region remains available to bind autoantigens of the protein extracts and cause them to immunoprecipitate. In the case of RNA-IP, the aim of the assay is to detect autoantibodies that immunoprecipitate ribonucleoprotein autoantigens from the cell extract. Ribonucleoproteins are macromolecular complexes that contain intrinsic small RNA molecules necessary for their function and are involved in many cellular processes [96]. After the immunoprecipitation, an acid nucleic extraction is performed, and specific RNA molecules that form part of the recognised ribonucleoproteins are obtained (**Figure 9**) [97].

Both in protein IP and RNA-IP, already described autoantibodies present in the sera are identified based on the molecular weight (MW) of the proteins and RNA molecules detected by polyacrylamide gel electrophoresis (PAGE) and by comparison with reference sera positive for known autoantibodies. Moreover, protein and RNA-IP also have the potential to identify new autoantibodies, as proteins and RNA molecules that are immunoprecipitated by not identified autoantibodies are also detected. For this reason, protein IP combined with mass spectrometry (MS) has long been and it is still the assay of choice for discovering new autoantibodies in SSc [89, 98]. It is also possible to identify new autoantibodies with RNA IP, by sequencing RNA molecules eluted from gels [99], but it is not usually performed due to its technical challenge. However, although protein and RNA IP are still considered the gold-standard assays for the detection of specific ANAs, it is very difficult for diagnostic routine laboratories to implement these techniques as they are time-consuming assays that can only be performed by highly trained personnel and rely on radioactive metabolic labelling, which requires legal authorisation and special waste disposal protocols.

1.7.2. Anti-centromere autoantibodies

The centromere is defined as the primary area of constriction in mitotic chromosomes containing the genetic locus necessary for the partitioning of chromosomes. Kinetochores are specialised protein macromolecular complexes found on the surface of centromeres, to

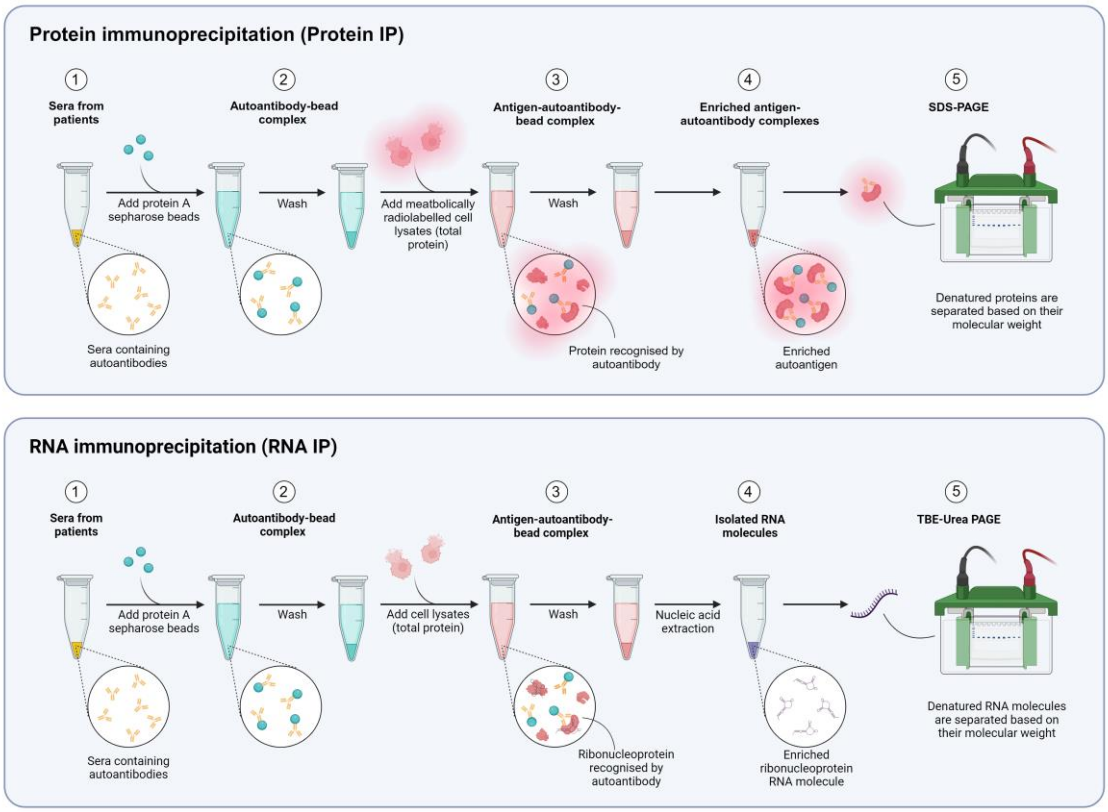
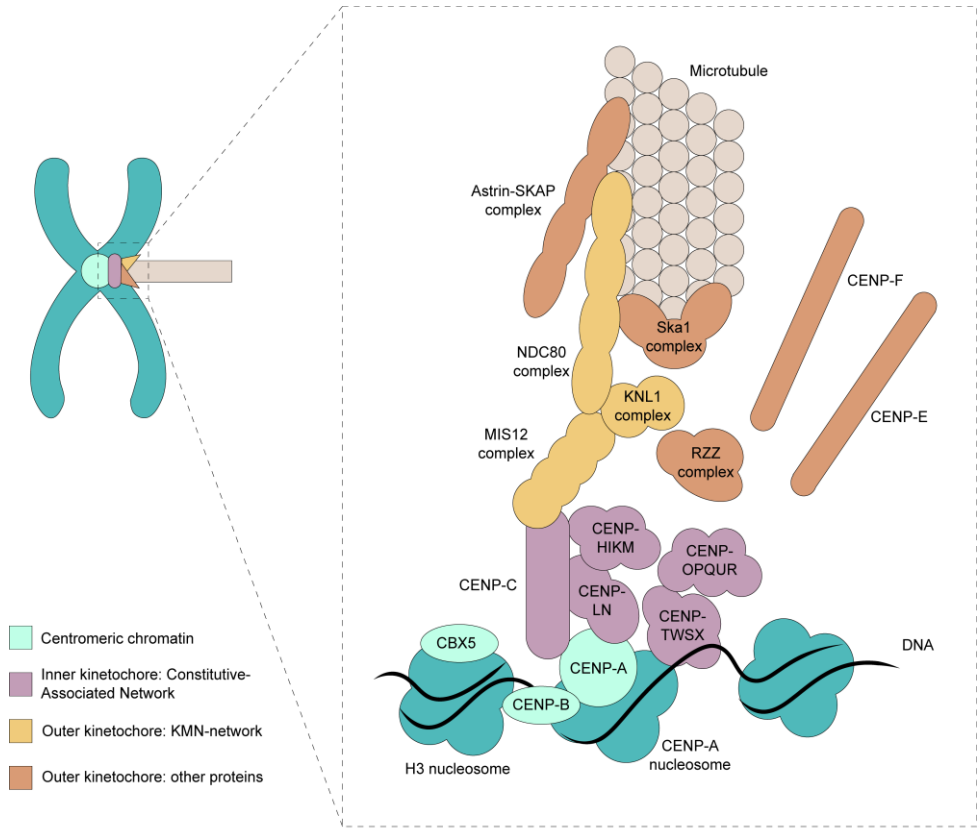


Figure 9. Graphical summary of protein and RNA immunoprecipitation methodologies. Patients' sera are incubated with beads coated with *Staphylococcus aureus* protein A, which binds IgG by the Fc region. Autoantibody-bead complexes are incubated with protein extracts obtained from human cell lines. Only autoantibodies will recognise proteins from the protein extracts by their Fab region and will get attached to the autoantibody-bead complex. In the case of RNA IP, the aim is to detect autoantibodies that immunoprecipitate ribonucleoprotein autoantigens from the cell extract. After immunoprecipitation, an acid nucleic extraction is performed, and specific RNA molecules that form part of the recognised ribonucleoproteins are obtained. Obtained protein and RNA molecules are separated by their molecular weight. *Created with BioRender.com*

which the spindle microtubules attach. Anti-centromere autoantibodies actually target proteins of the kinetochores, being centromere protein B (CENP-B) the major autoantigen, since it is recognised by virtually all sera with these autoantibodies. Moreover, autoantibodies against this protein are often present in titres orders of magnitude higher than those against other kinetochore proteins. Nevertheless, anti-centromere autoantibodies from SSc patients usually have specificities against more than one of the kinetochore antigens, being the second most autoantigenic centromere protein A (CENP-A). Finally, anti-centromere autoantibodies directed against other kinetochore proteins have also been described as centromere protein C (CENP-C) and chromobox protein homolog 5 (CBX5), but with a lower frequency (**Figure 10**) [100–102].

Anti-centromere autoantibodies were first described in 1980 by IIF. The IIF pattern of anti-centromere autoantibody positive patients (AC-3) normally shows discrete speckled

nuceoplasmic staining at the interphase and distinct centromeric dots that are visible for each chromosome pair in the metaphase plate, which reflects the localisation of kinetochore proteins. Furthermore, during interphase, the speckles are often seen around the nucleoli, as the location of CENP antigens is found within the nucleoli during this cell cycle phase [100, 102, 103].



Location	Antigen	Function
Centromeric chromatin	CENP-B	Centromere specific histone H3 variant that enables the generation of the centromere-specific nucleosome which serves as foundation for the kinetochore assembly
	CENP-A	DNA-binding domain specific for human α-satellite DNA, a family of highly repetitive 17bp DNA sequences localized exclusively at centromeres
	CBX5	Constitutes heterochromatin and interacts with MIS12 complex
Inner kinetochore (CCAN)	CENP-C CENP-HIKM CENP-TWSX CENP-OPQUR CENP-LN	Assembled on the centromeric chromatin serves as a platform for the outer kinetochore binding
Outer kinetochore (KMN-network)	KNL1 complex	Plays an important role in mitotic chekpoint control
	MIS12 complex	Links the inner kinetochore and NDC80C
	NDC80 complex	Binds to microtubules
Outher kinetochore (other proteins)	Astrin-SKAP	Interacts with NDC80 complex and stabilizes de kinetochore-microtubule binding
	CENP-E	Bind to microtubules and kinetochores and transport chromosomes along the microtubules
	CENP-F	Associated in kinetochore-microtubule binding and dynein regulation
	RZZ complex	Interacts with KNL1 complex and recruits dynein to kinetochore
	Ska1 complex	Is recruited by NDC80C and promotes the kinetochore-microtubule binding

Figure 10. Schematic representation of the structure of the kinetochore. Anti-centromere autoantibodies recognise proteins of the kinetochores, CENP-B being the major centromeric autoantigen. However, these autoantibodies have also been demonstrated to target other kinetochore proteins such as CENP-A, CENP-C and CBX5.

Anti-centromere autoantibodies are detected in 20-40% of SSc patients, being the most frequent autoantibody in the majority of SSc described cohorts, but it can also be detected in other autoimmune diseases as SjS and PBC, as already discussed. Anti-centromere autoantibodies are predictors of a more benign and protracted course, as they are highly associated with better survival and prognosis with less internal organ involvement. Specifically, it has been shown to be protective for the development of ILD, SRC and malignancy. In line with this, anti-centromere autoantibody is also associated with the lcSSc subset, as almost half of the patients within this subset present this autoantibody. On the other hand, anti-centromere autoantibody has traditionally been associated with PAH when compared to anti-Scl70 positive patients, the second most common autoantibody in SSc, who rarely present PAH. However, overall it has been described that only 10-20% of anti-centromere autoantibody positive patients present PAH, and some recent studies with large cohorts have not found an increased prevalence of PAH in this group of patients when compared with the frequency of this clinical characteristic in the whole cohort, or even when were compared with anti-Scl70 autoantibody positive patients [64, 65, 101, 102, 104].

1.7.3. Anti-Scl70/Topoisomerase I autoantibodies

DNA topoisomerases are a heterogeneous group of enzymes that change the tertiary structure of DNA, playing a central role in the transcription and replication of DNA. Topoisomerases perform their function either by relaxing supercoiled DNA through breaking and rejoining one strand at a time (type I enzymes) or by catalysing catenation/decatenation, knotting/unknotting of DNA rings through breaking and rejoining DNA in double-stranded fashion (type II enzymes). Specifically, topoisomerase I is a type Ib topoisomerase that transiently breaks the phosphodiester backbone of a single strand in the DNA duplex by binding covalently to a 3'-phosphate, followed by

resealing free DNA ends after a rotational event, acting as a swivel to relax both negative and positive supercoils. Moreover, a direct involvement of the enzyme in ribosomal RNA (rRNA) gene transcription has also been shown. In fact, the topoisomerase I is partitioned between nucleoli, specifically in the dense fibrillar component (DFC), and nucleoplasm during interphase. Significantly, the topological maintenance of rDNA by topoisomerase I may be required for the proper folding of rDNA and RNApol I into functional ribosomal transcriptional units (see section 1.7.5.). The distribution of the enzyme gives rise to a specific IIF pattern (AC-29) that is characterised by a fine specked nuclear staining and variable punctate nucleolar or perinucleolar staining during interphase together with strong

staining of chromatin and nucleolar organiser regions (NOR) in mitotic cells. Moreover, a weak cytoplasmic staining can be detected radiating from the perinuclear area towards the plasma membrane, which could be caused due to the significant homology between topoisomerase I and mitochondrial topoisomerase I [105, 106].

Anti-topoisomerase I autoantibodies were originally discovered by immunoblotting of sera of SSc patients reacting with a 70kDa nuclear protein from cell extracts and were termed consequently as anti-Scl70 autoantibodies. Later it was shown that this 70kDa protein was actually a breakdown product of DNA topoisomerase I, a 100kDa protein, produced during purification procedures. Anti-Scl70 autoantibodies are the second most common autoantibodies detected in SSc patients, with a prevalence of 20-35%, and contrary to anti-centromere autoantibodies, they have been associated with the dcSSc subset. However, anti-Scl70 autoantibodies are not restricted to the dcSSc subset, as they have also been found to be positive in lcSSc patients (25-48% of positive patients). Indeed, it has been recently proposed that anti-Scl70 autoantibody positive patients with dcSSc or lcSSc actually represent two different subsets of patients: patients with dcSSc and anti-Scl70 autoantibodies present poor prognosis, with increased mortality, ILD, proteinuria, musculoskeletal and cardiac involvement, while lcSSc patients with anti-Scl70 autoantibodies only present a high incidence of ILD, but other complications are rare, and present higher survival. While anti-Scl70 autoantibodies are considered to be present almost exclusively in SSc, the presence of anti-topoisomerase IIa autoantibodies has been demonstrated in a number of other autoimmune disorders in addition to SSc, including SLE, juvenile RA and insulin-dependent diabetes mellitus [64, 100, 105].

1.7.4. Anti-RNApol III autoantibodies

Transcription of the eukaryotic genome is mediated by three highly specialised nuclear DNA-dependent RNA polymerase (pol) enzymes [107, 108].

1. RNApol I synthesises a single transcript, the 47S pre-rRNA.
2. RNApol II synthesises a wide diversity of transcripts, including protein-coding messenger RNA (mRNA) and many non-coding RNAs (ncRNAs), such as micro RNAs (miRNA), small nuclear (snRNA), and small nucleolar RNAs (snoRNA).
3. RNApol III synthesises diverse transfer RNA (tRNA) and 5S rRNA, and also U6 snRNA and other non-coding small RNAs as the RNA components of RNase P and RNase for mitochondrial RNA Processing (MRP) endoribonucleases, the 7SL, vault

RNAs, Y RNAs, 7SK RNA, BC200RNAs, some virus-encoded RNAs (Adenovirus, Human Papillomavirus and Epstein-Barr virus RNA) and short interspersed repeated DNA-elements-encoded RNAs (SINEs) [109, 110].

RNApol I, II, and III contain 14, 12, and 17 subunits, respectively, as all three polymerases contain a ten-subunit catalytic core with additional peripheral subunits (**Table 6**, adapted from [111, 112]). RNA polymerases induce transcription by the assembly of initiation complexes on gene promoters. During initiation, the general transcription factors recognise promoter elements, recruit and orient the polymerase, and assist the polymerase in DNA opening and initial RNA synthesis [111].

Table 6. Molecular weight and expected function of protein constituents of the polymerase core of RNA polymerase I, II and III.

RNApol I		RNApol II		RNApol III		
Protein	MW (kDa)	Protein	MW (kDa)	Protein	MW (kDa)	**Protein function**
A190	194.8	Rpb1	217.2	C160	155.6	Active centre
A135	128.2	Rpb2	133.9	C128	127.8	Active centre
AC40	39.3	Rpb3	31.4	AC40	39.3	-
AC19	15.2	Rpb11	13.3	AC19	15.2	-
A12.2	13.9	Rpb9	14.5	C11	12.3	RNA cleavage
Rpb5	24.6	Rpb5	24.6	Rpb5	24.6	-
Rpb6	14.5	Rpb6	14.5	Rpb6	14.5	-
Rpb8	17.1	Rpb8	17.1	Rpb8	17.1	-
Rpb10	7.6	Rpb10	7.6	Rpb10	7.6	-
Rpb12	7.0	Rpb12	7.0	Rpb12	7.0	-

Autoantibodies directed to the three RNA polymerases have been detected in SSc patients by protein IP. The two largest subunits of human RNA polymerases are readily distinguishable by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), but the smaller subunits are less well defined and more difficult to visualise. Therefore, sera with anti-RNApol I, II and III autoantibodies are defined by IP of the two largest subunits of each RNA polymerase, with a weight of 190 and 128kDa, 220 and 138kDa, and 155 and 135kDa, respectively [113]. Frequently, sera from SSc patients present reactivity not only to one of the RNA polymerases but to two or all three of them. Indeed, several subunits are identical in two or all three of the enzyme complexes, and although some of the subunits are unique to each enzyme, they are similar regarding both sequence and conformation. Therefore, a single autoantibody may recognise either a unique or a shared RNA polymerase epitope or subunit and lead to immunoprecipitation of one, two or all three enzymes [114].

Three main groups of sera have been reported depending on their reactivity against RNA polymerases: anti-RNApol I/III sera, anti-RNApol I/II/III sera, and anti-RNApol II sera [89, 113–118]. Nevertheless, on rare occasions, IP patterns associated with other combinations of anti-RNApol recognition can also be observed. The first two groups are considered SSc specific, while anti-RNApol II autoantibodies have also been observed in other systemic autoimmune diseases such as SLE and overlap syndromes with or without SSc clinical characteristics and often accompanied by other autoantibodies [118]. Interestingly, anti-RNApol II autoantibodies in SSc are highly associated with anti-Scl70 autoantibodies when found without anti-RNApol I and III autoantibodies [114, 116, 117]. On the other hand, the bigger subunit of RNApol III has been shown to be the most antigenic subunit targeted by the vast majority of anti-RNA Pol I/III and anti-RNA Pol I/II/III sera [119]. Therefore, all the commercial assays that have been developed for the detection of these autoantibodies have used this antigen or its fragments, and therefore the majority of studies referring to clinical associations of autoantibodies against RNA polymerases are only reflecting positivity against RNApol III [120–122].

Anti-RNApol I/II/III and anti-RNApol I/III positive sera produce heterogeneous patterns by IIF. Typically, these sera produce a fine speckled nucleoplasmic stain with additional occasional bright dots. However, pure anti-RNA Pol I sera produces a punctate nucleolar staining in interphase cells, and several dots in the area of the condensed chromosomes in metaphase cells due to its localisation around NORs (AC-10) (see section 1.7.5.) [89, 113, 115, 123–125]. Moreover, the IIF staining patterns of some of these sera vary when diluted, probably due to different titres of specific autoantibodies against different subunits of RNA polymerases [119]. Indeed, part of the observed heterogeneity can be attributed to the fact that RNA polymerase I resides in nucleoli while RNA polymerase II and III are located in nucleoplasm [115, 126].

In Caucasians, a prevalence of anti-RNApol III of 5-20% has been reported, but this autoantibody is less frequently detected in Japanese and black patients. The presence of this autoantibody is highly associated with the dcSSc subset, although almost half of anti-RNA pol III positive patients may present the limited cutaneous form of the disease. Moreover, this autoantibody is associated with a decreased risk of developing ILD but a more extensive and rapid skin involvement and a higher risk of SRC, GAVE and synchronous malignancy [64, 113, 127–129].

Regarding the association of this autoantibody with malignancy (see section 1.5.9.), it has been estimated that patients with anti-RNApol III autoantibodies present an odds ratio (OR)

of 5-10 of having a diagnosis of cancer within an interval between 6 months before and 3 years after SSc onset, as compared with matched anti-RNApol III negative patients. Moreover, the prevalence of malignancy in this subset of patients is around 15% compared to an overall 10% prevalence in SSc patients [127, 129]. The close temporal relationship between cancer and SSc among patients positive for this autoantibody led to the hypothesis that SSc could represent a paraneoplastic disorder in this subset of patients. In this line, an increased nucleolar expression of RNApol III protein was found in tumour cells from patients with anti-RNApol III autoantibodies compared to cancer cells of those negative for this autoantibody and healthy control tissues. Moreover, somatic genetic alterations and loss of heterozygosity of the *POLR3A* gene locus, encoding RNApol III, was reported. These genetic alterations were found in a low fraction of neoplastic cells, suggesting that immunoediting of the neoplasm occurs, with these alterations being selected against during tumour growth. On the other hand, anti-RNApol III autoantibodies of patients with somatic mutations were demonstrated to cross-react with both the mutated and wild-type RNApol III protein. Considering all these data, it has been speculated that malignancies harbouring *POLR3A* mutations trigger SSc in most patients with anti-RNApol III autoantibodies. However, in the majority of these patients, the immune response eradicates the cancer by the time SSc develops since cancer is detected in only 15-20% of these patients [128, 130]. On the other hand, subsequent studies have demonstrated that patients presenting anti-RNApol I autoantibodies together with anti-RNApol III autoantibodies don't present the high association with malignancy as patients with isolated anti-RNApol III autoantibodies, indicating that there could be more than one physiological mechanism for the emergence of these autoantibodies [131]. These data could help to decide in which anti-RNA pol III positive patients intensive cancer detection strategies are needed [53–55].

1.7.5. Autoantibodies against protein components of the nucleolus

The nucleoli are the largest visible structures inside the nucleus and represent functionally and biophysically distinct compartments. Due to the absence of a delimiting membrane, nucleolar components are highly dynamic, exhibiting continual flux within the nucleolus and exchange with the surrounding nucleoplasm. In fact, nucleoli are fundamental for regulating cell homeostasis through the transient sequestration of key cell cycle regulators [132]. Moreover, it has been shown that the nucleoli also play important roles in sensing diverse stresses, including genotoxic and oxidative stress, heat shock, nutrient deprivation, oncogene activation and viral infection [133].

Despite being a membraneless organelle, the nucleolus normally exhibits a roughly spherical shape, although the nucleolar shape, size and number per cell nucleus change upon stress, viral infection or cancer [133–136]. In line with this, dozens of proteins exhibit a “rim-like” co-localization at the nucleolar surface, representing a quasi-two dimension nucleolar surfactant [137]. In addition, the nucleolus is often surrounded by a ring of condensed chromatin, known as perinucleolar chromatin, corresponding to inactive chromosomal regions [138]. Sometimes, projections of the perinucleolar chromatin may reach deeply inside the structure of the nucleolus, where they appear as nucleolar interstices or vacuoles [139].

The main function of the nucleoli is ribosome biogenesis. This process involves ribosomal DNA (rDNA) transcription, processing of rRNA and assembly of ribosomal proteins [140]. rDNA genes are arranged in arrays of head-to-tail tandem repeats clusters termed NORs which are located on the short arm of acrocentric chromosomes (13, 14, 15, 21 and 22) [141, 142]. Nucleoli form around these tandemly repeated clusters of rDNA, where they are transcribed by RNApol I to produce a pre-rRNA precursor transcript named 47S. This initial pre-rRNA transcript undergoes modification and processing to remove external transcribed spacers (ETSs) and internal transcribed spacers (ITSs) to yield mature 18S, 5.8S, and 28S rRNAs (Figure 11) [143, 144].

During interphase RNApol I complexes are mainly found at nucleoli, around NORs, guiding the rRNA transcription together with human upstream binding factor (hUBF), other transcription factors such as selective factor 1 (SL1), and topoisomerase I. At prophase, nucleoli disperse and disappear, but at least a portion of some nucleolar components, such as RNApol I, SL1, hUBF, and part of topoisomerase I, remain associated with NORs, while other components of the nucleoli are released, such as small nucleolar ribonucleoproteins (snoRNPs). Only a subset of NORs that are associated with RNApol I transcription machinery are transcriptionally active during interphase. By contrast, rDNA repeat units in inactive NORs are highly methylated and are not associated with the RNA Pol I machinery. During mitosis, those NORs that were transcriptionally active in the previous interphase form prominent chromosomal features termed secondary constrictions in which the chromatin is ten times less condensed than surrounding chromosomal regions. [145]. However, as little or no rRNA is synthesised during mitosis, it is hypothesised that RNApol I must be in a state of arrested transcription throughout most of metaphase. Beginning in telophase, RNApol I transcription resumes, and nucleoli begin to reform around individual active NORs, forming

multiple small nucleoli. Nucleolar fusion then occurs, resulting in the formation of larger mature nucleoli containing multiple active NORs [146, 147].

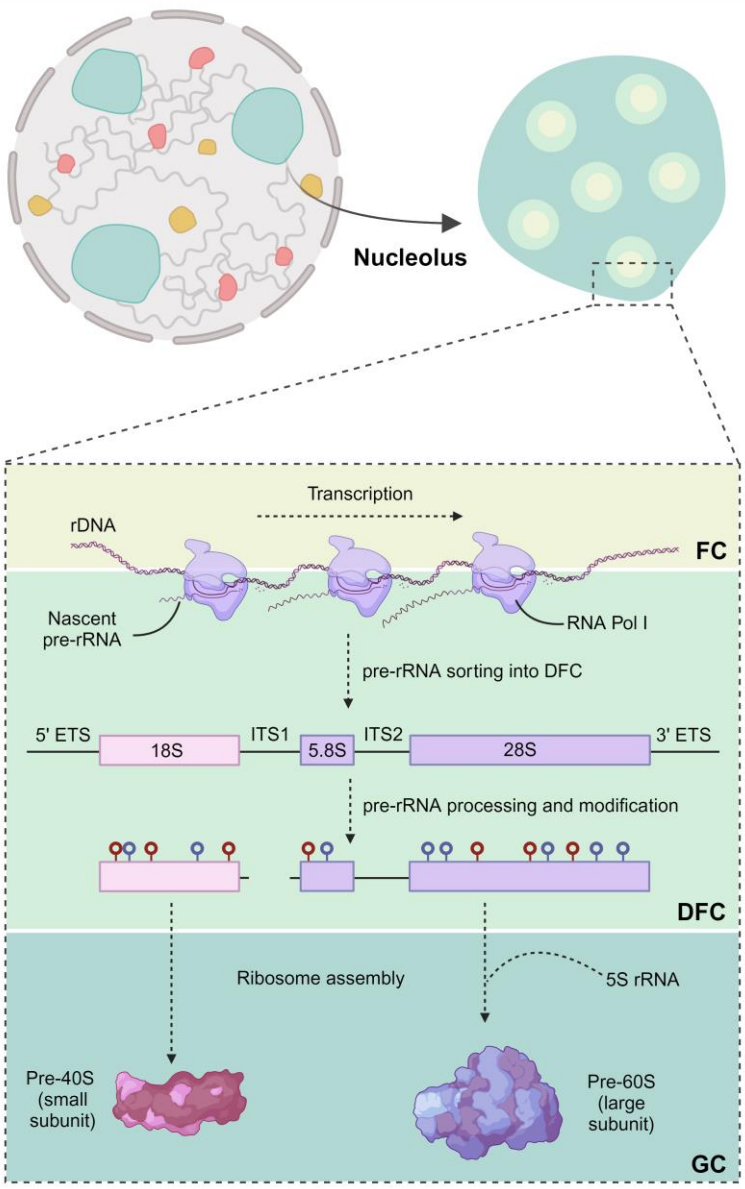


Figure 11. Ribosome biogenesis process. Ribosome biogenesis is a vectorial process, initiating in the FC and proceeding outward to the GC. rRNA transcription occurs at the FC/DFC interface where RNApol I complexes are enriched. Pre-rRNA transcript is sorted into the DFC., where it is excised in different molecules and pseudouridylation and 2'-O-methylation of around 200nt is catalysed by different snoRNPs. The processing of the pre-rRNA finally leads to the splitting of the initial 90S particle into pre-40S (small subunit of the ribosome) and pre-60S (large subunit of the ribosome) particles that are constituted in the GC. Created with *www.BioRender.com*

Once constituted, nucleoli are constituted by three distinct subcompartments that can be differentiated by electron microscope. The innermost fibrillar centres (FCs) are surrounded by DFCs, while FC/DFC compartments are in turn surrounded by granular components

(GCs) (**Figure 11**). Indeed, super-resolution microscopy has revealed that each human nucleolus consists of several dozen FC/DFC modules, the number of which can vary between cell types [148].

Current evidence suggests that ribosome biogenesis is a vectorial process, initiating in the FC and proceeding outward to the GC. FCs contain transcriptionally quiescent rDNA and act as a storage place for enzymes involved in rRNA transcription (RNApol I, topoisomerase I, hUBF). In this line, it has been confirmed that rRNA transcription occurs at the FC/DFC interface where RNApol I complexes are enriched, with each FC/DFC module containing two or three transcriptionally active rRNAs. RNApol I transcription leads to the synthesis of 47S pre-RNA, which is sorted to the DFC. A subset of ribosomal proteins associates with the 47S pre-rRNA along with numerous transacting factors, thus forming a large ribonucleoprotein particle or the 90S pre-ribosome. Within this particle and still in the DFC, the ETS and ITS are eliminated through endonucleolytic and exonucleolytic cleavages, leading to the splitting of the 90S particles into pre-40S and pre-60S particles, that are finally assembled within the GC (**Figure 11**). These particles are ultimately exported to the cytoplasm, where the last maturation steps take place [100, 149–152].

snoRNPs also localise in the nucleoli and catalyse pseudouridylation and 2'-O-methylation of ~200nt of pre-rRNA co-transcriptionally and post-transcriptionally (**Figure 11**). These modifications are important for ribosome biogenesis and translational accuracy. snoRNPs are assembled around guide RNAs, snoRNAs [153]. snoRNAs are 60 to 300nt long ncRNAs that are synthesised in the nucleoplasm but accumulate in the nucleolus. However, snoRNPs are first transported to Cajal bodies (CBs) before being routed to nucleoli, as it seems that they may interact with CB's proteins to end their assembly process [154]. Based on the presence of two types of conserved motifs denominated C/D and H/ACA box, that are necessary and sufficient for nucleolar localisation, snoRNAs are classified into two subsets. While methylation of the 2'-hydroxy groups of the riboses is directed by C/D box snoRNAs, conversion of uridines to pseudouridine is guided by box H/ACA snoRNAs [153, 154].

Both C/D box and H/ACA box snoRNAs are usually derived from introns. After the splicing reaction, introns are excised as lariats, which are then generally opened by the debranching enzyme and subsequently degraded. C/D box snoRNAs escape this degradation by forming a protein complex that consists of non-histone chromosome protein 2-like 1 (NHP2L1), nucleolar protein 56 (NOP56), nucleolar protein 58 (NOP58) and fibrillarin. Once the C/D box snoRNP is constituted, the antisense boxes of the snoRNA recognise complementary sequences in rRNA by base-pairing, while the 2'-O-methylation reaction is catalysed

specifically by fibrillarin, a S-adenosyl-L-methionine-dependent methyltransferase (SAM) [155, 156]. On the other hand, H/ACA box snoRNAs escape from degradation by assembling into a protein complex containing the pseudouridine synthetase dyskerin, and the structural proteins H/ACA ribonucleoprotein complex subunit 1 (GAR1), 2 (NHP2) and 3 (NOP10). Mature H/ACA box snoRNPs bind to rRNAs which allows recognition of the substrate uridine that is isomerised to pseudouridine by dyskerin [157].

In addition to the ribosome, other ribonucleoprotein particles are assembled, at least in part, in the nucleolus. This is notably the case for the signal recognition particle (SRP), involved in protein secretion and targeting to the endoplasmic reticulum of transmembrane proteins, and for telomerase, required for the maintenance of chromosome ends [158, 159]. There have also been reports of precursors of tRNAs, and even of mRNAs, transiting through the nucleolus [160]. Although it is not entirely clear why non-rRNAs transit through the nucleolus during their life cycle, it is largely assumed that they do so to benefit from the abundant assembly and modification machineries present there [139].

1.7.5.1. NOR90

Anti-NOR90 autoantibodies were first identified using sera from SSc patients that stained multiple discrete dots within the nucleoli and that recognised a nuclear protein doublet of approximately 90kDa in immunoblots from cell extracts. It was later demonstrated that these sera selectively stained NORs. Finally, the autoantigen recognised by anti-NOR90 autoantibodies was identified as the human RNApol I transcription factor hUBF, which exists as spliced variants of 97 and 94kDa and is involved in the regulation of the transcription of rRNA genes. Specifically, NOR90 associates with SL1 protein to form a stable pre-initiation complex within the core rDNA promoter region (see section 1.7.5.) [100].

Although autoantibodies against NOR90 were originally described in patients with SSc, it has been shown that they are not specific for the disease, as they can also be detected in patients with SjS, RA, undifferentiated connective tissue disease (UCTD) and malignancies. Moreover, no specific association with a particular clinical feature and this autoantibody in SSc has been demonstrated. Nevertheless, the lack of any strong association with clinical manifestations can be due to the rarity of this autoantibody, as it has been reported to have a very low prevalence of less than 3% [161, 162].

1.7.5.2. Anti-Th/To autoantibodies

In 1982, *Hardin et al.* observed that a serum specimen from a patient with SLE, termed "anti-Th", immunoprecipitated a small RNA denominated 7-2 RNA in addition to the Y RNAs normally immunoprecipitated by anti-Ro60/La autoantibodies [163]. In 1983, *Reddy et al.* independently detected autoantibodies immunoprecipitating 7-2 RNA and 8-2 RNA in a serum from an SSc patient designated as "anti-To" [164]. Further studies showed that the majority of sera co-immunoprecipitated both RNAs [165–169], and the anti-Th/To term was coined. Moreover, it was demonstrated that anti-La autoantibodies immunoprecipitated an immature and slower migrating form of 7-2 RNA [167], probably accounting for the reactivity observed by *Hardin et al.* [163]. Later studies identified 7-2 RNA and 8-2 RNA as the RNA components of RNase MRP and RNase P, respectively [168, 169].

RNase P is the endoribonuclease that removes 5' leader sequences from tRNA precursors, an essential step in tRNA maturation [170]. This enzyme is a ribonucleoprotein based on a catalytic RNA molecule (H1 or 8-2 RNA). The protein moiety of human RNase P is composed of a singular protein Pop1 and three subcomplexes, the Rpp20-Rpp25 heterodimer, Pop5-Rpp14-$(Rpp30)_2$-Rpp40 heteropentamer, and Rpp21-Rpp29-Rpp38 heterotrimer (**Figure 12**) [171]. Although the RNA component of human RNase P has been identified in the cytoplasm, nucleoplasm, perinucleolar compartment and nucleolus, the protein subunits of this complex are contained mainly in the nucleolus [172, 173]. Moreover, different studies have shown that gene clusters of some tRNA families localise in the nucleoli and that the 5' leader sequence removal happens in the same nuclear compartment [174, 175].

RNase P has long been recognised to share most of its protein components with another essential ribonucleoprotein, RNase MRP [176]. The RNase MRP was originally identified due to its ability to process an RNA transcript complementary to the light strand to generate RNA primers for the heavy-strand DNA replication in mitochondria, hence its name [177]. However, subcellular partitioning and *in situ* hybridisation experiments showed the presence of RNase MRP's RNA in both mitochondria and nucleoli, being the vast majority localised in the nucleoli [178]. Moreover, subsequent studies have demonstrated that mitochondrial RNase MRP and nuclear RNase MRP are distinct enzymes with differing activities and protein components but a common RNA subunit (Th or 7-2 RNA) [179]. In fact, nuclear RNase MRP is a non-canonical snoRNP that belongs to neither the C/D box nor H/ACA box classes and is involved in pre-rRNA processing [180–182]. In addition to this housekeeping function, more specific functions have been identified [182–184]. There is still controversy

about the exact protein composition of human RNase MRP, and consequently, its structural organisation is not clear, but it is clear that it shares most of its protein components with RNase P and probably presents a similar structure, as observed in order organisms (**Figure 12**) [176]. Moreover, although the RNA components of RNase P and RNase MRP are poorly conserved at the sequence level, their secondary structures, according to both experimental and phylogenetic data, suggest that they fold into similar cage-shaped structures [185, 186].

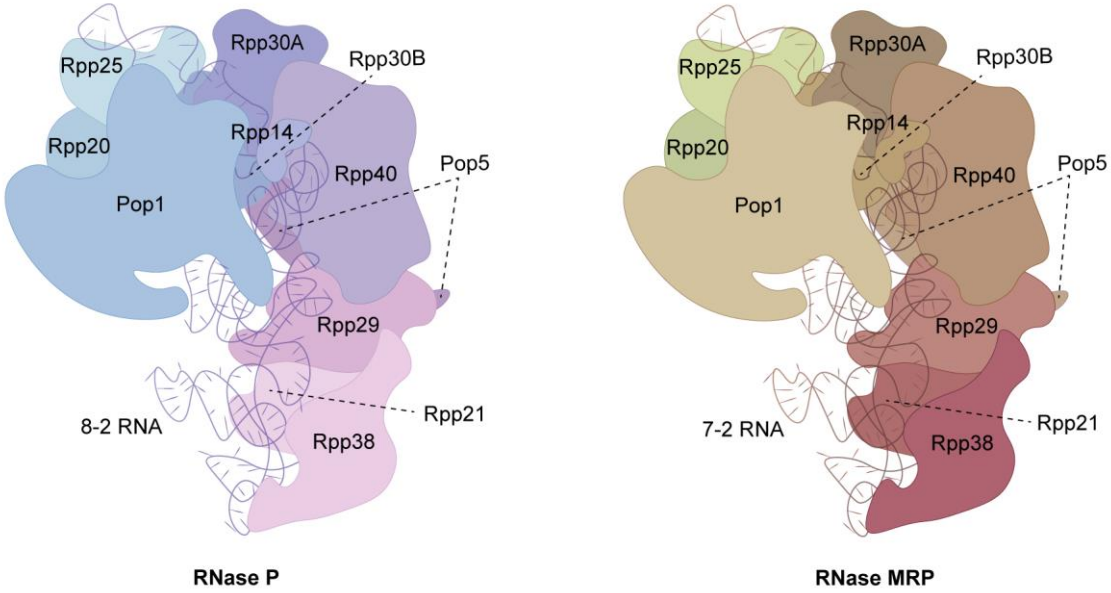


Figure 12. Schematic illustration of RNase P and RNase MRP structures. RNase P and RNase MRP are two different enzymes that share most of their protein components but present different associated RNA molecules, 8-2 and 7-2 RNA, respectively, that guide their catalytic activity.

Several protein components of RNase P and RNase MRP have been demonstrated to be the target anti-Th/To autoantibodies, including Rpp30, Pop5, Rpp14, Pop4, Rpp21, but the majority of patients recognise either Rpp38, Pop1 or Rpp25 [165, 187–192] Therefore, commercial assays using recombinant Pop1 [122, 193] and Rpp25 [192] proteins as antigens have been developed for the detection of anti-Th/To autoantibodies. These assays are specific enough, but their sensibility is not as high as expected because these assays do not recognise autoantibodies directed against other subunits of the RNase P or RNase MRP complexes. In fact, commercial assays using Pop1 and Rpp25 have been shown to detect only 50% and 53-63% of anti-Th/To autoantibodies compared to RNA IP, respectively [95, 192].

When the gold standard method, RNA IP, is used as a detection assay, the prevalence of anti-Th/To autoantibodies in SSc patients ranges between 0.8% and 4.6% [166, 194–197]. SSc patients with anti-Th/To autoantibodies were first shown to be associated with the lcSSc

subset and have a significantly lower total skin score, lower frequencies of joint contractures and palpable tendon friction rubs, but significantly higher frequencies of telangiectasias, subcutaneous calcification, hypothyroidism, puffy fingers, and small bowel involvement [194]. Nevertheless, most of these clinical features are considered to be correlated with lcSSc, rather than with the presence of anti-Th/To autoantibodies. When compared with anti-centromere autoantibody positive patients, anti-Th/To autoantibody positive patients have been reported to have more subtle cutaneous, sicca, vascular and gastrointestinal involvement (oesophageal dysmotility), but more often have symptoms commonly seen in dcSSc, as ILD, as well as reduced survival [166, 194, 198].

Although anti-Th/To autoantibodies are highly associated with SSc, they are not totally specific. Anti-Th/To autoantibodies have also been found at 4.0% frequency in Japanese patients with localised scleroderma, [199], similar to the prevalence observed in Japanese SSc patients [195]. Moreover, these autoantibodies are also found in some patients with primary RP [194] and other systemic autoimmune diseases [195] but with a much lower frequency. However, it has been shown that anti-Pop1 autoantibodies are highly specific of SSc, whereas reactivity to other proteins of the RNase P and RNase MRP complexes can also be seen in other diseases [195].

As discussed, anti-Th/To autoantibodies can be directed to different subunits of the RNase P or RNase MRP complexes, due to the subcellular localisation of both complexes, these sera always show homogeneous nucleolar staining (AC-8). In some cases, this nucleolar staining can appear associated with a nuclear fine speckled pattern [164, 166, 194, 199]

1.7.5.3. Anti-U3 snoRNP/fibrillarin autoantibodies

Anti-U3 snoRNP autoantibodies were first described when the serum from a SSc patient was shown to be able to immunoprecipitate U3 snoRNA [164]. U3 snoRNP is the most abundant C/D box snoRNP. but it does not methylate pre-RNA molecules during ribosome biogenesis like most of these ribonucleoproteins (see section 1.7.5.). Instead, U3 snoRNP guides endoribonucleolytic processing of the 5' ETS of the 47S pre-rRNA and plays a key role in the maturation of 18S rRNA (**Figure 11**) forming part of a bigger complex denominated small-subunit (SSU) processome [200]. SSU has been considered a ribosome assembly intermediate in some reports as it is associated with a subset of ribosomal proteins [201].

After the discovery of U3 snoRNA immunoprecipitation, it was demonstrated that the majority of sera presenting reactivity against U3 snoRNP recognized a 34kDa protein [202] and due to its localisation in FC/DFC regions of the nucleolus, it was named "fibrillarin" [203].

As previously mentioned, fibrillarin is a SAM that constitutes the catalytic subunit of C/D box snoRNPs [352, 353]. Consequently, fibrillarin is prominently found in the nucleoli during interphase, mainly in the transition FC/DFC area and in the DFC [196], although it can also be observed in CBs [354, 355]. In prophase, when the nucleolus is dispersed, fibrillarin is dispersed to the chromosomal periphery and remains there until anaphase [356, 357]. Due to its localisation during different cell cycle stages, when tested by IIF, anti-fibrillarin autoantibodies show a specific clumpy nucleolar staining (AC-9). This pattern consists of nucleolar and CB staining without nucleoplasmic staining in the interphase cells and reticular staining in the periphery of chromosomes on metaphase cells [125, 204].

On the other hand, although first studies reported that the majority of anti-U3 snoRNP autoantibodies recognised fibrillarin and U3 snoRNA, subsequent studies have shown that these autoantibodies can also immunoprecipitate other protein components [125, 187, 202, 205] and other RNA molecules [206], probably due to the association of fibrillarin with different proteins and snoRNAs. Nevertheless, all the commercial assays that have been developed for the detection of these autoantibodies have used fibrillarin as the target antigen. These assays have shown controversial results when compared to the gold-standard IP [95, 207, 208] or have not been compared to it [209].

Anti-fibrillarin autoantibodies are considered highly specific for SSc [125, 210, 211]. Moreover, this autoantibody has been associated with higher frequencies of dyspigmentation, calcinosis, digital pitting scars and ulcers, muscle, GI and cardiac involvement and PAH [210–212]. Interestingly, anti-fibrillarin autoantibodies have been described to be significantly more frequent in African American (17.0-50%) than in Caucasian (1.8-5%), Japanese (2.0-7.0%), or Hispanic (7.0%) SSc patients [197, 210, 212–217]. However, whether anti-fibrillarin autoantibodies are associated with the dcSSc subset in African American patients or Caucasian patients still remains controversial [210, 215, 217].

1.7.5.4. Anti-PM/Scl autoantibodies

The PM/Scl complex was first identified as an autoantigen using double immunodiffusion. First, it was detected that autoantibodies in sera of patients suffering from PM and overlap syndromes precipitated an antigen from calf thymus extract that was denominated “PM-1” [218]. Subsequent studies demonstrated that especially sera of patients with SSc-PM overlap syndrome showed reactivity against this autoantigen, which was therefore termed “PM/Scl” [219, 220]. Sera from patients with anti-PM/Scl reactivity were found to contain

autoantibodies that uniformly immunoprecipitated at least 11 proteins of an approximate molecular weight of 110, 90, 80, 39, 37, 33, 30, 27, 26, 22, and 20kDa, although additional polypeptides could also be seen [221, 222].

The PM/Scl antigen complex is actually the human RNA exosome [223–226]. The human RNA exosome is a ribonuclease complex with endoribonuclease and exoribonuclease activity involved in the degradation and processing of various RNA molecules [227]. After decapping by deadenylation of the polyA tail of mRNA molecules, RNA exosome acts as a 3' to 5' exoribonuclease degrading the mRNA [228–231]. In addition to this role in RNA decay, the RNA exosome also participates in the processing of snoRNAs and snRNAs and in the maturation of 5.8S rRNAs [232–235]. The RNA exosome consists of 10 conserved core protein subunits arranged into a ring-like structure. The macromolecular complex is comprised by a central 6-subunit ring formed by Rrp41, Rrp42, Rrp43, PM/Scl-75 (Rrp45), Rrp46 and myotubularin-related protein 3 (Mtr3); a 3-subunit cap formed by Rrp40, Csl4 and Rrp4; and an exoribonuclease subunit PM/Scl-100 (Rrp6) associated with the central ring [236]. Structural analysis has revealed that the RNA to be processed enters the RNA exosome through the central pore opening of its ring-like shape [237].

Because of the great diversity of processes the exosome is involved in, which occur in a variety of subcellular compartments, this complex is expected to be localised in the cytoplasm, nucleoplasm as well as nucleolus. In fact, it has been demonstrated that RNA exosome is present in the nucleus as well as in the cytoplasm, bound to compartment-specific cofactors [238]. Paradoxically, the results of immunofluorescence experiments have demonstrated that most exosome components are highly concentrated in the GC of nucleoli (see section 1.7.5.) [222, 239]. However, this type of distribution pattern is consistent with the estimated number of substrate RNA molecules for the exosome present in the different subcellular compartments in combination with the relative volumes of these compartments. For example, the number of rRNAs transcribed and processed in the nucleolus of growing cell lines in a certain time interval exceeds that of the other class of substrate RNAs in the nucleoplasm (snRNA and snoRNA) and cytoplasm (mRNA), whereas the volume of the nucleoli is much smaller than that of these other compartments. Therefore, if the distribution of the exosome would reflect the relative concentration of its substrates, the highest concentration of human RNA exosome is expected in the nucleoli [240]. Indeed, patients with anti-PM/Scl autoantibodies always show homogeneous nucleolar staining (AC-8) [161].

Immunoblot studies using total cell extracts or recombinant proteins and protein IP have demonstrated that most anti-PM/Scl autoantibody positive sera react with a protein

migrating at 100-110kDa, while some of them (30-50%) also show reactivity to a 70-75kDa migrating protein [221, 241–243]. Due to their apparent size on SDS-PAGE, these two proteins were named PM/Scl-100 and PM/Scl-75, respectively. Nevertheless, later studies reported that PM/Scl-75 is actually a 39.2kDa protein with an aberrant migration in SDS-PAGE because of its carboxyl-terminal half that is enriched in acidic residues [241], while PM/Scl-100 is in fact a 100.8kDa protein [242, 244, 245]. Although some studies using ELISAs have reported that some rare sera display a clear preference for PM/Scl-75 antigen with a weak PM/Scl-100 reactivity, co-occurrence of anti-PM/Scl-100 and anti-PM/Scl-75 seems to be particularly associated with the SSc-PM overlap syndrome [246].

As previously mentioned, a high association of anti-PM/Scl autoantibodies with SSc-PM overlap syndrome has been found, but these autoantibodies can also be detected in patients with SSc, PM and other autoimmune diseases [247]. However, while this autoantibody is only detected in 2-10% of SSc patients [125, 246–248] and in 5-10% of patients with PM [125, 220, 247, 249–251], 40-90% of the patients positive for anti-PM/Scl autoantibodies are indeed diagnosed as having a SSc/PM overlap syndrome. Patients diagnosed with SSc-PM overlap syndrome present clinical characteristics of both diseases, mainly RP, cutaneous changes, muscle involvement, arthritis and ILD, but with a rather benign course [247, 252, 253]. On the other hand, SSc patients with anti-PM/Scl autoantibodies display a consistently higher prevalence of muscle involvement as compared to anti-PM/Scl autoantibody negative SSc patients. On the other hand, as already discussed in section 1.5.7., the prevalence of muscle involvement in SSc may vary widely due to the lack of diagnostic consensus criteria for muscle involvement, and also the inclusion or exclusion of SSc-myositis overlap syndromes [125, 248]. Overall, most patients positive for anti-PM/Scl autoantibody present the limited cutaneous form of SSc (80-100%), with a higher frequency of calcinosis and usual visceral SSc involvements but with a good prognosis [125, 247]. In fact, myositis tends to be relatively responsive to treatment in all patients positive for anti-PM/Scl autoantibodies, either diagnosed with SSc-PM overlap or SSc [254]. Finally, contrary to what is seen for anti-fibrillarin autoantibodies, anti-PM/Scl autoantibodies are mainly detected in Caucasian patients and are rare in black and Asian patients [253, 255].

1.7.6. Anti-Ku autoantibodies

Anti-Ku autoantibodies were originally detected in a cohort of Japanese patients with different autoimmune diseases by immunodiffusion experiments using saline extracts of calf thymus as antigen source. In this first study, they were found to be highly associated with

SSc-PM overlap syndrome and good response to corticosteroid therapy [256]. Later on, anti-Ku autoantibodies were demonstrated to recognise two different proteins of 70 and 80KDa by IP [257] and were also found to be associated with other systemic autoimmune diseases, especially with SLE or SLE overlap syndromes [86, 87, 258–261]. In fact, it was shown that anti-Ku autoantibodies were found to be frequently associated with anti-Sm autoantibodies, a specificity almost exclusively seen in SLE [86, 87]. Co-presence of other autoantibodies, such as anti-Ro, anti-La, anti-U1 snRNP, anti-centromere and anti-Scl70, has also been detected in sera of SSc or SSc-PM patients positive for anti-Ku autoantibodies [85–87, 256–259, 261].

Anti-Ku reactivity has been investigated with different methodological approaches in different studies. While originally, immunodiffusion assays were used to detect anti-Ku autoantibodies, ELISA, immunoblot, and protein IP have also been used. The prevalence of anti-Ku autoantibodies in various autoimmune diseases varies widely, ranging from 3% with IB analysis to 55% using a capture ELISA [86, 87, 256, 260]. Moreover, differences in the frequency of this autoantibody have also been reported in African American SLE patients compared to Caucasians. In addition, while no anti-PM/Scl autoantibodies are detected in Japanese patients with inflammatory muscle disease, anti-Ku autoantibodies are found specifically in Japanese patients with SSc-PM overlap syndrome. Conversely, anti-Ku autoantibody is not so common in SSc-PM overlap syndrome in North Americans and it seems to be found more associated with SLE. It appears that a genetic or environmental influence may regulate the occurrence of anti-Ku and anti-PM/Scl autoantibodies [86, 262, 263]. Therefore, case selection and clinical and racial differences in the cohort of patients studied and methods employed to detect anti-Ku autoantibodies may account for some of the disagreements reported in the literature.

Ku70 and Ku80 are two of the central components of the non-homologous end-joining DNA repair pathway (NHEJ). NHEJ is based on unguided re-ligation of two double stranded DNA (dsDNA) ends created as a result of DNA breakage. Ku70 and Ku80 proteins form a heterodimer with a ring-shaped structure on which the central canal fits a dsDNA helix due in part to a high affinity for dsDNA ends. Ku70/80 recruits and activates the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) to the damage site, forming the DNA-dependent protein kinase complex (DNA-PK), which protects the dsDNA ends from degradation and juxtaposes the DNA ends in a synaptic complex. The DNA-PK complex phosphorylates a large number of substrate proteins, including itself, resulting in a conformational change that makes DNA ends more accessible for further processing and

repair. Depending on the DNA breakage, ligation may either take place directly or after limited processing of the DNA ends, the later resulting in deletion or insertion of a few nucleotides at the break site. Therefore, different set of proteins may be required for the repair process depending on the DNA breakage. In line with this, eroded DNA ends require processing by helicases and exonucleases, while minor DNA processing is catalysed directly by specific DNA polymerases. The final ligation step is promoted by a series of different proteins and carried out by Cernunnos and Ligase IV/XRCC4. More specifically, Cernunnos interacts with XRCC4, which stimulates Ligase IV to ligate the nick [264, 265].

Due to the main subcellular localisation of Ku70/80 in the nucleus, patients with anti-Ku autoantibodies normally present a nucleolar fine speckled pattern with no staining of the nucleoli or with nucleoli staining (AC-4) [85–87, 257, 261].

Autoantibodies against other molecules of the NHEJ pathway have also been described [81, 266]. Autoantibody binding could render Ku and other complex proteins more resistant to dissociation [81, 84]. This can alter the pattern of antigen processing and might enhance the presentation of abnormally processed peptides by antigen presenting cells. Stabilising autoantibodies might retard the dissociation of protein complexes inside endosomes, increasing their effective concentration, or might act as hindering structures that limit accessibility to proteases. Normally in patients with autoantibodies against other molecules of this pathway, autoantibodies against Ku are also found. These data support the idea that autoimmunity begins with an anti-Ku response that occasionally spreads to other components of the same pathway, but in some cases, a response to other molecules occurs first [75, 81, 266, 267].

1.7.7. Autoantibodies against components of the spliceosomes

In eukaryotes, mature RNA is formed by the removal of introns from a primary transcript. This process is denominated splicing and is always catalysed by a multicomponent complex denominated the spliceosome [268]. More specifically, two intron classes have been identified, a common U2-type and a rare U12-type (<1% introns). Splicing of U2-type introns is catalysed by the U2-dependent (major) spliceosome while splicing of U12-type introns is catalysed by the U12-dependent (minor) spliceosome [269–271].

U12-type introns were first identified on the basis of terminal AT-AC dinucleotides, which distinguished them from the typical U2-type introns with GT-AG terminal dinucleotides. However, it was soon realised that these termini are not exclusively present in U12-type introns. Moreover, GT-AG are also found to function as the terminal dinucleotides and are

actually the more common subtype also in U12-type introns. In fact, the real defining features of U12-type introns are the highly conserved 5' splice sites (5'ss) and branch point sequence (BPS), and the lack of a distinct polypyrimidine tract (PPT) that is typically found upstream of U2-type 3' splice sites (3'ss). Additionally, the distance from a U12-type BPS to the 3'ss has been shown to be an important factor for the recognition of U12-type introns, and is significantly shorter than in U2-type introns [272].

Recent studies have shown that the genomic regions of highly transcribed protein coding genes and their corresponding nascent pre-mRNAs are organized around nuclear speckles. A nuclear speckle is defined as a nuclear body that contains numerous splicing and processing proteins, including the spliceosome. Specifically, the inner structural core of nuclear speckles is composed of the splicing factors and other mRNA processing factors, whereas the periphery consists of chromatin and nascent pre-mRNA. Interestingly, the distance between a gene and a nuclear speckle is inversely correlated with its transcription level, and genes in highly transcribed neighbourhoods are preferentially proximal to speckles. Although splicing does not appear to occur within nuclear speckles, the splicing factors that are contained within them were shown to diffuse from the speckle to the nascent pre-mRNA (Figure 13). It is hypothesised that the compartmentalization of regulatory components around speckles increases the spatial concentration of splicing factors near nascent pre-mRNAs and, in this way, increases the rate of co-transcriptional splicing [273].

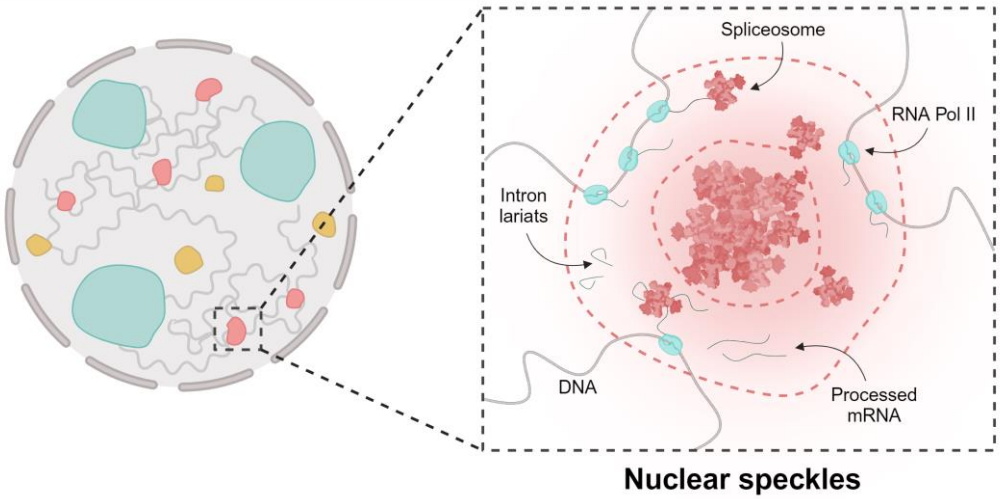


Figure 13. Schematic illustration of nuclear speckles. The inner structural core of nuclear speckles is composed of spliceosome complexes, splicing factors and other mRNA processing factors. Gene transcription occurs at the periphery of nuclear speckles, where the spliceosomes diffuse to the nascent pre-mRNA and excise introns in the form of lariats. Created with *www.BioRender.com.*

The major spliceosome is a nuclear assemblage that, when purified in complex with a pre-mRNA, is composed of at least 145 associated factors [274]. While the major spliceosome's size clearly illustrates its complexity, it is also highly dynamic. Large subassemblies are observed to dissociate and associate in an ordered manner *in vitro*. In this line, critical structural rearrangements are hypothesised to take place, not only between protein components but also between RNA subunits, to form the enzyme's active sites [275]. The major spliceosome function is based on five RNA-protein complexes termed uridine-rich small nucleolar ribonucleoprotein (U snRNP) particles, the U1, U2, U4, U5 and U6 snRNPs. The five U snRNPs are compositionally similar but functionally distinct and each is composed of a single uridine-rich small nuclear RNA (U snRNA), a set of seven Sm or Sm-like (LSm) proteins, and three or more U snRNP-specific proteins [275].

The seven Sm proteins (Sm-B/B', -D_1, -D_2, -D_3, -E, -F, and -G) are critical to the assembly, transport, and integrity of the U snRNPs [276]. In the absence of an U snRNA, Sm proteins associate as stable heterodimers ($SmD_1 \cdot D_2$ and $SmD_3 \cdot B$) and a heterotrimer (SmF·E·G) in the cytoplasm [277], but in the presence of an U snRNA, a single copy of each of seven Sm proteins interacts pairwise to assemble into a heptameric ring [278]. U1, U2, U4 and U5 snRNAs are RNApol II transcripts, and each is shuttled to the cytoplasm where their cotranscriptionally acquired 7-methylguaniosine (m_1G) cap is hypermethylated to a 2,2,7-trimethylguanosine (m_3G) cap (5' TMG cap). Once in the cytoplasm, Sm proteins recognise a short single stranded RNA sequence in U1, U2, U4, and U5 snRNAs called the Sm site. Subsequently, the Sm ring is formed around the U snRNA molecule, assuming a cartwheel-like arrangement where RNA bases radiate outward in varying orientations to interact with Sm protein residues. The 5' TMG cap and Sm core formation act as a bipartite recognition signal for the transport of the pre-assembled U snRNP to the nucleus where additional U snRNP-specific proteins bind [276]. In contrast, U6 snRNA is an RNApol III transcript that remains in the nucleus and does not contain a Sm binding site. A single-stranded sequence at the 3' end of U6 snRNA is recognised by a set of seven proteins homologous to the Sm proteins, the LSm proteins [279].

The U12-dependent (minor) spliceosome, responsible for the excision of the U12-type introns, is structurally similar to the U2-type spliceosome. It contains protein subunits and the U5 snRNA as well as the U11, U12, U4atac, and U6atac spliceosomal snRNAs that are functionally and structurally related to the U1, U2, U4 and U6 snRNAs of the major spliceosome, respectively [269, 270, 280, 281]. Much of the specificity in the splicing reaction of both spliceosomes is accomplished by pairing with snRNAs. Although the overall

assembly and catalytic steps of intron removal are very similar between the two spliceosomes [269, 270, 282], there is a significant difference in the intron recognition step, which for minor introns is carried out by a preformed U11/U12 di-snRNP complex, while the major introns are recognised independently by individual U1 and U2 snRNPs (**Figure 14**) [272, 283–285].

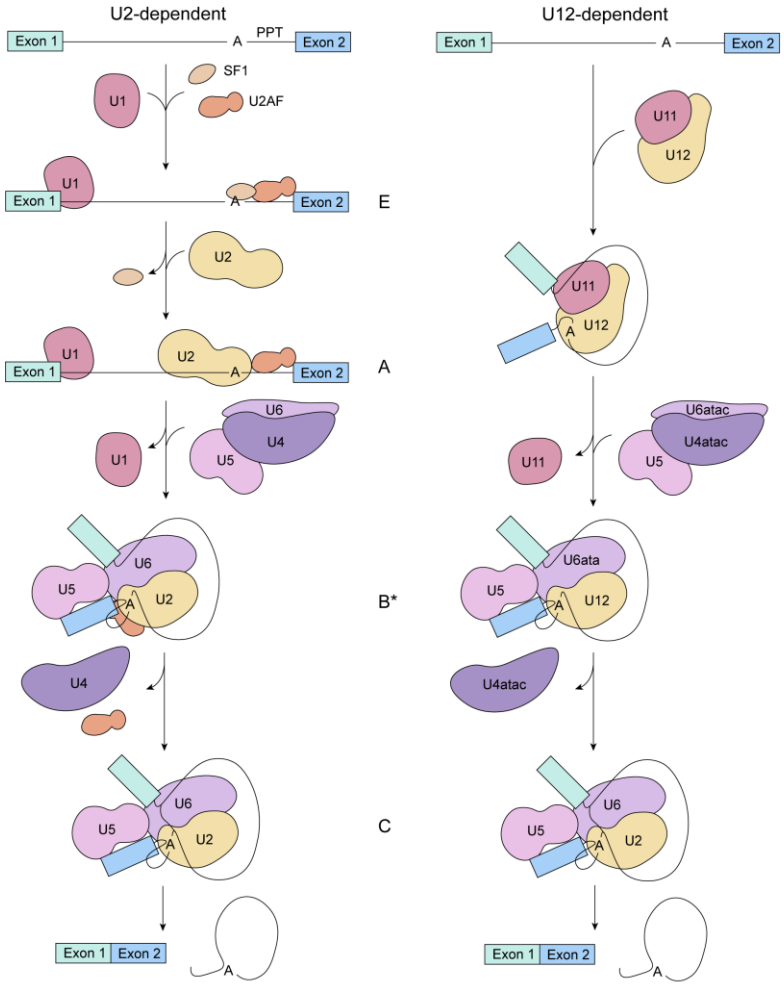


Figure 14. Graphical summary of U2-type and U12-type intron splicing. In U2-type introns, 5'ss is initially recognized by U1 snRNP, while the BPS, PPT and 3'ss are recognized by the protein factors SF1, U2AF^{65} and U2AF^{35}, respectively, together forming the spliceosomal commitment (or E) complex. During the formation of the pre-spliceosome, or A complex, U2 snRNP replaces SF1 at the BPS. At later stages, the U4/U6.U5 tri-snRNP stably associates with the spliceosome (B complex). U6 snRNP then replaces U1 snRNP at the 5'ss and the U4/U6 di-snRNP structure unwinds. U2 and U6 snRNPs base pair with each other to form the catalytic core structure within B* complex in which the reactive A residue at the BP and the 5'ss are juxtaposed for the first transesterification reaction. During this process, both the U1 and U4 snRNPs and U2AF are released from the spliceosome, giving rise to the C complex that catalyses the second transesterification needed for the excision of the intron as a lariat. In U12-type introns, the preformed U11/U12 di-snRNP binds the intron as a unit, and the 5'ss and BPS are recognized in a cooperative manner within the pre-spliceosomal or A complex. However, U11 snRNP/5'ss base pairing still precedes the formation of stable U12 snRNP/BPS base pairing. After initial recognition for the splice sites, overall assembly and catalytic steps of intron removal are very similar.

This functional difference is reflected in the composition of the U11/U12 di-snRNP, which, in addition to the two unique U snRNAs, also contains 7 protein species that are unique to the minor spliceosome: RNA-binding region-containing protein 3 (RNPC3/65K), programmed cell death protein 7 (PDCD7/59K), 48K, 35K, zinc finger CCHC-type and RNA-binding motif-containing protein 1 (ZCRB1/31K), 25K and zinc finger matrin-type protein 5 (ZMAT5/20K) [286–288]. Moreover, the notion that specific protein components of the minor spliceosome are needed only during the intron recognition phase and not in the later assembly steps has been challenged recently. A recent report of the cryo-EM structure of the minor catalytic spliceosome complex revealed that RNA-binding protein 48 (RBM48), armadillo repeat-containing protein 7 (ARMC7), sodium channel modifier 1 (SCNM1) and cysteine-rich PDZ-binding protein (CRIPT) are also specific components of the minor spliceosome [289]. Moreover, it has been described that centrosomal AT-AC splicing factor (CENATAC) and thioredoxin-like protein 4B (TXNL4B), are specific for the U4atac/U6atac di-snRNP and U4atac/U6atac.U5 tri-snRNP [290]. However, from the more than 300 proteins that have been identified in spliceosome complexes, including U snRNP-associated proteins (Sm proteins, LSm proteins, and U snRNP-specific peptides) and non-snRNP associated proteins (splicing factors), many are shared between the minor and major spliceosomes [291].

Before accumulating around nuclear speckles, all U snRNP transiently transit through CBs [292, 293]. CBs are nuclear compartments enriched in specific proteins and RNA components not limited by any lipid bilayer. CBs owe their formation and maintenance to a protein named coilin with no known function other than scaffolding CBs. This protein binds RNA molecules and contains a N-terminal domain that mediates coilin-coilin “self” interactions, giving rise to CB’s characteristic structure (coiled bodies). However, other proteins are also present in CBs. In fact, the protein deficient in spinal muscular atrophy (SMA), spinal motor neuron protein (SMN), cooperates with coilin to maintain CB integrity [294]. On the other hand, CBs are enriched in U snRNAs and snoRNAs, as different steps of biogenesis and assembly of snRNPs and snoRNPs take place in these nuclear compartments. Indeed, CB homeostasis requires ongoing U snRNP biogenesis, as a perturbation of SMN and other factors involved in U snRNP biogenesis cause CB to disassemble. In this line, CB consistently associates with specific loci on multiple chromosomes, many of which include U snRNA gene arrays, snoRNAs, as well as histone gene clusters. On the other hand, without CB formation, these chromosomal regions are no longer clustered, and the expression of many of the associated U snRNA loci is significantly reduced (Figure 15) [295].

Regarding U snRNPs biogenesis, CBs participate in the transcription, pre-processing and formation of the export complex prior to the exit of immature U snRNPs to the cytoplasm. After its assembly, the core U snRNP complexes return to the CB, where snRNPs-specific proteins are added, and tri-snRNP assembly occurs [296]. The SMN complex, which includes SMN and several tightly associated proteins, collectively called Gemins, regulates the entire cytoplasmic phase of the snRNP cycle, including Sm core assembly and 5' TMG cap formation. Following the import and localisation of newly assembled U snRNPs to the CB, coilin may function to disrupt the SMN-snRNP complex and facilitate higher-order U snRNP formation. Once released from the SMN complex, the newly assembled U snRNP is free to diffuse throughout the interchromatin space (**Figure 15**). The fate of the SMN complex following snRNP release in the CB has not been demonstrated, but it has been proposed that it can accumulate in distinct nuclear substructures called Gemini bodies, or Gems, constituted solely by components of the SMN complex [295].

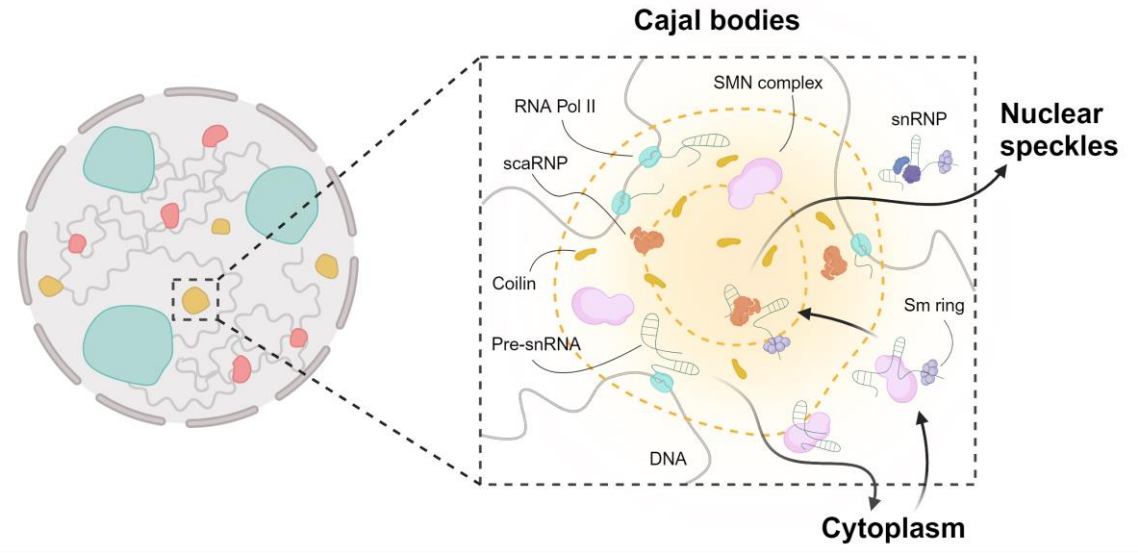


Figure 15. Graphical summary of U snRNP biogenesis. Pre-U snRNAs are transcribed by RNApol II at CBs, where they associate with the SMN complex. SMN regulates the cytoplasmic phase of the U snRNP biogenesis including Sm ring and 5' TMG cap formation, which act as a recognition signal for transport of the pre-assembled U snRNP to CBs again. In the CB, the SMN complex is released, U snRNA is modified by scaRNPs and specific proteins are added to each complex after diffusing to nuclear speckles as mature U snRNP complexes. Created with *www.BioRender.com*.

On the other hand, U snRNAs are modified in the CB by specific snoRNAs that guide 2'-O-methylation and pseudouridylation. These specific snoRNAs localise to the CB and not nucleoli and are named scaRNAs (**Figure 15**). scaRNAs can either contain a pair of box H/ACA, a pair of box C/D motifs, or the 2 motifs at the same time. scaRNAs presenting C/D box motifs bind to the same proteins as C/D box snoRNPs, while scaRNAs with H/ACA box

motifs bind to the same proteins as H/ACA box snoRNPs . However, scaRNAs are characterised by containing additional elements that retain them in CBs. This localisation signal is named CAB box, and it is required for the binding of scaRNAs to the telomerase cajal body protein 1 (TCAB1), which is also essential for CB maintenance. Indeed, TCAB1 protein is associated with SMN and coilin, and is believed to recruit the SMN complex to CBs [154, 294, 297, 298]. Interestingly, human telomerase RNA resembles scaRNAs and also localises to CBs. However, characteristically this RNA leaves CBs to accumulate at certain telomeres during the S-phase of the interphase. In addition to scaRNAs, canonical nucleolar snoRNAs that are synthesised in the nucleoplasm traffic through CBs *en route* to nucleoli, likely to undergo final snoRNP assembly steps. In fact, fibrillarin and GAR1 have been suggested to become stably associated with snoRNPs in CBs (see section 1.7.5.). As it has been demonstrated that these proteins interact with SMN, it has been proposed that the assembly process of snoRNPs may be facilitated by the SMN complex[154]. Similar to what has been reported in nuclear speckles, the accumulation of snoRNAs and U snRNAs in CBs has been proposed to enhance ribonucleoprotein biogenesis by favouring interactions between partners. Coilin may have a principal role in promoting this interplay as it directly interacts with many CB proteins as well as numerous non-coding RNAs, including snoRNAs and scaRNAs [154].

1.7.7.1. Autoantibodies against the major spliceosome

Autoantibodies against U1 snRNP and Sm are among the most common autoantibody specificities in systemic autoimmune rheumatic diseases. Both autoantibodies were originally identified by double immunodiffusion using calf thymus extracts as source of autoantigens. Although anti-U1 snRNP was recognised as a distinctive precipitin line from anti-Sm, both precipitin lines partially fused, suggesting some shared antigenic component. Subsequent studies using RNA and protein IP assays, helped in the identification of the target of these autoantibodies and helped in the elucidation and function of key components of the spliceosome. Since then, it has been demonstrated that virtually all anti-Sm positive patients also target U1 snRNP, as U1 snRNP is constituted in part by Sm proteins, while anti-U1 snRNP positive patients may not be positive for anti-Sm, in the case of patients recognising proteins specific of the U1 snRNP particle [299].

Although anti-U1 snRNP and anti-Sm often coexist, there are major differences in the clinical significance of these two specificities. Anti-Sm autoantibodies are highly specific for the diagnosis of SLE, and although they present a not very high sensitivity (5-30%), their

presence is one of the serologic criteria in the ACR classification criteria for SLE. On the other hand, anti-U1 snRNP autoantibodies are found in patients with various diseases, including SSc, SLE, SjS, RA, inflammatory myopathies and UCTD. In fact, by definition, all patients with mixed connective tissue disease (MCTD) are positive for anti-U1 snRNP autoantibodies. MCTD is characterized by overlapping features of SLE, SSc and inflammatory myopathy, as patients present a high prevalence of RP, oedema of the fingers, arthritis or arthralgia, myositis, serositis, favourable response to steroid treatment, and a relative absence of renal disease, with a good overall prognosis [299].

Autoantibodies to other U snRNPs are far more uncommon but have also been reported in rare cases. Anti-U4/U6 di-snRNP autoantibodies were initially reported in the serum of a patient with SSc and subsequently in a Japanese patient with primary SjS [300, 301]. Anti-U5 snRNP autoantibodies were identified in the serum of a Caucasian patient with SSc-PM overlap syndrome and later in a Japanese patient with a similar overlap syndrome and large cell carcinoma of the lung [302, 303]. Isolated anti-U2 snRNP autoantibodies were also reported in a patient with SSc-PM overlap syndrome [304], while anti-U1/U2 di-snRNP autoantibodies were described in a cohort of patients with different autoimmune diseases (SLE, MCTD, RA and SjS) [305]. Additionally, autoantibodies to the 5' TMG cap recognizing all U series RNAs (U1, U2, U3, U4, U5, U8, U11, U12, U13) except U6 RNA have also been reported in patients with SSc [205, 306].

Anti-SMN autoantibodies were first described in two patients with SSc-PM overlap syndrome and one patient with PM that presented an uncommon pattern of protein IP. These patients presented reactivity against proteins Sm-D_1, D_2, E, F and G, but did not immunoprecipitate the rest of the components of the major spliceosome neither none of the related U snRNAs by RNA IP. Additionally, they immunoprecipitated 4 proteins that were identified as SMN, Gemin2, Gemin3 and Gemin4 [307]. However, further studies have demonstrated that anti-SMN autoantibodies are not specific for SSc-PM overlap syndrome, as they are more common in MCTD (36% of patients), co-existing with anti-U1 snRNP autoantibodies. Furthermore, anti-SMN autoantibodies seem to be specifically associated with the subset of MCTD patients presenting with clinical characteristics of SLE, SSc and inflammatory myopathy, and a higher prevalence of PAH and ILD and overall poorer prognosis [308].

Anti-U1 snRNP and anti-Sm are typically correlated with a coarse speckled nuclear pattern staining when screened by IIF (AC-5) due to the subcellular localization of the spliceosome in the speckles of the nucleus [299]. Similarly, all sera containing anti-U4/U6 di-snRNP, -U5

snRNP, -U2 snRNP or -5' TMG cap autoantibodies, show the same coarse speckled nuclear pattern (AC-5) by IIF [300–304]. On the other hand, anti-SMN positive patients present a few nuclear dots pattern by IIF (AC-7), as SMN is mainly localised in CB. However, as the majority of patients with anti-SMN autoantibodies are also positive for anti-U1 snRNP, they normally present an overlap of both IIF patterns, AC-5 and AC-7, in which the few nuclear dots pattern is difficult to distinguish [307, 308].

1.7.7.2. Anti-U11/U12 snRNP autoantibodies

Anti-U11 snRNP autoantibodies were first reported in 1993 in a patient with SSc. In this first study, they only detected U11 snRNA by RNA IP, but this autoantibody was expected to also recognise U12 snRNP since U11/U12 di-snRNP form part of the minor spliceosome as a stable complex [306, 309]. It was suggested that the low abundance of U12 snRNA molecule in the cell extract could be responsible for the lack of its detection. Until 2009, there were no more reports about this autoantibody, when it was detected by RNA IP in two consecutive series of SSc patients. In this study, the term anti-U11/U12 snRNP was coined, although the band corresponding to U12 snRNA was not clearly seen by RNA IP in positive patients [97]. The majority of sera positive for these autoantibodies recognised a 65-68kDa protein that was shown to be RNPC3 subunit of the minor spliceosome [97, 306, 310]. Since then, anti-U11/U12 snRNP autoantibodies have been tested by three different assays: RNA IP [311], IP of 35^S-methionine labelled protein generated by in vitro transcription and translation (IVTT) from complementary DNA (cDNA) encoding full-length RNPC3 [312, 313] and a particle-based multi-analyte technology (PMAT) that uses particles coated with recombinant RNPC3 [314]. IP of 35^S-methionine labelled RNPC3 produced by IVTT showed a 95% sensibility when compared with RNA IP [313], whereas there is not any comparative study of the PMAT technology for detecting these autoantibodies published in the literature.

Anti-U11/U12 snRNP autoantibodies are considered SSc-specific, as they are not detected in sera of patients with idiopathic pulmonary fibrosis and other systemic autoimmune disease [97]. These autoantibodies are found both in dcSSc and lcSSc subsets at similar rate, but are highly associated with ILD, often severe, even when compared with anti-Scl70 autoantibody positive patients [97, 313, 314]. Moreover, they have also been reported to be associated with a more severe oesophageal dysmotility [97, 313] and increased risk of cancer at the time of the first clinical manifestations of SSc [312]. Although there are very few studies about anti-U11/U12 snRNP autoantibodies, their prevalence has been reported to range between 3,2% to 8% in SSc patients [97, 314, 315].

Anti-U11/U12 snRNP autoantibodies do not present a clear IIF pattern, probably due to the low abundancy of this ribonucleoprotein complex in HEp-2 cells. In line with this, although most positive sera resemble a weak nuclear fine speckled pattern (AC-4), other ANA patterns have also been reported [312, 314]. However, the reported heterogeneity could be associated with the co-presence of other SSc-associated autoantibodies in anti-U11/U12 snRNP positive patient [97, 313, 314].

1.7.8. Anti-EIF2B autoantibodies

Anti-EIF2B autoantibodies were first discovered by protein IP, in which a few SSc patients showed immunoprecipitation of a 30kDa protein that was identified to be the Eukaryotic Initiation Factor 2B (EIF2B). After this first report in 2016, there have not been many more positive cases published in the literature, as it is a rare autoantibody with an estimated prevalence of around 1%. However, these autoantibodies have consistently been associated with SSc overlap syndrome (inflammatory myopathy and RA) and ILD, and they are slightly more common in the diffuse form of SSc. By contrast, these autoantibodies have not been detected in other diseases and are considered SSc-specific [316–319].

The translation of mRNAs into proteins is a highly energy-intensive and resource-intensive cellular process. Therefore, it requires a large number of regulating factors to ensure its reliability and accuracy. The regulation most commonly occurs during the initiation phase of translation. In eukaryotes, start codons are selected by mRNA scanning mediated by the 40S ribosomal subunit in collaboration with various protein factors, the eukaryotic translation initiation factors (eIFs). eIF2, which is bound to GTP, guides the initiator tRNA to the ribosome, participates in the scanning of the mRNA molecule, and supports the selection of the start codon in the physiological state. Once the start codon is recognised, the hydrolysis of GTP by eIF2 commits the 40S ribosome to translation initiation and leads to the recruitment of the 60S subunit and entry into the elongation phase of translation initiation. After the release of GDP from the ribosome, GDP is exchanged for GTP by the guanine nucleotide exchange factor EIF2B, preparing the complex for another round of initiation [320, 321]. Therefore, as EIF2B is localised in the cytoplasm, this autoantibody is associated with a cytoplasmatic dense fine speckled staining IIF pattern (AC-19) [316–319].

1.7.9. Anti-RuvBL1/2 autoantibodies

Autoantibodies against RuvBL1 and RuvBL2 proteins were first reported in 2014. These autoantibodies were detected by protein IP as a doublet of bands at a molecular weight of

around 50kDa that were later identified by MS as RuvB-like protein 1 (RuvBL1) and 2 (RuvBL2). The prevalence of this autoantibody was estimated to be 1.5-5.9% in this initial report, in which two Japanese cohorts of SSc patients and a cohort from Pittsburgh were studied [322]. However, in a subsequent study in which a very large number of sera samples from SSc patients of the United States of America and Canada were studied by protein IP, only two samples were positive, with a prevalence of 0.06% [318]. Moreover, only a few additional reports of patients positive for anti-RuvBL1/2 autoantibodies have been published, indicating that it is indeed a rare autoantibody. Regarding the clinical associations of this autoantibody, it has consistently been associated with SSc-myositis overlap syndrome and, interestingly, with the diffuse cutaneous form of the disease, opposite to which is observed in anti-PM/Scl and anti-Ku autoantibody positive patients [318, 322–326]. Considering the association of this autoantibody with SSc-myositis overlap syndrome, heterogeneity on the used clinical exclusion or inclusion criteria could be accountable for the observed prevalence variability [47–49].

RuvBL1 and RuvBL2 are ATPases of the AAA (ATPase Associated with diverse cellular Activities) family and have been implicated in many cellular pathways. RuvBL1 and RuvBL2 hexamerize, forming a double-hexameric ring, and it has been suggested that this complex could act as a scaffolding protein, explaining its appearance in diverse cellular protein complexes as various chromatin remodelling complexes and complexes related to assembly and maturation of snoRNPs [327]. Although the diverse functions of the RuvBL1/2 complex and protein expression assays demonstrate that this complex is found in both the cytoplasm and nucleus of cells, patients with anti-RuvBL1/2 autoantibodies demonstrate a very strong fine speckled nuclear pattern by IIF (AC-4) [322].

1.7.10. Anti-telomerase autoantibodies

In a study from 2021, autoantibodies targeting telomerase and sheltering proteins were specifically reported in SSc patients, as they were rarely present in other autoimmune diseases such as RA, inflammatory myopathies or healthy donors. Specifically, autoantibodies directed against the telomeric repeat-binding factor 1 (TERF1), a sheltering protein, were present in 9% of SSc patients of the study. These autoantibodies were found to be associated with severe lung disease. However, non-white race was strongly associated with severe lung disease and was also associated with the presence of TERF1 autoantibodies, and the association between TERF1 autoantibodies and severe lung disease was not statistically significant after adjusting for race. On the other hand, there was also an

association of anti-TERF1 autoantibodies with a history of severe muscle disease and inflammatory arthritis, but the co-presence of anti-U1 snRNP and anti-Ku autoantibodies that was detected could have represented a co-founding factor. Additionally, less common autoantibodies against other sheltering proteins of the telomere and showing reactivity against the catalytic RNA molecule of the telomerase holoenzyme have also been reported [328, 329]. Due to the few reports about autoantibodies against telomerase and the sheltering proteins, there is still no clear ANA IIF pattern to which they are associated.

Telomeres are repetitive DNA structures located at the ends of chromosomes, offering stability and protection of the chromosomes and, hence, of the genome as a whole. Because of the end of replication, with each DNA replication of cells, some DNA at the ends of the chromosomes is lost. Without telomeres, this loss per replication would be of chromosomal genetic information, but due to the presence of telomeres, only not critical genetic information is lost. With time, telomere repeats shorten, and at a certain point, cells with short telomeres are activated to enter cellular senescence or apoptosis and, therefore, irreversibly lose function. One of nature's ways to overcome this ageing process is via the telomerase holoenzyme. This ribonucleoprotein complex is involved in the natural replenishment of telomere repeats in selected cell types, including stem cells, sperms and lymphocytes, increasing the cell's proliferative lifespan. While the telomerase holoenzyme solves the end replication problem, linear chromosomes must also solve the end protection problem. The end protection problem occurs when the natural ends of linear chromosomes are misrecognised by the DNA damage response and repair machinery as double strand breaks requiring repair. The sheltering complex, constituted by protection of telomeres protein 1 (POT1), tripeptidyl-peptidase 1 (TPP1), TERF1-interacting nuclear factor 2 (TIN2), TERF1, telomeric repeat-binding factor 2 (TERF2), and Rap1, solves the end protection problem by coating telomeric DNA [330, 331].

The protein subunit of telomerase, called Telomerase Reverse Transcriptase (TERT), contains a catalytic reverse transcriptase domain for DNA synthesis, while the RNA subunit of telomerase, called telomerase RNA component (TERC), contains, amongst other elements, the template for telomeric repeat retrotranscription. In addition to these two catalytic core subunits, many additional proteins are necessary for assembly, subcellular trafficking, and telomere association of a functional telomerase holoenzyme. TERC stability requires precursor co-transcriptional assembly as a H/ACA snoRNP with dyskerin, NOP10, NHP2, and the chaperone nuclear assembly factor 1 (NAF1), which is later replaced by GAR1. After initial TERC biogenesis, a fraction of the biologically stable TERC snoRNP

associates with TERT through multiple direct protein-RNA interactions. TERC snoRNP is initially assembled in the DFC and is retained in the nucleoli through the interaction between TERT and nucleolin. However, in the S phase of the cell cycle, TERC snoRNP separates from nucleoli and is transported to CB for subsequent recruitment to telomeric chromatin. Within the CB, TERC snoRNP binds to TCAB1, increasing the steady-state CB association of TERC and a subset of other H/ACA snoRNAs that also contain CAB box. Although they do not contribute to telomerase catalytic activation, these interactions are necessary for TERT-TERC recruitment to telomeres and their extension. In fact, mutations in TCAB1 prevent telomerase from elongating telomeres by disrupting telomerase localisation to CB [330, 332].

1.7.11. Anti-Ro60/SSA and anti-La/SSB autoantibodies

Anti-Ro autoantibodies were first discovered in patients with LES by double immunodiffusion, and were named by the first two letters of the patient's name in whose serum they were detected. Later on, a supposedly new autoantibody was detected in patients with SjS by immunodiffusion and named Sjögren's syndrome antigen A (SSA). After an exchange of sera, it was demonstrated that these independently described autoantigens and their respective autoantibodies were indeed identical. The Ro/SSA antigen is a 60kDA ribonucleoprotein that contains one of several short, uridine-rich, stem-loop, structural RNAs termed Y RNAs. There are four molecule species of Y RNAs (Y1, Y3, Y4, Y5 RNA), and stoichiometry studies show that each Ro60/SSA RNP contain a 60kDa protein and one Y RNA in an equal molar ratio. Crystallography studies have demonstrated that Ro60/SSA particle has a ring-shaped structure with a central cavity where misfolded single-stranded noncoding RNAs bind. Therefore, it is hypothesised that the Ro60/SSA particle functions as a quality control check for ncRNAs. However, the function of the Ro60/SSA particle is influenced by the Y RNA species it contains, as Y RNAs may influence the subcellular location of the particle and regulate the binding of other RNAs to the 60kDa protein. In addition, Y RNAs may bind to other proteins, which could also affect the function of the particle. For example, La protein is physically associated with the Ro60/SSA particle, in part because of Y RNA binding [333].

La or Sjögren syndrome antigen B (SSB) was first detected by immunodiffusion in sera from patients with SjS. The La/SSB protein is a 47kDa protein involved in diverse aspects of RNA metabolism by binding precursor RNA molecules or acting as an RNA-chaperone in order to protect them from nuclease-mediated decay and to facilitate their correct processing, including folding and maturation by specific ribonucleases. La is normally part of the Ro/La complex, constituted by the 60kDa Ro (Ro60/SSA), 52kDa Ro (Ro52), and La (SSB), as well

as one of the four Y RNAs. As a result, normally, RNA immunoprecipitation of anti-Ro60/SSA results in immunoprecipitation of the four Y RNAs, while immunoprecipitation of anti-La results in immunoprecipitation of Y RNAs together with various newly synthesised RNA transcripts [334].

Both the presence of anti-Ro60/SSA or anti-La/SSB is considered as a criterion for classification of a patient with SjS, as they are both highly associated with this autoimmune disease. However, although anti-La/SSB is highly specific for SjS, anti-Ro60/SSA is also detected in patients with other autoimmune diseases such as LES, RA, inflammatory myopathies and SSc, among others. In the case of SSc, anti-Ro60/SSA autoantibodies are associated with lcSSc and good prognosis in patients with SSc-SjS overlap syndrome and anti-centromere positivity. However, it seems that when anti-Ro60/SSA is not detected in combination with a SSc-specific autoantibody such as anti-centromere, anti-Scl70 or anti-RNApol III autoantibody, or is detected together with anti-U1 snRNP, it could be associated with a higher risk of presenting ILD [333–335].

1.7.12. Anti-IFI16 autoantibodies

Similar to anti-Ro60/SSA autoantibodies, more recently anti-interferon gamma inducible protein 16 (IFI16) autoantibodies have also been described as SSc-associated autoantibodies [67, 336–338]. In fact, anti-IFI16 autoantibodies have been described as the third most prevalent autoantibodies in SSc patients, being detected in 20-30% of SSc patients [67, 336, 337]. However, this autoantibody is detected in a great number of patients with other autoimmune diseases, with a prevalence of 20-60% in SLE [339–341], 50-70% in SjS [336, 342], 0-10% in RA and 3% in SSc-PM overlap syndrome [339, 342].

Up-regulation of IFN-inducible genes, also known as “IFN signature”, has been reported in SSc and other systemic autoimmune diseases [343]. *IFI16* is a member of the HIN200 gene family, which encodes evolutionarily-related nuclear phospho-proteins IFI16, pyrin and HIN domain-containing protein 1 (PYHIN1), myeloid cell nuclear differentiation antigen (MNDA) and AIM2. These proteins act as innate pattern recognition receptors that sense dsDNA from invading pathogens in both the nucleus and cytoplasm [344, 345]. In line with this, IFI16 protein has been found to be up-regulated across several transcriptomic studies in specific subsets of SSc patients [346–348]. Moreover, elevated levels of circulating IFI16 have also been found in SSc, SLE, SS and RA patients [349]. Therefore, it has been suggested that the high expression of this protein could lead to the development of anti-IFI16 autoantibodies.

Although anti-IFI16 are not SSc-specific, as they are present in 25%-30% of SSc patients who are negative for anti-Scl70 and anti-centromere autoantibody, and they have been associated with lcSSc, they could be useful prognostic biomarkers [67, 336]. Moreover, it has been shown that anti-centromere positive SSc patients who were also positive for anti-IFI16 autoantibodies had an increased risk of digital vascular events during the course of the disease, further supporting the idea that this autoantibody could be a good prognostic biomarker in SSc [338]. As IFI16 is not highly expressed by HEp-2 cells, no association is found between anti-IFI16 autoantibody positivity and any ANA IIF pattern, so a specific test is required for their detection [67].

2. HYPOTHESES

SSc-related autoantibodies are associated with clinical phenotypes that define different subsets of SSc patients, serving as good prognostic biomarkers.

On the basis of this concept, we describe the hypotheses of this thesis by which we postulate that:

1. SSc patients in which no specific autoantibodies are detected do indeed present not identified autoantibodies associated with specific clinical manifestations.
2. SSc-related autoantibodies not detectable by commercial assays are associated with clinical phenotypes that define different subsets of SSc patients.

Therefore, identifying SSc-related autoantibodies and their association with specific clinical manifestations of a heterogeneous disease as SSc will enable better classification of patients in different subsets and the establishment of distinct follow-up approaches. More specifically, this thesis focuses on the development of novel autoantibody detection strategies based on the gold standard assay IP and ELISA and their application on SSc patients in which commercial assays detect no specific autoantibodies, as we postulate that these patients do indeed present not identified autoantibodies.

3. OBJECTIVES

The specific objectives of this thesis are:

1- To develop novel autoantibody detection methods based on the gold standard assay IP:
   a. To develop a novel RNA IP assay coupled with high throughput sequencing to identify new SSc-related autoantibodies against ribonucleoproteins.
   b. To develop a novel protein IP assay based on bio-orthogonal metabolic labelling to identify new SSc-related autoantibodies.
   c. Testing of a cohort of SSc patients in which no specific autoantibodies were detected. Identification and characterisation of new SSc-related autoantibodies and definition of clinical phenotypes associated with them.

2- To develop a novel ELISA assay for anti-IFI16 autoantibody detection.
   a. Testing of a cohort of SSc patients in which no specific autoantibodies were detected and analysis of the clinical manifestations associated with anti-IFI16 autoantibodies.
   b. Testing of a cohort of SSc patients positive for anti-centromere and analysis of the clinical manifestations associated with the co-positivity of both autoantibodies.

4. RESULTS

4.1. Chapter 1

4.1.1. Objective

To develop novel autoantibody detection methods based on the gold standard assay IP:

1- To develop a novel RNA IP assay coupled with high throughput sequencing to identify new SSc-related autoantibodies against ribonucleoproteins.
2- To develop a novel protein IP assay based on bio-orthogonal metabolic labelling to identify new SSc-related autoantibodies.
3- Testing of a cohort of SSc patients in which no specific autoantibodies were detected. Identification and characterisation of new SSc-related autoantibodies and definition of clinical phenotypes associated with them.

4.1.2. Articles

Expanding the landscape of Systemic Sclerosis-related autoantibodies through RNA immunoprecipitation coupled with massive parallel sequencing. Perurena-Prieto J, Sanz-Martínez MT, Viñas-Giménez L, Codina-Clavaguera C, Triginer L, Gordillo-González F, Andrés-León E, Batlle-Masó L, Martin J, Selva-O'Callaghan A, Pujol R, McHugh NJ, Tansley SL, Colobran R, Guillen-Del-Castillo A, Simeón-Aznar CP

Journal of Autoimmunity, Submitted.

Anti-nuclear valosin-containing protein-like autoantibody is associated with calcinosis and higher risk of cancer in systemic sclerosis. Perurena-Prieto J, Viñas-Giménez L, Sanz-Martínez MT, Selva-O'Callaghan A, Callejas-Moraga EL, Colobran R, Guillén-Del-Castillo A, Simeón-Aznar CP.

Rheumatology (Oxford). 2024 Aug 1;63(8):2278-2283. PMID: 37769243. DOI: 10.1093/rheumatology/kead520.

4.1.3. Previous considerations

The interest in the utility of autoantibodies as diagnostic and prognostic biomarkers from both the Translational Immunology and the Systemic Autoimmune Diseases research groups of the Vall d'Hebron Barcelona Hospital Campus led to the possibility of testing the vast majority of SSc-related autoantibodies by commercial methods at our laboratory. After some years of using these methods, we realised that in a significant group of SSc patients from our cohort presenting heterogeneous clinical manifestations, these assays detected no specific SSc-related autoantibodies. Up to 20% of SSc patients from our cohort were included in this group, while only 10% were ANA negative by IIF, similar to what is reported in the literature [350]. As the presence of serum autoantibodies is a serological hallmark of SSc, the lack of detection of specific autoantibodies in this large group of patients, together with the presence of ANAs by IIF on half of them, led us to believe that these patients could present specific autoantibodies that were not detected by commercial assays.

On this basis, as protein and RNA IP are still considered the gold-standard techniques for detecting the majority of specific SSc-related autoantibodies due to the limited sensitivity and specificity of commercial assays [93, 94, 193], the first objective of the project was to develop novel assays based on these two techniques to be able to identify already described and possible new SSc-related autoantibodies.

In the last decades, RNA IP coupled with massive parallel RNA sequencing (RIP-Seq) has opened the possibility of studying ribonucleoprotein interactions with RNA molecules with unprecedented depth [351]. As RNA IP is based on immunoprecipitation of ribonucleoproteins by autoantibodies, the development of a new RIP-Seq assay addressed to autoantibody detection could lead to the discovery of novel autoantibodies in SSc. Therefore, we decided to develop a new RIP-Seq assay for autoantibody detection that was validated by comparison with the gold-standard traditional RNA IP methodology. On the other hand, as traditional protein IP relies on the use of radioisotopes, it makes it impossible for many laboratories to use this gold-standard assay, as it implies having legal authorisation for the use of radioisotopes, special waste disposal protocols and well-trained personnel. In this line, we resolved that it would be very useful to develop a new protein IP method that avoids the utilisation of radioisotopes by using biorthogonal noncanonical amino acid tagging that could be easily implemented in routine laboratories.

4.1.4. Summary of the results

In total, 307 SSc patients consulting during the 2013 to 2020 period in Vall d'Hebron University Hospital were included in the study. First, all recollected samples were tested by commercial assays. Sixty-eight patients from the cohort (22.1%) were negative for all SSc-specific autoantibodies, whereas only 36 (11.7%) were ANA negative by IIF. These 68 samples were tested by RNA IP and RIP-Seq. When tested by traditional RNA IP, 16 patients were positive for already known SSc-specific autoantibodies. In particular, six patients were positive for anti-Th/To, one for anti-U3 snoRNP and nine for anti-U11/U12 snRNP autoantibodies. The six patients positive for anti-Th/To autoantibodies presented the associated homogeneous nucleolar pattern (AC-8) by IIF, while the only patient positive for anti-fibrillarin presented the characteristic nucleolar fibrillar staining pattern (AC-9). Patients with anti-U11/U12 snRNP did not exhibit a specific IIF pattern: six were ANA-negative, two had a nuclear speckled pattern, and one showed a nucleolar pattern.

Fifty-seven samples of the 68 samples assayed by RNA IP were also tested by a newly developed RIP-Seq methodology. Following this approach, 197 RNA molecules were detected as possible candidates of forming part of ribonucleoproteins targeted by SSc-related autoantibodies. All RNAs previously found to be immunoprecipitated by known SSc-related autoantibodies (7-2 RNA, 8-2 RNA, U3 snoRNA, U11 snRNA and Y RNAs) were selected as candidate RNAs by RIP-Seq. Furthermore, these RNA molecules were almost exclusively enriched in samples that tested positive for the corresponding autoantibody by RNA IP. However, RIP-Seq detected Y RNAs, associated with anti-Ro60 autoantibodies, in a patient who tested negative for anti-Ro60 using RNA-IP (**Table 7**).

C/D box snoRNAs were the most abundant final RNA candidates obtained by RIP-Seq. The most abundant C/D box snoRNP, U3 snoRNP, has been reported to be immunoprecipitated by anti-fibrillarin autoantibodies. The only sample in which U3 snoRNA immunoprecipitation was detected by RIP-Seq was the one positive for anti-anti-U3 snoRNP by traditional RNA IP. However, in our study, we found that samples exhibiting nucleolar fibrillar staining pattern on IIF (AC-9) that did not immunoprecipitate U3 snoRNA, neither by traditional RNA IP nor RIP-Seq, immunoprecipitated a wide repertoire of different C/D box snoRNA molecules (**Table 7**). Moreover, patients with autoantibodies against C/D box snoRNPs exhibited a heterogeneous profile, as shown by the different RNA immunoprecipitation patterns determined by RNA-IP and RIP-Seq. Of particular note, these differing patterns were associated with two distinct clinical phenotypes: some patients exhibited a more severe

phenotype similar to that reported for classical anti-fibrillarin autoantibodies, while others had a much milder phenotype. Commercial tests failed to distinguish between these groups, as all patients displayed the same AC-9 IIF pattern and were variably positive for anti-fibrillarin by immunoblotting. However, RNA IP was able to differentiate between the two subsets: patients whose samples showed immunoprecipitation of U3, U8, or U13 snoRNA presented higher rates of dcSSc, heart, lung, arthritis and myositis involvement when compared to those not immunoprecipitating any RNA molecule. Furthermore, weighted correlation network analysis (WGCNA) of co-immunoprecipitation of RNA molecules detected by RIP-Seq revealed that patients presenting an AC-9 ANA pattern by IIF with a more severe clinical phenotype clustered together, while patients with the same ANA pattern but a milder phenotype did not.

In addition, we identified several novel autoantibody targets by RIP-Seq, including H/ACA box snoRNPs, mitochondrial tRNA synthetases, 7SK snRNP and the cytoplasmic vault complex (**Table 7**). However, the small number of patients with potential new autoantibodies prevented us from establishing statistically significant associations between these autoantibodies and specific clinical phenotypes or ANA patterns by IIF.

Of the 52 samples that tested negative for all SSc-specific autoantibodies by commercial assays and traditional RNA IP, 51 were available for further analysis by a newly developed protein IP based on biorthogonal metabolic labelling. When assayed by this novel technique, five samples showed a 100-110kDa band that was demonstrated to be nuclear valosin-containing protein-like (NVL) by MS analysis and IP coupled with western blot (IP-WB) (**Table 7**). As all these five patients showed a homogeneous nucleolar pattern (AC-8) by IIF, patients from the initial cohort of 307 patients with a nucleolar pattern were also tested for anti-NVL reactivity by IP-WB. One additional patient initially considered anti-PM/Scl positive due to a faint reactivity against this protein in a commercial immunoblot and a compatible IIF pattern, was demonstrated to also be anti-NVL positive by IP-WB. More importantly, we found that anti-NVL autoantibody was associated with a specific clinical phenotype characterised by a higher prevalence of calcinosis and cancer, specifically synchronous cancer.

Although not included in the published articles, patients immunoprecipitating other proteins were also found by the newly developed assay. Specifically, one patient was shown to be positive for anti-RuvBL1/2, one positive for anti-Ki/SL and one positive both for anti-RPA and anti-Ki/SL autoantibodies. Moreover, we detected protein bands with a MW not corresponding to any known SSc-related autoantigen in 11 additional samples, indicating that these patients do indeed present ANAs that have not previously been described (**Figure**

16). However, no more than two patients presented the same immunoprecipitation pattern and were not further studied (Table 7).

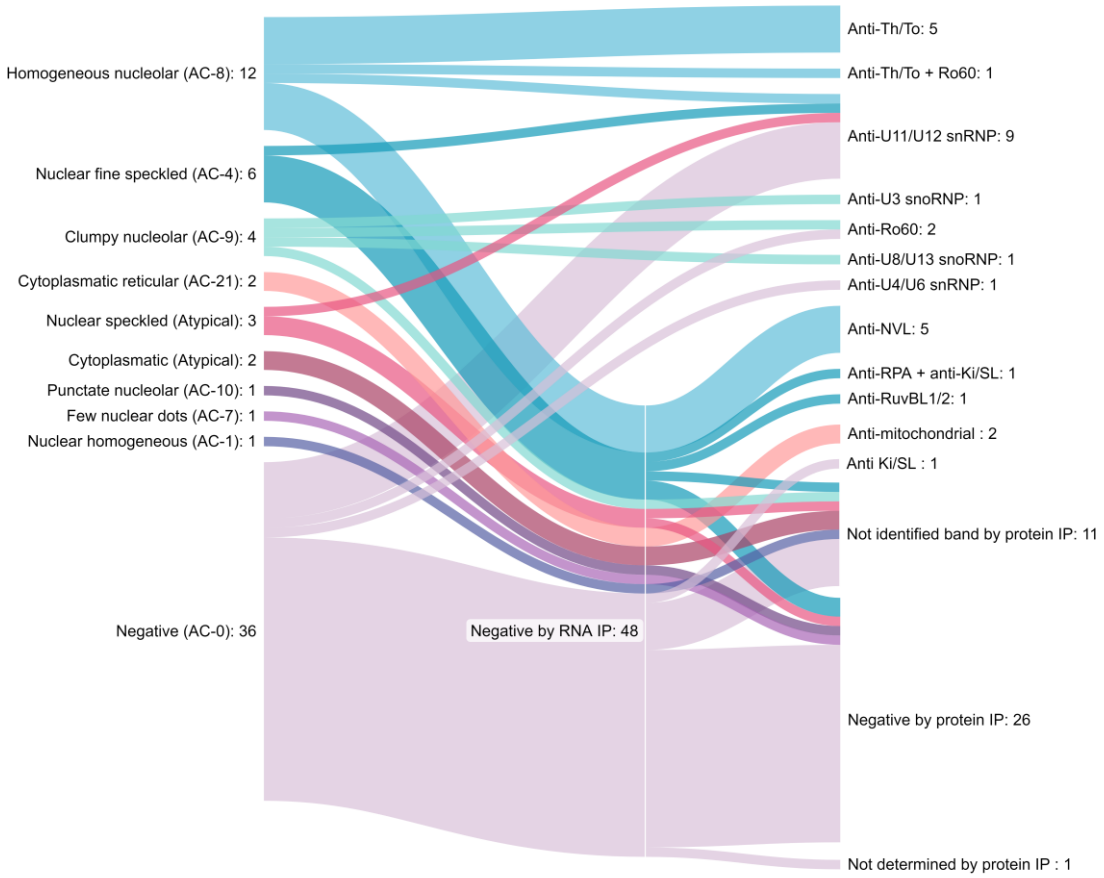


Figure 16. Sankey diagram of IIF patterns and detected specific SSc-related autoantibodies. Graphical summary of the relation between ANA patterns identified by IIF and RNA IP and protein IP results. Results from RIP-Seq were omitted as grouping patients would have been impossible due to the large number of different autoantibodies detected by this approach. Created with *www.SankeyMATIC.com*.

Considering the IIF pattern of the 68 patients tested by RNA IP and protein IP, all patients with a homogeneous nucleolar pattern (AC-8) tested positive for at least one autoantibody by the used methodologies (Figure 16, Table 8). Of note, 50.0% (n=6) were positive for anti-Th/To autoantibodies, while 41.7% (n=5) were positive for the newly identified anti-NVL autoantibody. Additionally, although not all patients presenting a clumpy nucleolar (AC-9) ANA pattern were tested by RIP-Seq, three of the four patients that were analysed by this approach presented reactivity against different C/D box snoRNPs (Table 7). On the other hand, patients with nuclear fine speckled (AC-4) ANA pattern by IIF were positive against different autoantibodies, as anti-U11/U12 snRNP, RuvBL1/2 and RPA, while some still remained negative for any specific autoantibody despite the used approaches (Figure 16, Table 8). Nevertheless, as already mentioned, anti-U11/U12 snRNP autoantibodies were not

Table 7. Summary of the results obtained by IIF, RNA IP, RIP-Seq and protein IP of the 68 patients that were initially considered negative for all specific autoantibodies tested by commercial assays. In the case of detection of a not identified band by protein IP the approximate molecular weight of the protein has been noted.

No.	Patient	IIF ANA pattern	RNA IP	RIP-Seq	Protein IP	Anti-IFI16
1	SSc_01	Homogeneous nucleolar (AC-8)	Anti-Th/To	Anti-Th/To + Anti-Vault complex	ND	N
2	SSc_02	N (AC-0)	N	Anti-MT-TM	100kDa	P
3	SSc_03	Cytoplasmatic (Atypical)	N	Anti-Vault complex	90kDa	N
4	SSc_04	N (AC-0)	Anti-U11/U12 snRNP	Anti-U11/U12 snRNP	ND	P
5	SSc_05	N (AC-0)	N	N	N	N
6	SSc_06	N (AC-0)	N	N	N	N
7	SSc_07	N (AC-0)	N	N	ND	N
8	SSc_08	Nuclear fine speckled (AC-4)	N	N	120kDa	N
9	SSc_09	N (AC-0)	N	N	80kDa	P
10	SSc_10	Nuclear fine speckled (AC-4)	N	N	N	N
11	SSc_11	Cytoplasmatic reticular (AC-21)	N	N	Anti-mitochondrial	N
12	SSc_12	N (AC-0)	Anti-U11/U12 snRNP	Anti-U11/U12 snRNP + Anti-U5 snRNP	ND	N
13	SSc_13	N (AC-0)	N	N	N	N
14	SSc_14	Nuclear speckled (Atypical)	N	Anti-Ro60 + Anti-Vault complex + Anti-7SK snRNP	N	N
15	SSc_15	Homogeneous nucleolar (AC-8)	N	N	Anti-NVL	N
16	SSc_16	Punctate nucleolar (AC-10)	N	N	N	N
17	SSc_17	N (AC-0)	N	N	N	N
18	SSc_18	N (AC-0)	N	Anti-Vault complex	120kDa	N
19	SSc_19	N (AC-0)	N	Anti-7SK snRNP	N	N
20	SSc_20	N (AC-0)	N	N	N	N

Table 7 (continued)

No.	Patient	IIF ANA pattern	RNA IP	RIP-Seq	Protein IP	Anti-IFI16
21	SSc_21	Nuclear homogeneous (AC-1)	N	N	25, 50kDa	N
22	SSc_22	Homogeneous nucleolar (AC-8)	N	N	Anti-NVL	N
23	SSc_23	N (AC-0)	Anti-U4/U6 snRNP	Anti-U5 snRNP + Anti-U4 snRNP	N	N
24	SSc_24	Nuclear fine speckled (AC-4)	N	N	N	N
25	SSc_25	N (AC-0)	N	N	N	N
26	SSc_26	N (AC-0)	N	Anti-7SK	N	N
27	SSc_27	Cytoplasmatic reticular (AC-21)	N	Anti-MT-TM + Anti-SRP	Anti-mitochondrial	N
28	SSc_29	N (AC-0)	Anti-U11/U12 snRNP	Anti-U11/U12 snRNP + Anti-U5 snRNP	ND	N
29	SSc_30	Clumpy nucleolar (AC-9)	Anti-U3 snoRNP	Anti-U3 snoRNP	ND	N
30	SSc_31	Nuclear fine speckled (AC-4)	N	N	Anti-RuvBL1/2	N
31	SSc_32	N (AC-0)	N	N	N	N
32	SSc_33	Nuclear fine speckled (AC-4)	Anti-U11/U12 snRNP	Anti-U11/U12 snRNP + Anti-U5 snRNP	ND	P
33	SSc_34	N (AC-0)	N	Anti-MT-TM	N	P
34	SSc_35	N (AC-0)	N	N	N	P
35	SSc_36	N (AC-0)	N	N	N	N
36	SSc_37	Nuclear speckled (Atypical)	N	N	90kDa	N
37	SSc_38	Few nuclear dots (AC-7)	N	N	N	N
38	SSc_39	N (AC-0)	Anti-U11/U12 snRNP	Anti-U11/U12 snRNP + Anti-U5 snRNP + Anti-SRP	ND	N
39	SSc_40	N (AC-0)	N	Anti-MT-TM	100kDa	N
40	SSc_41	N (AC-0)	Anti-U11/U12 snRNP	Anti-U11/U12 snRNP	ND	N
41	SSc_42	Homogeneous nucleolar (AC-8)	Anti-Th/To	Anti-Th/To + Anti-Vault complex	ND	P

Table 7 (continued)

No.	Patient	IIF ANA pattern	RNA IP	RIP-Seq	Protein IP	Anti-IFI16
42	SSc_43	Homogeneous nucleolar (AC-8)	N	N	Anti-NVL	N
43	SSc_44	Nuclear speckled (Atypical)	Anti-U11/U12 snRNP	Anti-U11/U12 snRNP + Anti-H/ACA snoRNP	ND	N
44	SSc_45	N (AC-0)	N	N	N	N
45	SSc_46	Clumpy nucleolar (AC-9)	Anti-U8 snoRNA + Anti-U13 snoRNA	Anti-C/D box snoRNP	40kDa	P
46	SSc_47	N (AC-0)	Anti-U11/U12 snRNP	Anti-U11/U12 snRNP + Anti-MT-TH	ND	N
47	SSc_48	Clumpy nucleolar (AC-9)	Anti-Ro60	Anti-Ro60 + Anti-C/D box snoRNP + Anti-Vault complex + Anti-MT-TE	N	P
48	SSc_49	N (AC-0)	N	N	N	N
49	SSc_50	N (AC-0)	N	N	N	N
50	SSc_51	Homogeneous nucleolar (AC-8)	Anti-Th/To	Anti-Th/To + Anti-Vault complex	ND	N
51	SSc_52	N (AC-0)	Anti-Ro60	Anti-Ro60	N	N
52	SSc_53	Homogeneous nucleolar (AC-8)	Anti-Th/To	Anti-Th/To + Anti-Vault complex + Anti-MT-TM	ND	N
53	SSc_54	Nuclear fine speckled (AC-4)	N	N	Anti-RPA + Anti-Ki/SL	P
54	SSc_55	N (AC-0)	N	N	125kDa	N
55	SSc_56	Homogeneous nucleolar (AC-8)	N	N	Anti-NVL	N
56	SSc_57	N (AC-0)	N	N	N	N
57	SSc_58	Homogeneous nucleolar (AC-8)	Anti-Th/To + Anti-Ro60	Anti-Th/To + Anti-Ro60 + Anti-U6 snRNP + Anti-MT-TM + Anti-MT-TE + Anti-7SK snRNP	ND	N
58	SSc_66	Homogeneous nucleolar (AC-8)	Anti-U11/U12 snRNP	ND	ND	ND
59	SSc_68	N (AC-0)	N	ND	N	ND
60	SSc_69	N (AC-0)	N	ND	Anti-Ki/SL	ND

Table 7 (continued)

No.	Patient	IIF ANA pattern	RNA IP	RIP-Seq	Protein IP	Anti-IFI16
61	SSc_70	Clumpy nucleolar (AC-9)	N	ND	40kDa	ND
62	SSc_71	Homogeneous nucleolar (AC-8)	Anti-Th/To	ND	ND	ND
63	SSc_72	Homogeneous nucleolar (AC-8)	N	ND	Anti-NVL	ND
64	SSc_73	N (AC-0)	N	ND	N	ND
65	SSc_74	Cytoplasmatic (Atypical)	N	ND	40kDa	ND
66	SSc_75	N (AC-0)	N	ND	N	ND
67	SSc_76	N (AC-0)	N	ND	N	ND
68	SSc_77	N (AC-0)	N	ND	N	ND

N, negative; ND, not determined; P, positive; Vault complex, cytoplasmatic vault complex.

found to associate with AC-4 or any other specific ANA pattern by IIF, as more than half of the positive patients were ANA negative by IIF (**Figure 16**). Overall, considering the results of RNA IP and protein IP, 26 patients, 8.5% of the overall cohort, tested negative for all autoantibodies despite the used approach. Of these 26 patients, 21 (80.8%) were ANA negative by IIF (AC-0) (**Figure 16**). Interestingly, 4 of these 26 patients presented reactivity against different ribonucleoproteins when tested by RIP-Seq (**Table 7**).

Table 8. Summary of results obtained from the combination of traditional RNA IP and protein IP by subsets of patients showing different ANA patterns by IIF Results from RIP-Seq were omitted as grouping patients would have been impossible due to the large number of different autoantibodies detected by this approach.

IIF ANA pattern	N (%) of patients	Detected reactivity
Homogeneous nucleolar (AC-8)	5 (42)	Anti-Th/To
	1 (8)	Anti-Th/To + Ro60
	1 (8)	Anti-U11/U12 snRNP
	5 (42)	Anti-NVL
Nuclear fine speckled (AC-4)	1 (17)	Anti-U11/U12 snRNP
	1 (17)	Anti-RPA + Anti-Ki/SL
	1 (17)	Anti-RuvBL1/2
	1 (17)	Not identified band by protein IP
	2 (33)	Negative
Clumpy nucleolar (AC-9)	1 (25)	Anti-U3 snoRNP
	1 (25)	Not identified band by protein IP
	1 (25)	Ro60
	1 (25)	Anti-U8 snoRNP + Anti-U13 snoRNP
Nuclear speckled (Atypical)	1 (33)	Anti-U11/U12 snRNP
	1 (33)	Not identified band by protein IP
	1 (33)	Negative
Cytoplasmatic reticular (AC-21)	2 (100)	Anti-mitochondrial
Cytoplasmatic (Atypical)	2 (100)	Not identified band by protein IP
Punctate nucleolar (AC-10)	1 (100)	Negative
Few nuclear dots (AC-7)	1 (100)	Negative
Nuclear homogeneous (AC-1)	1 (100)	Not identified band by protein IP
Negative (AC-0)	6 (17)	Anti-U11/U12 snRNP
	1 (3)	Anti-Ro60
	1 (3)	Anti-Ki/SL
	1 (3)	Anti-U4/U6 snRNP
	1 (3)	Not determined by protein IP
	5 (14)	Not identified band by protein IP
	21 (58)	Negative

4.1.5. Expanding the landscape of systemic sclerosis-related autoantibodies through RNA immunoprecipitation coupled with massive parallel sequencing

TITLE PAGE

Article type: Research paper

Title: Expanding the landscape of systemic sclerosis-related autoantibodies through RNA immunoprecipitation coupled with massive parallel sequencing

Running Title: Expanding the landscape of SSc-related autoantibodies through RIP-Seq

Authors: Janire Perurena-Prieto[a,b,c], María Teresa Sanz-Martínez[a,b], Laura Viñas-Giménez[a,b], Claudia Codina-Clavaguera[d,e], Laura Triginer[e], Fernando Gordillo-González[f,1], Eduardo Andrés-León[f], Laura Batlle-Masó [g,h,i], Javier Martin[f], Albert Selva-O'Callaghan[d,e], Ricardo Pujol[c,j], Neil J McHugh[k], Sarah L Tansley[k], Roger Colobran [a,b,c,l*], Alfredo Guillen-Del-Castillo [d,e*], Carmen Pilar Simeón-Aznar[d,e]

[a]Immunology Division, Vall d'Hebron University Hospital (HUVH), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain

[b]Translational Immunology Group, Vall d'Hebron Research Institute (VHIR), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain

[c]Department of Cell Biology, Physiology and Immunology, Autonomous University of Barcelona (UAB), Bellaterra, Spain

[d]Systemic Autoimmune Diseases Unit, Internal Medicine Department, Vall d'Hebron University Hospital (HUVH), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain.

[e]Systemic Autoimmune Diseases Group, Vall d'Hebron Research Institute (VHIR), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain

[f]Institute of Parasitology and Biomedicine "López-Neyra", CSIC (IPBLN-CSIC), Granada, Spain

[g]Infection and Immunity in Pediatric Patients Research Group, Vall d'Hebron Research Institute (VHIR), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain.

[h]Pediatric Infectious Diseases and Immunodeficiencies Unit, Children's Hospital. Hospital Universitari Vall d'Hebron (HUVH), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain.

[i]Pompeu Fabra University (UPF), Barcelona, Spain.

[j]Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain.

[k]Department of Pharmacy and Pharmacology, University of Bath, Bath, UK.

[l]Department of Clinical and Molecular Genetics, Vall d'Hebron University Hospital (HUVH), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain.

[1]Computational Biomedicine Laboratory, Principe Felipe Research Center (CIPF), Valencia, Spain.

*These are both corresponding authors.

Correspondence:roger.colobran@vallhebron.cat, alfredo.guillen@vallhebron.cat

Funding: This study was supported by the Carlos III Health Institute [grants PI16/02088, PI20/00761 and PI22/01804] co-financed by the European Regional Development Fund (ERDF). Laura Batlle-Masó has received research support from the European Union-NextGenerationEU, Ministry of Universities and Recovery, Transformation and Resilience Plan, through a call from Pompeu Fabra University (Barcelona).

Competing interests: All authors declare that they received no financial support or other benefits from commercial sources for the work reported in the manuscript, and have no other financial interests that could create a potential conflict of interest or appearance of a conflict of interest with regard to the study.

Keywords: systemic sclerosis, autoantibodies, RIP-seq, RNA immunoprecipitation, anti-fibrillarin.

ABSTRACT

Objectives: Systemic sclerosis (SSc)-related autoantibodies are widely used diagnostic and prognostic biomarkers. This study aimed to develop a new assay for detecting anti-ribonucleoprotein autoantibodies in SSc based on RNA immunoprecipitation (RNA IP) coupled with massive parallel sequencing.

Methods: Serum samples and clinical data were collected from 307 SSc patients. Among these, 57 samples underwent analysis using a new protocol that combines RNA IP with massive parallel sequencing (RIP-Seq). Filtering strategies and statistical outlier detection methods were applied to select RNA molecules that could represent novel ribonucleoprotein autoantigens associated with SSc.

Results: Among the 30,966 different RNA molecules identified by RIP-Seq in 57 SSc patients, 197 were ultimately selected. These included all RNA molecules previously identified by RNA IP, which were found to exhibit high counts almost exclusively in samples positive for the autoantibodies associated to the corresponding RNA molecule, indicating high sensitivity and specificity of the RIP-Seq technique. C/D box snoRNAs were the most abundant RNA type identified. The immunoprecipitation patterns of the detected C/D box snoRNAs varied among patients and could be associated with different clinical phenotypes. In addition, other ribonucleoproteins were identified, which could be potential targets for previously undescribed SSc-related autoantibodies. These include H/ACA box snoRNPs, vault complexes, mitochondrial tRNA synthetases, and 7SK snRNP.

Conclusion: A novel RIP-Seq assay has been developed to detect autoantibodies targeting ribonucleoprotein complexes in SSc patients. This method successfully identified RNA molecules associated with ribonucleoproteins known to be targeted by SSc-related autoantibodies, validating both the assay and the analysis strategy. Additionally, this approach uncovered RNA molecules associated with ribonucleoproteins that were not previously identified as targets of SSc patients' sera, suggesting potential new autoantibody candidates in this disease.

INTRODUCTION

Systemic sclerosis (SSc) is a systemic autoimmune disease with a multifaceted etiology, characterized by inflammation and fibrosis, predominantly affecting the skin, microcirculation, and internal organs [1]. The presence of serum antinuclear autoantibodies (ANA) is a serological hallmark of SSc, being detected in more than 95% of patients. In addition, SSc-related autoantibodies are linked to the clinical presentation, organ involvement, and prognosis of the disease [1-3] [62, 350, 352]. Due to the significant heterogeneity of SSc, it is essential to identify biomarkers, such as SSc-related autoantibodies, that can help predict the clinical course of each patient.

Over the last few decades, a number of commercial assays have become available for detecting SSc-related autoantibodies. However, due to the limited sensitivity and specificity of these tests, protein immunoprecipitation and RNA immunoprecipitation (RNA IP) remain the gold-standard techniques for this purpose [4-6]. Both these methods rely on incubating patient sera with protein extracts obtained from human cell lines. Autoantibodies present in the sera bind to autoantigens in the protein extracts, causing them to immunoprecipitate. In the case of RNA IP, the aim is to detect autoantibodies that bind to ribonucleoprotein autoantigens within a cell extract. Ribonucleoproteins, which are involved in many cellular processes, are macromolecular complexes containing small RNA molecules essential for their function [7]. Following immunoprecipitation, nucleic acids are extracted, yielding specific RNA molecules associated with ribonucleoproteins recognized by autoantibodies. If the autoantibodies present in the sera have been previously described, the corresponding proteins and RNAs can be identified by analyzing their molecular weights and comparing them with those in reference sera known to contain specific autoantibodies.

Protein immunoprecipitation and RNA IP can also identify new autoantibodies by detecting proteins and RNA molecules bound by previously undescribed autoantibodies. Protein immunoprecipitation combined with mass spectrometry has long been the preferred method for discovering novel autoantibodies in SSc [8,9]. Although new autoantibodies can also be identified using traditional RNA IP—by sequencing specific RNA molecules eluted from gels [10]—this method is not commonly used due to its technical challenges.

In recent decades, the development of affordable massive parallel sequencing technologies has provided new opportunities for research in molecular biology and genetics. This includes a more in-depth study of ribonucleoprotein interactions with RNA molecules. One such approach that can be used for this purpose is RNA IP coupled with next generation RNA

sequencing (RIP-Seq) [11]. As RIP-Seq is expected to be more sensitive than traditional RNA IP for detecting RNA molecules, developing a dedicated RIP-Seq protocol for autoantibody detection could potentially lead to the discovery of new autoantibodies in SSc. Nonetheless, RIP-Seq technology has not yet been applied to detect autoantibodies in systemic autoimmune diseases.

The aim of this study was to create a new method for detecting autoantibodies in SSc based on a modified RNA IP protocol coupled with high-throughput RNA sequencing. After developing the technique, it was validated by comparing the results with those of traditional RNA IP. The sensitivity of the new RIP-Seq assay was superior and showed potential for uncovering novel autoantibodies in SSc.

PATIENTS AND METHODS

A complete description of the patients and methods is provided in the supplementary material.

In total, 307 SSc patients consulting during the 2013 to 2020 period in Vall d'Hebron University Hospital were included in the study. All patients met the LeRoy and Medsger criteria for SSc [12], and 84.0% (n=258) fulfilled the 2013 ACR/EULAR classification criteria [13]. The study was approved by the hospital Ethics Committee for Clinical Research (PG(AG)07/2015), and written informed consent was obtained from all participants. Patients who tested negative for all autoantibodies as assessed by commercial assays were retested using traditional RNA IP and the new RIP-Seq technique (**Figure 1**).

RESULTS

Identification of SSc-related autoantibodies by commercial assays and traditional RNA IP

The 307 SSc patients were analyzed by immunoblotting (IB), chemiluminescence immunoassay (CLIA), and indirect immunofluorescence (IIF) commercial assays. Overall, 40.4% (n=124), 19.9% (n=61) and 8.1% (n=25) were positive for anti-centromere, anti-Scl70 and anti-RNApol III autoantibodies, respectively. Less common autoantibodies (anti-PM/Scl, fibrillarin, Th/To, Ku, or U1-RNP) were found in 9.4% (n=29) of patients. Sixty-eight SSc patients (22.1%) tested negative for all the aforementioned autoantibodies (**Supplementary Figure S1**). Among these 68 patients, 17 (25%) presented a predominantly nucleolar pattern (AC-8, 9, 10), nine (13.2%) a nuclear speckled pattern (AC-4, 5, atypical), four (5.9%) a cytoplasmatic pattern (AC-21, atypical), one (1.5%) a homogeneous nuclear

pattern (AC-1), and one (1.5%) a few nuclear dots pattern (AC-7) using IIF on Hep-2 cells, whereas 36 (52.9%) patients tested negative (AC-0) (**Supplementary Table S1**).

The RNA IP results in samples from these 68 patients are detailed in **Supplementary Table S1**. Six samples showed immunoprecipitation of 7-2 and 8-2 RNA, identifying these patients as positive for anti-Th/To antibodies. All six samples displayed a homogeneous nucleolar pattern (AC-8) on IIF (**Figure 2B**). Among patients with either a nucleolar (AC-8, 9, 10) or specifically homogeneous nucleolar (AC-8) IIF pattern, 35.3% (6/17) and 50% (6/12) were positive for anti-Th/To antibodies, respectively, on RNA IP analysis. One patient tested positive for anti-fibrillarin, as indicated by the presence of U3 RNA on RNA IP (**Figure 2A**) and a clumpy nucleolar pattern (AC-9) by IIF (**Supplementary Figure S2**). Anti-U11/U12 RNP autoantibodies were detected in nine patients whose sera immunoprecipitated U11 RNA (**Figure 2C**), but U12 RNA was not clearly identified in all samples. Patients with anti-U11/U12 RNP did not exhibit a specific IIF pattern: six were ANA-negative, two had a nuclear speckled pattern, and one showed a nucleolar pattern (**Supplementary Table S1**). Samples from three patients were found to immunoprecipitate Y1 RNA, Y3 RNA, Y4 RNA, and Y5 RNA, which are associated with anti-Ro60 antibodies (**Supplementary Figure S3**). Additionally, samples SSc_23 and SSc_46 displayed two bands with electrophoretic mobility similar to that of U4 and U6 snRNA.

Evaluation of RNA immunoprecipitation coupled with massive parallel sequencing for the detection of known SSc-related autoantibodies (RIP-Seq)

Samples from 57 SSc patients, out of the 68 testing negative for all autoantibodies assessed using commercial assays, were retested by RIP-Seq (**Supplementary Figure S1**). Sera from three healthy donors were included as negative controls. Overall, 30,966 different RNA molecules (RNAs) were detected in at least one of the 60 samples tested. A large number of RNAs were present only at low counts in all samples and were considered nonspecific immunoprecipitation events (ie, background noise). Based on the distribution of median counts of protein-coding RNAs, which are not specifically immunoprecipitated as they are not stably associated with any ribonucleoprotein, a cut-off of 100 median counts was established (**Supplementary Figure S4**). The 2239 RNAs with median counts above this cut-off were selected for further analysis.

Since SSc-related autoantibodies are not present in all patients, we aimed to identify RNA molecules with high counts in a subset of samples, and low counts in the others. To this end, we used four methods for outlier detection, searching for count outliers related to each RNA

molecule across samples. Following this process, 400 molecules with at least one outlier value by the four methods were selected (**Supplementary Figure S5**).

Despite these filtering criteria, 187 protein-coding RNAs remained in the selection. To further refine the list and exclude nonspecific RNA molecules, the 400 RNAs were ranked using a ratio calculated as the square of the highest outlier count divided by the median count of each specific RNA molecule across all samples (max^2/median ratio). This approach aimed to identify RNAs that showed a significant difference in abundance between samples precipitating the molecule and those that did not. The effectiveness of this criterion was evaluated using a receiver operating characteristic (ROC) curve, assuming that all protein-coding RNA molecules were false positives (**Supplementary Figure S6**). The area under the curve (AUC) was 0.908, and the cut-off with the best sensitivity (80.8%) and specificity (85.6%) was found to be a max^2/median ratio $\geq$ 21,779. Applying this cut-off, 200 RNAs were selected (**Supplementary Table S2**).

To verify the validity of the technique and filtering criteria, we compared the RIP-Seq results with those of RNA IP. All RNAs previously identified by RNA IP (7-2 RNA, 8-2 RNA, U3 RNA, U11 RNA, Y1 RNA, Y3 RNA, Y4 RNA and Y5 RNA) were among the 200 selected RNA candidate molecules detected by RIP-Seq (**Supplementary Table S2**). The 7-2 RNA and 8-2 RNA molecules were highly enriched only in samples containing anti-Th/To autoantibodies (**Figure 3A**), while U3 RNA showed high counts exclusively in the single patient positive for anti-fibrillarin (**Figure 3B**). U11 RNA had significantly higher counts in samples positive for anti-U11/U12 RNP (**Figure 3C**). In contrast, U12 RNA enrichment was not specific, as it was detected in both anti-U11/U12 RNP positive samples and other samples (**Supplementary Figure S7**). Y1 RNA, Y3 RNA, Y4 RNA, and Y5 RNA were enriched in samples that had immunoprecipitated these molecules, but were also present in a patient who tested negative by RNA IP (**Supplementary Figure S2, Figure 3D, 3E**).

Application of RNA immunoprecipitation coupled with massive parallel sequencing for detection of new autoantibodies in SSc

After applying the filters and outlier detection strategies, 200 RNA molecules were selected. This set included RNAs bound by known autoantibodies as well as a variety of other RNAs. Notably, 68.5% of the molecules selected were small nucleolar RNAs (snoRNAs) or snoRNA host genes (SNGH). C/D box snoRNAs were particularly prevalent, comprising 59.5% of the selected RNA molecules. The next most represented group consisted of protein-coding RNAs, with 28 protein-coding RNA molecules included despite the filtering criteria. A summary of the frequencies of all RNA types is provided in **Supplementary Table S3**.

Since the number of RNAs in the protein-coding group was higher than expected, we investigated their specific locations of immunoprecipitated sequences within the corresponding genes. We discovered that some of these protein-coding RNAs actually mapped to intronic regions that code for snoRNAs (**Supplementary Figure S8**). In total, we identified 7 protein-coding RNAs that were actually snoRNAs that had been misclassified as such due to the annotation approach used, which prioritized RNA sequence location within gene regions over accurate snoRNA identification (**Supplementary Figure S8**). Consequently, these 7 RNAs were reclassified into the appropriate category (**Supplementary Table S4**). Additionally, the immunoprecipitation results for patient SSc_58 showed a specific RNA intronic sequence from the ANK1 protein-coding gene (**Supplementary Figure S8**), likely corresponding to a small RNA molecule yet to be identified. Considering this further analysis, we updated the overall frequency of RNA types (**Supplementary Table S3**) and finalized a curated list of 197 RNA molecules as potential components of ribonucleoproteins targeted by SSc-related autoantibodies (**Table 1**).

Autoantibodies against C/D box snoRNPs: beyond anti-fibrillarin autoantibodies

Since C/D box snoRNAs were the most abundant candidate RNAs, we aimed to identify which samples specifically immunoprecipitated these molecules. Among the patient samples, SSc_30, SSc_35, SSc_46, SSc_48, and SSc_58 had the highest C/D box snoRNA counts (**Supplementary Figure S9**). However, these snoRNAs were not immunoprecipitated equally across samples. To investigate whether the C/D box snoRNAs and the samples immunoprecipitating them could be clustered, we performed a weighted correlation network analysis (WGCNA), which grouped RNAs into modules based on their co-immunoprecipitation patterns. The RNA molecules assigned to each module are summarized in **Supplementary Table S5**.

We also evaluated the correlations between each sample and the various modules (**Figure 4A, Supplementary Figure S10**). Out of the nine modules generated by the analysis, six (B, D, E, F, G and H) mainly consisted of C/D box snoRNAs and/or SNHGs (**Figure 4B**). Notably, samples SSc_30 and SSc_46 highly correlated with module H. In contrast, SSc_48 was strongly associated with module F, which showed no significant correlations with either SSc_30 or SSc_46 (**Figure 4A, Supplementary Figure S10**). These findings suggest that SSc_30 and SSc_46 may present autoantibodies targeting a common group of snoRNPs, whereas SSc_48 is positive for an autoantibody directed against a different subset of snoRNPs.

The autoantibody commonly known as anti-fibrillarin recognizes U3 snoRNP, the most abundant C/D box snoRNP. U3 snoRNP consists of U3 snoRNA (also denominated U3 RNA) and various associated proteins, including fibrillarin. However, samples from some patients with anti-fibrillarin autoantibodies also immunoprecipitate other snoRNA molecules, specifically U8 and U13 snoRNA. Among the samples analyzed by RNA IP and RIP-Seq, only SSc_30 tested positive for anti-fibrillarin autoantibody. Although SSc_46 and SSc_48 did not immunoprecipitate U3 snoRNA using these methods, both samples showed the same clumpy nucleolar pattern (AC-9) as SSc_30 on IIF (**Supplementary Figure S2**), which is specifically associated with the anti-fibrillarin autoantibody. In addition, SSc_46 immunoprecipitated two RNA molecules that were compatible in size with U8 snoRNA and U13 snoRNA by RNA IP. In contrast, SSc_48 only immunoprecipitated Y RNAs, which are associated with anti-Ro60 autoantibody (**Supplementary Figure S11**). Interestingly, U13 snoRNA (SNORD13) was not a selected RNA candidate in the RIP-Seq analysis (**Table 1**), whereas U8 snoRNA (SNORD118) was predominantly immunoprecipitated by SSc_58 and SSc_48, as indicated by the RIP-Seq data (**Supplementary Table S5**).

In view of these results, and considering that SSc_46 and SSc_48 could present a previously undescribed autoantibody against C/D box snoRNPs, we decided to study all SSc patients with an AC-9 IIF pattern for whom we had an available sample. Thus, in addition to samples SSC_30, SSc_46, and SSc_48, the ten remaining samples with an AC-9 IIF pattern from the 307 patient cohort were analyzed (**Supplementary Figure S2**). Among all 13 samples with an AC-9 IIF pattern, only 3 (23.1%) immunoprecipitated U3 snoRNA by RNA IP, but 8 (61.6%) immunoprecipitated two RNA molecules that could correspond to U8 and U13 snoRNA (**Supplementary Figure S11**). The various RNA immunoprecipitation patterns observed did not correlate with the commercial IB results (**Table 2**). The clinical characteristics of the 13 patients were examined to determine if there were differences between those whose samples immunoprecipitated U3, U8, or U13 snoRNA and the remaining patients showing an AC-9 IIF pattern. While no statistically significant differences were found between these groups, likely due to the limited number of patients, those showing immunoprecipitation of U3, U8, or U13 snoRNAs exhibited more severe clinical features than patients testing negative (**Table 2**). These manifestations included higher rates of heart (67% vs. 25%) and lung (33% vs. 0%) involvement, as well as arthritis (22% vs. 0%), myositis (22% vs. 0%), and diffuse cutaneous systemic sclerosis (dcSSc) (44% vs. 25%). In the subset of patients showing only U3 snoRNA immunoprecipitation, the differences were less pronounced, likely due to the small sample size (n=3).

Autoantibodies against anti-H/ACA box snoRNPs and other possible new ribonucleoprotein targets in SSc

Regarding H/ACA box snoRNAs, the third most represented type of candidate RNA (**Supplementary Table S3**), WGCNA showed that module I was highly enriched with these molecules and with scaRNAs (**Figure 4B**). The sample showing the highest correlation with this module was SSc_44 (**Figure 4A**), which was anti-U11/U12 RNP positive by RNA IP but presented an atypical speckled IIF pattern with peri-chromosomal staining of metaphase plates at very high titers. This suggests the presence of an additional autoantibody. Clinically, patient SSc_44 exhibited dcSSc and interstitial lung disease (ILD).

Finally, to investigate whether additional ribonucleoproteins might be targeted by autoantibodies in SSc, we examined RNA molecules detected by RIP-Seq other than snoRNAs (**Supplementary Table S6**). A cut-off based on the interquartile range of control RNA molecules (7-2 RNA, 8-2 RNA, U11 RNA, U3 RNA and Y RNAs) was established to classify each sample as positive or negative for each RNA molecule **(Figure 4C).** Vault RNAs (VTRNA1-2, VTRNA2-1, VTRNA1-1) were found to be immunoprecipitated in samples positive for anti-Th/To and anti-Ro60, as well as sample SSc_18. Various U5 and U4 snRNAs were immunoprecipitated in samples positive for anti-U11/U12 RNP autoantibodies and in sample SSc_23, positive for U4/U6 snRNP by RNA IP. The analysis also disclosed immunoprecipitation of various mitochondrial tRNAs (MT-TM, MT-TE, MT-TH), RN7SL1, and 7SK snRNA.

DISCUSSION

The use of RNA immunoprecipitation coupled with massive parallel sequencing has provided the means to investigate interactions between ribonucleoproteins and RNA molecules with unprecedented high resolution [11]. In this study, we describe the development and validation of a novel RIP-Seq technique for detecting SSc-related autoantibodies targeting ribonucleoproteins.

In a cohort of 307 SSc patients, 22.1% tested negative for all SSc-related autoantibodies using commercial assays. However, when these negative samples were analyzed using the gold-standard RNA IP technique, six patients tested positive for anti-Th/To autoantibodies. Among patients with a homogeneous nucleolar pattern (AC-8) by IIF who were negative by commercial assays, 50% were found to be anti-Th/To positive. In the overall cohort (n=307), only 2 patients were classified as anti-Th/To positive by the commercial assay, indicating a sensitivity of 25% for detecting this autoantibody, consistent with previous reports [14][95].

Hence, we suggest RNA IP analysis for anti-Th/To autoantibodies in all SSc patients who show an AC-8 IIF pattern and negative testing for SSc-specific autoantibodies by commercial methods.

Anti-U11/U12 RNP was the autoantibody most frequently detected by RNA IP: 13.2% of patients who tested negative by commercial assays showed a positive result for this autoantibody. Although anti-U11/U12 RNP is strongly linked to severe ILD [15,16][97, 314], there are no currently available commercial tests for its detection. Since this autoantibody does not produce a specific IIF pattern, performing RNA IP on SSc samples testing negative for SSc-specific autoantibodies by commercial assays should be considered.

Massive parallel sequencing technology is highly sensitive; therefore, filtering strategies were used to eliminate nonspecific binding and detect RNA molecules presenting high variance. Samples testing positive by RNA IP were also tested by RIP-Seq to validate this new technique and the data analysis approach. All RNAs previously found to be immunoprecipitated by known SSc-related autoantibodies were selected as candidate RNAs by RIP-Seq. Moreover, these RNA molecules were almost exclusively enriched in samples that had tested positive for the corresponding autoantibody by RNA IP. However, RIP-Seq detected Y RNAs, which is associated with anti-Ro60 autoantibodies, in a patient who tested negative for anti-Ro60 using RNA IP. This finding suggests that the RIP-Seq technique may be more sensitive than RNA IP.

On RIP-Seq analysis, U12 RNA enrichment appeared to be nonspecific, as it was also detected in samples lacking anti-U11/U12 RNP according to RNA IP. Typically, the band for U12 RNA is not clearly visible in RNA IP assays of patients positive for anti-U11/U12 RNP autoantibodies [15,17][97, 306]. In fact, U12 RNA was not detected through immunoprecipitation when this autoantibody was first described [17][306]. Nonetheless, it was expected that anti-U11/U12 RNP autoantibodies also immunoprecipitated U12 RNA, since U11/U12 RNP forms part of the minor spliceosome as a stable complex [18][309]. Thus, the failure to detect U12 RNA was thought to be due to its low abundance in cell extracts. However, given the high sensitivity of RIP-Seq, this explanation seems unlikely in relation to our findings. Our results suggest that the absence of specific immunoprecipitation of U12 RNA in conjunction with U11 RNA indicates that anti-U11/U12 RNP-denominated autoantibodies specifically recognize only U11 RNP.

C/D box snoRNAs and H/ACA box snoRNAs were the most abundant final RNA candidates obtained by RIP-Seq. These RNA molecules serve as guides for C/D box and H/ACA box snoRNPs, a large group of nucleolar ribonucleoproteins characterized by catalysing 2'-O-

methylation and pseudouridylation of pre-rRNA molecules, respectively. However, the most abundant C/D box snoRNP, U3 snoRNP, does not participate in methylation, but instead guides the endoribonucleolytic processing of pre-rRNA together with U8 and U13 snoRNP[180]. [19]. Historically, anti-fibrillarin autoantibodies were believed to target a 34-kDa protein within the U3 snoRNP complex, resulting in immunoprecipitation of U3 snoRNA. However, it has been shown that anti-fibrillarin autoantibodies immunoprecipitate additional protein components and RNA molecules, such as U8 snoRNA, U13 snoRNA, and other smaller, unidentified RNA molecules [20,21]. Our study found that samples from patients exhibiting the same nucleolar fibrillar staining pattern on IIF (AC-9) as classical anti-fibrillarin-positive patients immunoprecipitated a wide repertoire of C/D box snoRNA molecules, and not exclusively U3 snoRNA.

In this line, recent reports indicate that C/D box snoRNPs may exhibit various supramolecular structures and consist of several proteins, not necessarily including fibrillarin [19,22]. Thus, previous research has identified autoantibodies targeting C/D box snoRNPs that do not react with fibrillarin or U3 snoRNA [21][353]. In our study, patients with autoantibodies against C/D box snoRNPs exhibited a heterogeneous profile, as shown by the different RNA immunoprecipitation patterns determined by RNA IP and RIP-Seq. Of particular note, these differing patterns were associated with two distinct clinical phenotypes: some patients exhibited a more severe phenotype similar to that reported for classical anti-fibrillarin autoantibodies [23][215], while others had a much milder phenotype. Commercial tests failed to distinguish between these groups, as all patients displayed the same AC-9 IIF pattern and were variably positive for anti-fibrillarin by immunoblotting. However, RNA IP was able to differentiate between the two subsets: patients whose samples did not immunoprecipitate any RNA molecule presented a very mild clinical course, while those showing immunoprecipitation of U3, U8, or U13 snoRNA had a more severe course. In addition, WGCNA revealed that among the three patients with autoantibodies against C/D box snoRNPs using both RIP-Seq and RNA IP, the two patients with a more severe clinical phenotype (SSc_30 and SSc_46) clustered together, while the patient negative by RNA IP and with a milder phenotype (SSc_48) did not. These findings suggest that recognition of different C/D box snoRNP sets may correlate with different clinical outcomes. Therefore, it could be beneficial to use RNA IP to assess SSc patients with an IIF AC-9 pattern to better predict the likelihood of developing severe clinical symptoms.

Regarding the identification of new SSc-associated autoantibodies, we found that sample SSc_44 was RNA IP-positive for anti-U11/U12 RNP and also positive for an anti-H/ACA box

snoRNP autoantibody through RIP-Seq, which showed immunoprecipitation of various H/ACA box snoRNAs and scaRNAs. As there are few reports on autoantibodies against H/ACA box snoRNPs, their clinical implications and specific targets remain unclear [21,24][329, 353]. Our patient had dcSSc and ILD, but as ILD is highly associated with anti-U11/U12 RNP autoantibodies, we could not determine whether anti-H/ACA box snoRNP autoantibodies alone are associated with ILD.

As expected, our RIP-Seq results showed that various U5 snRNA and U4 snRNA molecules, which are part of the major and minor spliceosomes, were immunoprecipitated in samples positive for anti-U11/U12 RNP and U4/U6 snRNP. Unexpectedly, U6 snRNA (major spliceosome) and U6-atac snRNA (minor spliceosome) were predominantly immunoprecipitated by SSc_58, a sample that was not known to recognize any protein of the major or minor spliceosome. Interestingly, this sample also immunoprecipitated 7SK RNA. Both U6 and U6-atac snRNA, and 7SK present a γ-monomethyl phosphate cap at their 5' ends [25,26], indicating that sample SSc_58 could present autoantibodies against this structure. However, 7SK RNA was also immunoprecipitated by samples SSc_14, SSc_19 and SSc_26, which do not immunoprecipitate U6 or U6-atac snRNA molecules, suggesting that these patients could present autoantibodies targeting protein components of the 7SK RNP complex. Additionally, we identified several other novel autoantibody targets, including mitochondrial tRNA synthetases, 7SK snRNP, and the cytoplasmic vault complex. Recently, autoantibodies against the major vault protein (MVP) have been reported in systemic lupus erythematosus and other systemic autoimmune diseases. This suggests that anti-Vault complex autoantibodies may be not specific to SSc, but could be associated with the disease, similar to anti-Ro60 antibodies [27]. In any case, this study is the first to report 7SK RNP, mitochondrial tRNA synthetases and cytoplasmic vault complexes as autoantibody targets in SSc.

The main limitation of this study is that the small number of patients with potential new autoantibodies prevented us from establishing statistically significant associations between these autoantibodies and specific clinical phenotypes. To confirm our findings, a larger patient sample is needed. Given the rarity of the disease and the low frequency of these autoantibodies, multicenter studies may be required to achieve a sufficient number of patients. Moreover, further research is needed to confirm whether SSc patients do indeed present autoantibodies against the candidate ribonucleoproteins we identified, as we have not yet demonstrated direct interactions between these autoantibodies and the protein components of the ribonucleoproteins.

CONCLUSION

When the overall SSc cohort was analyzed using commercial assays and the gold-standard RNA IP assay, 16.3% of patients remained negative for all SSc-related autoantibodies. This indicates that there is still a significant subset of SSc patients in which we are not able to detect any specific autoantibody. Since autoantibodies have proven to be excellent prognostic biomarkers in this heterogeneous disease, it is crucial to identify potential novel autoantibodies in patients currently classified as negative. In this study, we used a new technique for this purpose that combines RNA immunoprecipitation with massive parallel sequencing. This approach enabled the detection of RNA molecules known to be immunoprecipitated by SSc-related autoantibodies, confirming the assay's capability to detect autoantibodies targeting ribonucleoprotein complexes. Additionally, we identified RNA molecules that form part of ribonucleoproteins not previously known to be targets of SSc patient sera, reporting possible new autoantibody candidates in this disease.

ACKNOWLEDGMENTS

The authors are grateful to Professor Vicent Fonollosa (Vall d'Hebron University Hospital) for his commitment to the establishment of a systemic sclerosis research line at our hospital.

REFERENCES

1. Denton, C. P. & Khanna, D. Systemic sclerosis. Lancet 390, 1685–1699 (2017). DOI: 10.1016/S0140-6736(17)30933-9.
2. Masi, A. T. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Arthritis Rheum. 23, 581–590 (1980). DOI: 10.1002/art.1780230510.
3. Shah, S. & Denton, C. P. Scleroderma autoantibodies in guiding monitoring and treatment decisions. Curr. Opin. Rheumatol. 34, 302–310 (2022). DOI: 10.1097/BOR.0000000000000904.
4. Mahler, M. et al. Autoantibodies to the Rpp25 component of the Th/To complex are the most common antibodies in patients with systemic sclerosis without antibodies detectable by widely available commercial tests. J. Rheumatol. 41, 1334–1343 (2014). DOI: 10.3899/jrheum.131450.
5. Bonroy, C. et al. Optimization and diagnostic performance of a single multiparameter lineblot in the serological workup of systemic sclerosis. J. Immunol. Methods 379, 53–60 (2012). DOI: 10.1016/j.jim.2012.03.001.
6. Van Praet, J. T. et al. Specific anti-nuclear antibodies in systemic sclerosis patients with and without skin involvement: an extended methodological approach. Rheumatology (Oxford). 50, 1302–1309 (2011). DOI: 10.1093/rheumatology/keq446.
7. Gerovac, M., Vogel, J. & Smirnov, A. The World of Stable Ribonucleoproteins and Its Mapping With Grad-Seq and Related Approaches. Front. Mol. Biosci. 8, 661448 (2021). DOI: 10.3389/fmolb.2021.661448.
8. Perurena-Prieto, J. et al. Anti-nuclear valosin-containing protein-like autoantibody is associated with calcinosis and higher risk of cancer in systemic sclerosis. Rheumatology (Oxford). Online, (2023). DOI: 10.1093/rheumatology/kead520.
9. Kuwana, M., Kaburaki, J., Mimori, T., Tojo, T. & Homma, M. Autoantibody reactive with three classes of RNA polymerases in sera from patients with systemic sclerosis. J. Clin. Invest. 91, 1399–1404 (1993). DOI: 10.1172/JCI116343.
10. Montzka, K. A. & Steitz, J. A. Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. Proc. Natl. Acad. Sci. U. S. A. 85, 8885–8889 (1988). DOI: 10.1073/pnas.85.23.8885.
11. Moore, K. S. & 't Hoen, P. A. C. Computational approaches for the analysis of RNA-protein interactions: A primer for biologists. J. Biol. Chem. 294, 1–9 (2019). DOI: 10.1074/jbc.REV118.004842.
12. LeRoy, E. C. & Medsger, T. A. J. Criteria for the classification of early systemic sclerosis. J. Rheumatol. 28, 1573–1576 (2001).
13. Van Den Hoogen, F. et al. 2013 classification criteria for systemic sclerosis: An american college of rheumatology/European league against rheumatism collaborative initiative. Arthritis Rheum. 65, 2737–2747 (2013). DOI: 10.1002/art.38098.
14. Hamaguchi, Y., Kuwana, M. & Takehara, K. Performance evaluation of a commercial line blot assay system for detection of myositis- and systemic sclerosis-related autoantibodies. Clin. Rheumatol. 39, 3489–3497 (2020). DOI: 10.1007/s10067-020-04973-0.
15. Fertig, N. et al. Anti-U11/U12 RNP antibodies in systemic sclerosis: A new serologic marker associated with pulmonary fibrosis. Arthritis Care Res. 61, 958–965 (2009). DOI: 10.1002/art.24586.
16. Callejas-Moraga, E. L. et al. Anti-RNPC-3 antibody predicts poor prognosis in patients with interstitial lung disease associated to systemic sclerosis. Rheumatol. (United Kingdom) 61, 154–162 (2022). DOI: 10.1093/rheumatology/keab279.

17. Gilliam, A. C. & Steitz, J. A. Rare scleroderma autoantibodies to the U11 small nuclear ribonucleoprotein and to the trimethylguanosine cap of U small nuclear RNAs. Proc. Natl. Acad. Sci. U. S. A. 90, 6781–6785 (1993). DOI: 10.1073/pnas.90.14.6781.
18. Verma, B., Akinyi, M. V, Norppa, A. J. & Frilander, M. J. Minor spliceosome and disease. Semin. Cell Dev. Biol. 79, 103–112 (2018). DOI: 10.1016/j.semcdb.2017.09.036.
19. Ojha, S., Malla, S. & Lyons, S. M. snoRNPs: Functions in Ribosome Biogenesis. Biomolecules 10, 783 (2020). DOI: 10.3390/biom10050783.
20. Tyc, K. & Steitz, J. A. U3, U8 and U13 comprise a new class of mammalian snRNPs localized in the cell nucleolus. EMBO J. 8, 3113–3119 (1989). DOI: 10.1002/j.1460-2075.1989.tb08463.x.
21. Van Eenennaam, H. et al. Autoantibodies against small nucleolar ribonucleoprotein complexes and their clinical associations. Clin. Exp. Immunol. 130, 532–540 (2002). DOI: 10.1046/j.1365-2249.2002.01991.x.
22. Stamm, S. & Lodmell, J. S. C/D box snoRNAs in viral infections: RNA viruses use old dogs for new tricks. Non-coding RNA Res. 4, 46–53 (2019). DOI: 10.1016/j.ncrna.2019.02.001.
23. Tormey, V. J., Bunn, C. C., Denton, C. P. & Black, C. M. Anti-fibrillarin antibodies in systemic sclerosis. Rheumatology 40, 1157–1162 (2001). DOI: 10.1093/rheumatology/40.10.1157.
24. Vulsteke, J.-B. et al. Identification of new telomere- and telomerase-associated autoantigens in systemic sclerosis. J. Autoimmun. 135, 102988 (2023). DOI: 10.1016/j.jaut.2022.102988.
25. Gupta, S., Busch, R. K., Singh, R. & Reddy, R. Characterization of U6 small nuclear RNA cap-specific antibodies. Identification of gamma-monomethyl-GTP cap structure in 7SK and several other human small RNAs. J. Biol. Chem. 265, 19137–19142 (1990).
26. Tarn, W.-Y. & Steitz, J. A. Highly Diverged U4 and U6 Small Nuclear RNAs Required for Splicing Rare AT-AC Introns. Science (80-.). 273, 1824–1832 (1996). DOI: 10.1126/science.273.5283.1824.
27. Moadab, F. et al. Argonaute, Vault, and Ribosomal Proteins Targeted by Autoantibodies in Systemic Lupus Erythematosus. J. Rheumatol. 50, 1136–1144 (2023). DOI: 10.3899/jrheum.2022-1327.

Table 1: Candidates selected after protein-coding RNA molecule re-analysis based on max^2/median ratio (cut-off ≥21,778,93). RNA molecules immunoprecipitated by known SSc-related autoantibodies using traditional RNA IP are marked in blue.

RNA molecule name	ENSEMBL ID	RNA molecule type	Max^2/median	Known immunoprecipitated RNA molecule
SNORD3B-1	**ENSG00000265185.6**	**C/D box snoRNA**	**4.44E+09**	**U3 RNA**
RNY4	**ENSG00000252316.1**	**Y RNA**	**1.34E+09**	**Y4 RNA**
RNU11	**ENSG00000274978.1**	**snRNA**	**7.03E+08**	**U11 RNA**
VTRNA1-2	ENSG00000202111.1	Vault RNA	1.49E+08	-
RNY1	**ENSG00000201098.1**	**Y RNA**	**1.43E+08**	**Y1 RNA**
RNY5	**ENSG00000286171.1**	**Y RNA**	**7.62E+07**	**Y5 RNA**
SNORD14C	ENSG00000202252.1	C/D box snoRNA	5.51E+07	-
RMRP	**ENSG00000269900.3**	**lncRNA**	**3.80E+07**	**7-2 RNA**
VTRNA2-1	ENSG00000270123.4	Vault RNA	3.57E+07	-
SNORD49A	ENSG00000277370.1	C/D box snoRNA	1.84E+07	-
SNORD30	ENSG00000277846.1	C/D box snoRNA	1.46E+07	-
SNHG5	ENSG00000203875.13	SNRHG	1.41E+07	-
SNORD45A	ENSG00000207241.1	C/D box snoRNA	1.07E+07	-
SNORD68	ENSG00000200084	C/D box snoRNA	1.03E+07	-
GAS5	ENSG00000234741.8	SNRHG	1.02E+07	-
SNORD46	ENSG00000200913.1	C/D box snoRNA	7.92E+06	-
SNORD100	ENSG00000221500.1	C/D box snoRNA	7.90E+06	-
SNORD63	ENSG00000206989.1	C/D box snoRNA	7.72E+06	-
RNY3	**ENSG00000202354.1**	**Y RNA**	**7.69E+06**	**Y3 RNA**
CSKMT	ENSG00000199352.1	rRNA	6.31E+06	-
SNORD2	ENSG00000238942.1	C/D box snoRNA	4.78E+06	-
7SK	ENSG00000202198.1	miscRNA	4.70E+06	-
VTRNA1-1	ENSG00000199990.1	Vault RNA	4.34E+06	-
SNORD21	ENSG00000206680.1	C/D box snoRNA	4.02E+06	-
SNHG29	ENSG00000175061.18	SNRHG	3.36E+06	-
SNORD57	ENSG00000226572.1	C/D box snoRNA	3.10E+06	-
SNORD101	ENSG00000206754.1	C/D box snoRNA	2.87E+06	-
SNORD12	ENSG00000212304.1	C/D box snoRNA	2.72E+06	-
SNORD27	ENSG00000275996.1	C/D box snoRNA	2.69E+06	-
SNHG32	ENSG00000204387.14	SNRHG	2.64E+06	-

Table 1 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
RNA5SP376	ENSG00000212251.1	rRNA	2.57E+06	-
SNORD114-1	ENSG00000199575.1	C/D box snoRNA	2.46E+06	-
RPPH1	**ENSG00000259001.3**	**lncRNA**	**2.08E+06**	**8-2 RNA**
SNORD18C	ENSG00000199574.1	C/D box snoRNA	1.92E+06	-
SNORD91B	ENSG00000275084.4	C/D box snoRNA	1.78E+06	-
SNORD18A	ENSG00000200623.1	C/D box snoRNA	1.62E+06	-
SNORD25	ENSG00000275043.1	C/D box snoRNA	1.55E+06	-
SNORD99	ENSG00000221539.1	C/D box snoRNA	1.38E+06	-
SNORD92	ENSG00000264994.1	C/D box snoRNA	1.37E+06	-
SNORD12B	ENSG00000222365.1	C/D box snoRNA	1.23E+06	-
SNORD3A	ENSG00000263934.5	C/D box snoRNA	1.22E+06	-
ENSG10010134635.1	ENSG00000222185.1	C/D box snoRNA	1.17E+06	-
SNORD63B	ENSG00000222937.1	C/D box snoRNA	1.15E+06	-
SNORD111B	ENSG00000221514.1	C/D box snoRNA	1.00E+06	-
SNORD45C	ENSG00000206620.1	C/D box snoRNA	9.46E+05	-
SNORD114-12	ENSG00000202270.1	C/D box snoRNA	9.37E+05	-
SNORD86	ENSG00000212498.1	C/D box snoRNA	9.00E+05	-
SNORD95	ENSG00000264549.1	C/D box snoRNA	8.13E+05	-
SNORD1B	ENSG00000199961.1	C/D box snoRNA	7.45E+05	-
SNORD6	ENSG00000202314.1	C/D box snoRNA	7.40E+05	-
SNORD82	ENSG00000202400.1	C/D box snoRNA	6.55E+05	-
SNORD43	ENSG00000263764.1	C/D box snoRNA	6.22E+05	-
SNORD58A	ENSG00000206602.1	C/D box snoRNA	5.95E+05	-
MALAT1	ENSG00000251562.8	lncRNA	5.95E+05	-
SNORD24	ENSG00000206611.1	C/D box snoRNA	5.93E+05	-
SNORD114-23	ENSG00000200406.1	C/D box snoRNA	5.76E+05	-
SNORA3B	ENSG00000212607.1	H/ACA box snoRNA	5.72E+05	-
SNORD104	ENSG00000199753.1	C/D box snoRNA	5.58E+05	-
RNU5A-1	ENSG00000199568.1	snRNA	5.47E+05	U5 RNA
RNA5SP289	ENSG00000199202.1	rRNA	5.42E+05	-
SNORD110	ENSG00000221116.1	C/D box snoRNA	5.13E+05	-
SNORD60	ENSG00000206630.1	C/D box snoRNA	4.61E+05	-

Table 1 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
SNORD114-22	ENSG00000202293.1	C/D box snoRNA	4.58E+05	-
SNORD102	ENSG00000207500.1	C/D box snoRNA	4.07E+05	-
SNORD52	ENSG00000201754.1	C/D box snoRNA	4.04E+05	-
RNU5B-1	ENSG00000200156.1	snRNA	3.99E+05	U5 RNA
SNORD114-25	ENSG00000200612.1	C/D box snoRNA	3.98E+05	-
SNORA16A	ENSG00000280498.1	H/ACA box snoRNA	3.95E+05	-
SNORD114-26	ENSG00000200413.1	C/D box snoRNA	3.92E+05	-
SNORD114-3	ENSG00000201839.1	C/D box snoRNA	3.84E+05	-
MT-TE	ENSG00000210194.1	Mitochondrial tRNA	3.79E+05	-
SNORD126	ENSG00000238344.1	C/D box snoRNA	3.74E+05	-
RNU5F-1	ENSG00000199377.1	snRNA	3.54E+05	U5 RNA
SNORD33	ENSG00000199631.1	C/D box snoRNA	3.42E+05	-
SNORD37	ENSG00000206775.1	C/D box snoRNA	3.35E+05	-
SNORD113-8	ENSG00000200367.1	C/D box snoRNA	3.21E+05	-
SNORD127	ENSG00000239043.1	C/D box snoRNA	3.17E+05	-
SNORD69	ENSG00000212452.1	C/D box snoRNA	3.00E+05	-
SNORD111	ENSG00000221066.1	C/D box snoRNA	2.97E+05	-
SNORD105B	ENSG00000238531.1	C/D box snoRNA	2.81E+05	-
SNORD42A	ENSG00000238649.1	C/D box snoRNA	2.73E+05	-
G3BP1	ENSG00000145907.16	Protein coding	2.70E+05	-
RNU6-5P	ENSG00000206965.1	snRNA	2.45E+05	U6 RNA
ENSG00000280494	ENSG00000280494.2	miRNA	2.43E+05	-
SNORD12C	ENSG00000209042.1	C/D box snoRNA	2.25E+05	-
SNORA61	ENSG00000278274.1	H/ACA box snoRNA	2.24E+05	-
SNORA20	ENSG00000207392.1	H/ACA box snoRNA	2.23E+05	-
SNORD34	ENSG00000202503.1	C/D box snoRNA	2.08E+05	-
SNORD58C	ENSG00000202093.1	C/D box snoRNA	1.89E+05	-
7SK	ENSG00000271394.1	miscRNA	1.78E+05	-
SNORD65	ENSG00000277512.1	C/D box snoRNA	1.77E+05	-
SNORD72	ENSG00000212296.1	C/D box snoRNA	1.74E+05	-
SNORD28	ENSG00000274544.1	C/D box snoRNA	1.55E+05	-
SNORA3A	ENSG00000200983.1	H/ACA box snoRNA	1.54E+05	-

Table 1 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
SNORD66	ENSG00000212158.1	C/D box snoRNA	1.50E+05	-
SNORA18	ENSG00000207145.1	H/ACA box snoRNA	1.42E+05	-
SNORD83A	ENSG00000209482.1	C/D box snoRNA	1.42E+05	-
SNORD61	ENSG00000206979.1	C/D box snoRNA	1.35E+05	-
RNU12	ENSG00000276027.1	snRNA	1.35E+05	U12 RNA
SNORD114-11	ENSG00000200608.1	C/D box snoRNA	1.35E+05	-
SNORD49B	ENSG00000277108.1	C/D box snoRNA	1.29E+05	-
SNORD19C	ENSG00000222345.1	C/D box snoRNA	1.18E+05	-
SNORD41	ENSG00000209702.1	C/D box snoRNA	1.18E+05	-
SNORD53B	ENSG00000265706.1	C/D box snoRNA	1.16E+05	-
SNORD71	ENSG00000223224.1	C/D box snoRNA	1.15E+05	-
SNORA2C	ENSG00000221491.2	H/ACA box snoRNA	1.12E+05	-
SNORD114-14	ENSG00000199593.1	C/D box snoRNA	1.07E+05	-
SNORD38A	ENSG00000202031.1	C/D box snoRNA	1.07E+05	-
SNORD54	ENSG00000238650.1	C/D box snoRNA	1.07E+05	-
SNORA66	ENSG00000207523.1	H/ACA box snoRNA	1.06E+05	-
SNORD51	ENSG00000207047.2	C/D box snoRNA	1.05E+05	-
SCARNA22	ENSG00000249784.1	scaRNA	1.01E+05	-
SNORD19B	ENSG00000238862.1	C/D box snoRNA	9.98E+04	-
SNORD4A	ENSG00000238578.1	H/ACA box snoRNA	9.83E+04	-
SNORD59A	ENSG00000207031.1	C/D box snoRNA	9.81E+04	-
SNORD105	ENSG00000209645.1	C/D box snoRNA	9.71E+04	-
SNORD114-9	ENSG00000201240.1	C/D box snoRNA	9.53E+04	-
SNORD114-17	ENSG00000201569.1	C/D box snoRNA	9.28E+04	-
TTC3	ENSG00000182670.13	Protein coding	8.87E+04	-
SNORD50B	ENSG00000275072.1	C/D box snoRNA	8.84E+04	-
ENSG10010137917.1	ENSG00000280554.1	C/D box snoRNA	8.74E+04	-
SNORD5	ENSG00000239195.1	C/D box snoRNA	8.43E+04	-
MT-CO1	ENSG00000198804.2	Protein coding	8.40E+04	-
SNORD3B-2	ENSG00000262074.7	C/D box snoRNA	8.36E+04	-
SNORD36C	ENSG00000252542.1	C/D box snoRNA	8.28E+04	-
SNORD113-6	ENSG00000200215.3	C/D box snoRNA	7.96E+04	-

Table 1 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
CH17-3B23.3	ENSG00000287979.1	lncRNA	7.84E+04	-
SNORD91A	ENSG00000212163.6	C/D box snoRNA	7.67E+04	-
SNORD38B	ENSG00000281859.1	C/D box snoRNA	7.36E+04	-
SNORD87	ENSG00000254341.2	C/D box snoRNA	7.19E+04	-
MT-TV	ENSG00000210077.1	Mitochondrial tRNA	7.02E+04	-
RNU6ATAC	ENSG00000221676.1	snRNA	7.00E+04	U6 RNA
SNORD1A	ENSG00000278261.1	C/D box snoRNA	6.54E+04	-
SNORD84	ENSG00000265236.1	C/D box snoRNA	6.23E+04	-
SNORD114-13	ENSG00000201247.1	C/D box snoRNA	6.18E+04	-
MT-TM	ENSG00000210112.1	Mitochondrial tRNA	6.05E+04	-
SNORD11B	ENSG00000271852.1	C/D box snoRNA	5.97E+04	-
RP11-596C23.6	ENSG00000282885.2	lncRNA	5.93E+04	-
SNORD14B	ENSG00000201403.1	C/D box snoRNA	5.86E+04	-
SNORD45B	ENSG00000201487.1	C/D box snoRNA	5.86E+04	-
SCARNA15	ENSG00000277864.1	scaRNA	5.76E+04	-
FBXL20	ENSG00000108306.13	Protein coding	5.71E+04	-
SNORD7	ENSG00000207297.1	C/D box snoRNA	5.49E+04	-
SNORD114-28	ENSG00000200480.1	C/D box snoRNA	5.39E+04	-
SNORD113-7	ENSG00000200632.1	C/D box snoRNA	5.31E+04	-
LINC02739	ENSG00000255008.3	lncRNA	5.28E+04	-
RANBP2	ENSG00000153201.16	Protein coding	5.24E+04	-
RN7SL1	ENSG00000276168.1	miscRNA	5.15E+04	-
PPP1CB	ENSG00000213639.10	Protein coding	5.04E+04	-
SNORD35A	ENSG00000200259.1	C/D box snoRNA	4.88E+04	-
SNORD94	ENSG00000208772.1	C/D box snoRNA	4.87E+04	-
SNORD70	ENSG00000212534.1	C/D box snoRNA	4.79E+04	-
SNORD1C	ENSG00000274091.1	C/D box snoRNA	4.66E+04	-
SNORA58B	ENSG00000201129.1	H/ACA box snoRNA	4.66E+04	-
SNORA77B	ENSG00000264346	H/ACA box snoRNA	4.65E+04	-
SNORA46	ENSG00000207493.1	H/ACA box snoRNA	4.51E+04	-
SNORD73A	ENSG00000208797.1	C/D box snoRNA	4.10E+04	-
KMT2A	ENSG00000118058.23	Protein coding	4.08E+04	-

Table 1 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
SNORD14D	ENSG00000207118.1	C/D box snoRNA	4.03E+04	-
SNORD98	ENSG00000283551.1	C/D box snoRNA	4.01E+04	-
TUFM	ENSG00000178952.11	Protein coding	4.01E+04	-
AK2	ENSG00000004455.17	Protein coding	3.99E+04	-
SNORD4B	ENSG00000238597.1	C/D box snoRNA	3.92E+04	-
SNORD90	ENSG00000212447.1	C/D box snoRNA	3.86E+04	-
EP300	ENSG00000100393.14	Protein coding	3.82E+04	-
ATRX	ENSG00000085224.23	Protein coding	3.81E+04	-
GAPDH	ENSG00000111640.15	Protein coding	3.68E+04	-
AC084082.3	ENSG00000253190.4	lncRNA	3.59E+04	-
RPS18	ENSG00000231500.7	Protein coding	3.57E+04	-
SNORD83B	ENSG00000209480.1	C/D box snoRNA	3.30E+04	-
SNORD16	ENSG00000199673.1	C/D box snoRNA	3.22E+04	-
ZMYND8	ENSG00000101040.19	Protein coding	3.21E+04	-
SNORA57	ENSG00000206597	H/ACA box snoRNA	3.12E+04	-
IST1	ENSG00000182149.21	Protein coding	3.08E+04	-
SNORD57	ENSG00000226572	Protein coding	3.08E+04	-
SNORD93	ENSG00000221740.1	C/D box snoRNA	3.02E+04	-
ENSG10010138968.2	ENSG00000252787.2	C/D box snoRNA	2.91E+04	-
SNORD114-10	ENSG00000200279.1	C/D box snoRNA	2.85E+04	-
TPT1	ENSG00000133112.17	Protein coding	2.81E+04	-
SNORD89	ENSG00000212283.1	C/D box snoRNA	2.76E+04	-
SNORD62A	ENSG00000235284.1	C/D box snoRNA	2.68E+04	-
SNORD11	ENSG00000238317.2	C/D box snoRNA	2.66E+04	-
SNHG17	ENSG00000196756.13	SNRHG	2.59E+04	-
MT-TH	ENSG00000210176.1	Mitochondrial tRNA	2.59E+04	-
RNA5S9	ENSG00000201321.1	rRNA	2.53E+04	-
SNORD113-5	ENSG00000272474.1	C/D box snoRNA	2.52E+04	-
LPIN1	ENSG00000134324.12	Protein coding	2.51E+04	-
RNU4-1	ENSG00000200795.1	snRNA	2.48E+04	U4 RNA
SNORD118	ENSG00000200463.1	C/D box snoRNA	2.44E+04	-
GAN	ENSG00000261609.8	Protein coding	2.40E+04	-

Table 1 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
SNORD14A	ENSG00000272034.1	C/D box snoRNA	2.39E+04	-
ENSG00000253389*	ENSG00000253389*	Not identified RNA sequence	2.28E+04	-
SNORD17	ENSG00000212232.1	C/D box snoRNA	2.28E+04	-
H2AC18	ENSG00000203812.2	Protein coding	2.23E+04	-
PI4KAP2	ENSG00000183506.17	Transcribed unitary pseudogene	2.21E+04	-
SNORD14E	ENSG00000200879.1	C/D box snoRNA	2.18E+04	-
NUP214	ENSG00000126883.17	Protein coding	2.18E+04	-

lncRNA, long non-coding RNA; miRNA, micro RNA; miscRNA, miscellaneous RNA; rRNA, ribosomal RNA; SNRHG, small nucleolar RNA host gene; snRNA, small nuclear RNA; snoRNA, small nucleolar RNA; tRNA, transfer RNA. *Reads were mapped to a specific intronic sequence of ENSG00000253389 (CAGCACCAGCGTGACCAACAGCTGGGTGACGAAAGTCCACATCCTCGCCT)

Table 2: Demographic and clinical characteristics, immunoblotting and RNA IP results in the 12 patients with an AC-9 pattern included in the study. Patients were classified into 3 groups: anti-U3 snoRNP positive, anti-U3, -U8 or -U13 snoRNP positive, or RNA IP negative.

Patient	SSc_30	SSc_78	SSc_85	SSc_46	SSc_48	SSc_70	SSc_79	SSc_80	SSc_81	SSc_82	SSc_83	SSc_84	SSc_86	Anti-U3 snoRNP positive (%)	Anti-U3/U8/U13 snoRNP positive (%)	RNA IP negative (%)
Commercial IB result	N	3+	2+	N	N	(+)	2+	1+	(+)	(+)	3+	1+	(+)	-	-	-
RNA IP result (U3 snoRNA band)	P	P	P	N	N	N	N	N	N	N	N	N	N	-	-	-
RNA IP result (U8/U13 snoRNA bands)	N	P	P	P	N	N	P	N	P	P	P	N	P	-	-	-
Gender	F	F	M	F	F	M	F	F	M	F	F	F	F	-	-	-
SSc cutaneous subset	lcSSc	lcSSc	dcSSc	lcSSc	lcSSc	dcSSc	dcSSc	ssSSc	lcSSc	dcSSc	dcSSc	ssSSc	lcSSc	-	-	-
Age at disease onset, y	48.3	56.7	55.4	12.0	16.0	26.1	36.2	34.8	43.3	55.2	60.5	17.3	11.4	53	42.1	23.53
Age at onset of first non-RP symptom, y	51.8	56.7	55.4	37.6	50.3	26.3	36.2	37.2	44.3	55.7	60.5	18.3	19.4	55	46.4	33.01
2013 ACR/EULAR criteria	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	67	89	50
Cutolo late pattern	No	No	Yes	Yes	Yes	No	Yes	No	No	Yes	Yes	No	No	33	56	25
Peripheral vascular manifestations	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	100	100	100
RP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	100	100	100
DU	No	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	No	No	0	44	25
Telangiectasias	Yes	No	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	67	89	50
Gastrointestinal involvement	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	67	78	50
Oesophageal	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	67	78	75
Gastric	No	No	Yes	No	No	No	No	No	No	No	No	No	No	33	11	0
Intestinal	No	No	No	No	No	No	Yes	No	No	Yes	Yes	No	No	0	33	0
Lung involvement	Yes	No	No	Yes	No	No	No	No	No	No	Yes	No	No	33	33	0
ILD	Yes	No	No	No	No	No	No	No	No	No	Yes	No	No	33	22	0
PAH	Yes	No	No	Yes	No	No	No	No	No	Yes	No	No	No	33	33	0
Scleroderma renal crisis	No	No	No	No	No	No	No	No	No	No	No	No	No	0	0	0

Table 2 (continued)

Patient	SSc_30	SSc_78	SSc_85	SSc_46	SSc_48	SSc_70	SSc_79	SSc_80	SSc_81	SSc_82	SSc_83	SSc_84	SSc_86	Anti-U3 snoRNP positive (%)	Anti-U3/U8/U13 snoRNP positive (%)	RNA IP negative (%)
Heart involvement	Yes	No	No	Yes	No	No	No	No	Yes	Yes	Yes	Yes	Yes	33	67	25
Ischemic heart disease	Yes	No	No	No	No	No	No	No	No	Yes	Yes	No	No	33	33	0
Pericardial involvement	No	No	No	Yes	No	No	No	No	No	Yes	No	Yes	No	0	22	25
Myocardial fibrosis	No	No	No	No	No	No	Yes	No	No	Yes	Yes	No	No	0	33	0
Conduction abnormalities	Yes	No	No	No	No	No	No	No	Yes	Yes	No	No	No	33	33	0
Diastolic dysfunction	Yes	No	No	Yes	No	No	No	No	No	Yes	Yes	No	No	33	44	0
LVEF <50%	No	No	No	Yes	No	No	No	No	No	Yes	No	No	No	0	22	0
Musculoskeletal involvement	No	No	Yes	No	Yes	No	Yes	No	No	Yes	Yes	No	No	33	44	25
Arthralgia	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	100	89	100
Arthritis	No	No	Yes	No	No	No	Yes	No	No	No	No	No	No	33	22	0
Tendon friction rubs	No	No	No	No	No	No	Yes	No	No	Yes	Yes	No	No	0	33	0
Contractures	No	No	Yes	No	Yes	No	Yes	No	No	Yes	Yes	No	No	33	44	25
Myositis	No	No	No	No	No	No	Yes	No	No	Yes	No	No	No	0	22	0
Calcinosis	No	No	No	No	No	No	No	No	No	No	No	No	No	0	0	0
Cancer	No	No	No	No	No	No	No	No	No	No	No	No	No	0	0	0
Death	No	No	No	No	No	No	No	No	No	Yes	No	No	No	0	11	0

DU, digital ulcer; dcSSc, diffuse cutaneous systemic sclerosis; IB, immunoblot; ILD, interstitial lung disease; lcSSc, limited cutaneous systemic sclerosis; LVEF, left ventricular ejection fraction; N, negative; PAH, pulmonary arterial hypertension; P, positive; RNA IP; RNA immunoprecipitation; RP, Raynaud's phenomenon; ssSSc, sine scleroderma; y, years.

FIGURE LEGENDS

Figure 1. Graphical summary of the RNA IP and RIP-Seq protocols. Protein-A Sepharose beads were coupled with autoantibodies present in patient sera and incubated with ribonucleoprotein extracts obtained from K562 cells. RNA was isolated by classical phenol:chloroform:isoamyl alcohol extraction. In traditional RNA IP, immunoprecipitated RNA was analyzed by TBE-Urea-PAGE. For RIP-Seq, RNA was purified, libraries were prepared, and size selection was performed before massive parallel sequencing. Reads were mapped and counted using the GRCh38 genome and GENCODE v38 annotation as a reference.

Figure 2: RNA IP results and IIF patterns of anti-Th/To positive samples. (A) RNA IP results for anti-Th/To- and anti-fibrillarin-positive samples. Samples SSc_01, SSc_42, SSc_51, SSc_53, SSc_58 and SSc_71 immunoprecipitated 7-2 and 8-2 RNA molecules, associated with anti-Th/To autoantibodies, while sample SSc_30 immunoprecipitated U3 RNA, associated with anti-fibrillarin autoantibodies. (B) Homogeneous nucleolar IIF pattern (AC-8) observed on HEp-2 cells in samples SSc_01 (a), SSc_42 (b), SSc_51 (c), SSc_53 (d) and SSc_58 (e) and SSc_71 (f). (C) RNA IP results for anti-U11/U12 RNP-positive samples (SSc_04, SSc_12, SSc_29, SSc_33, SSc_39, SSc_41, SSc_44, SSc_47 and SSc_66).

Figure 3: RIP-Seq results for RNA molecules previously found to be immunoprecipitated by RNA IP. RNA count numbers for molecules immunoprecipitated by RNA IP in each sample. Counts obtained for (A) 7-2 RNA (blue) and 8-2 RNA (orange), (B) U3 RNA, (C) U11 RNA, (D) Y1 RNA (blue), Y3 RNA (orange), (E) Y4 RNA (blue) and Y5 RNA (orange).

Figure 4: RIP-Seq results for RNA molecules not known to be immunoprecipitated by RNA IP. (A) Correlation between individual samples and modules generated by WGCNA. (B) Proportional distribution of C/D box snoRNA, SNHG, H/ACA box snoRNA, and scaRNA molecules across each module. The total number of RNA molecules within each module is provided beneath the module name, and the percentage of C/D box snoRNA plus SNHG RNA molecules is indicated within each module. (C) RNA molecules other than snoRNAs detected using RIP-Seq analysis. A cut-off based on the interquartile range (IQR) was established taking into consideration the results obtained for RNA molecules in samples positive for known SSc-related autoantibodies by RNA IP. Samples were classified as positive (red) or negative (white) for each RNA molecule using the criterion >Q3+5·IQR.

Figure 1

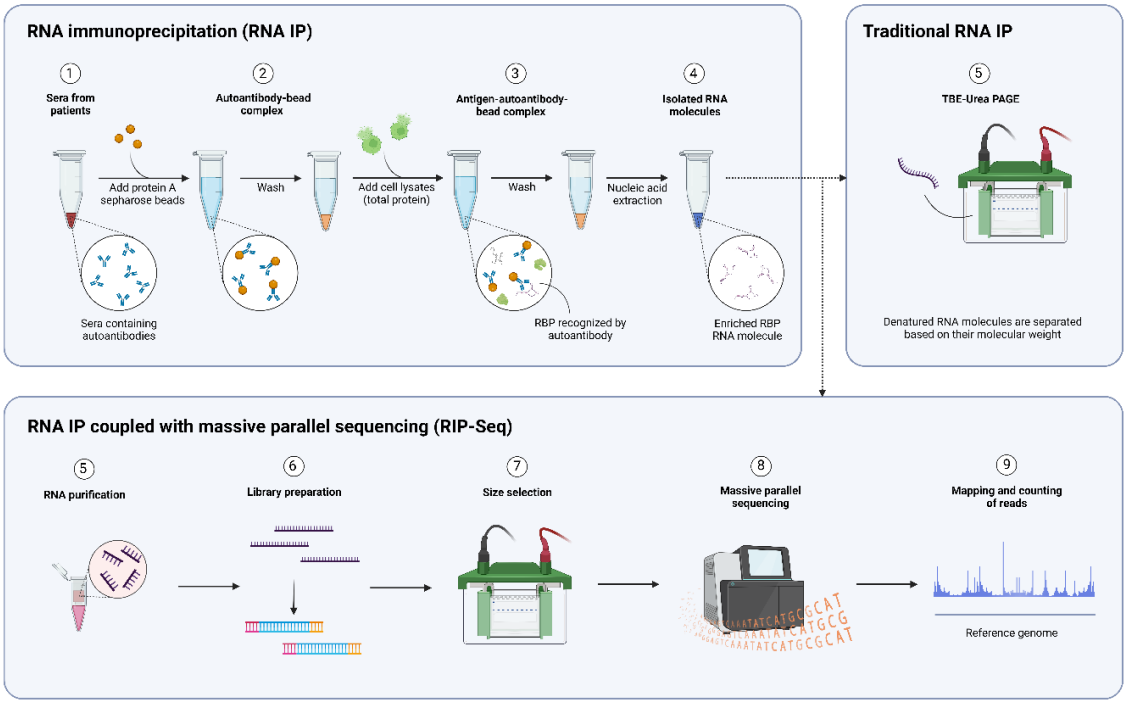

RNA immunoprecipitation (RNA IP)
1
Sera from patients
Add protein A sepharose beads
2
Autoantibody-bead complex
Wash
Add cell lysates (total protein)
3
Antigen-autoantibody-bead complex
Wash
Nucleic acid extraction
4
Isolated RNA molecules
Sera containing autoantibodies
RBP recognized by autoantibody
Enriched RBP RNA molecule
Traditional RNA IP
5
TBE-Urea PAGE
Denatured RNA molecules are separated based on their molecular weight
RNA IP coupled with massive parallel sequencing (RIP-Seq)
5
RNA purification
6
Library preparation
7
Size selection
8
Massive parallel sequencing
9
Mapping and counting of reads
Reference genome

Figure 2

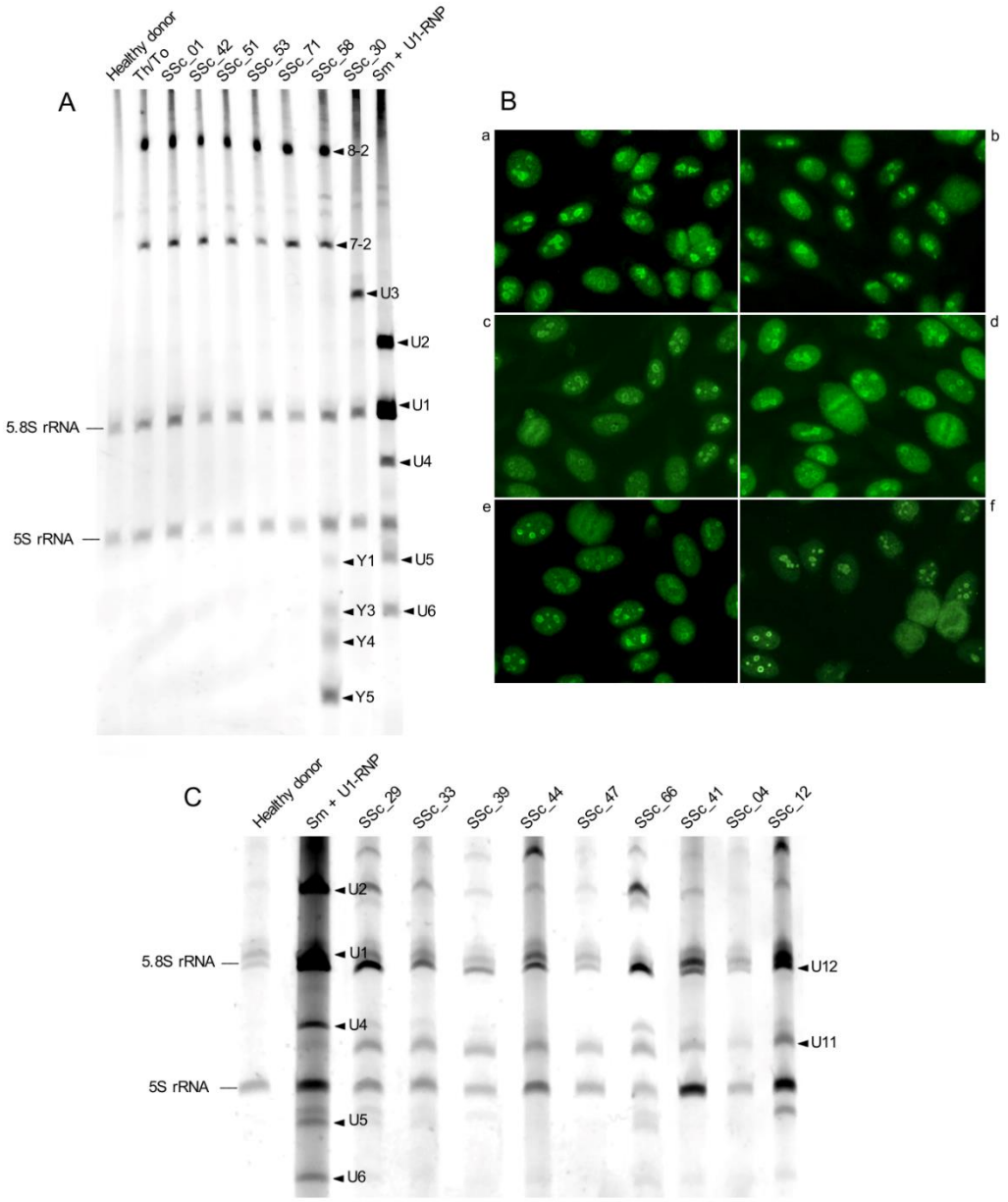

A
Healthy donor
Th/To
SSc_01
SSc_42
SSc_51
SSc_53
SSc_71
SSc_58
SSc_30
Sm + U1-RNP
8-2
7-2
U3
U2
U1
U4
U5
U6
Y1
Y3
Y4
Y5
5.8S rRNA
5S rRNA
B
a
b
c
d
e
f
C
Healthy donor
Sm + U1-RNP
SSc_29
SSc_33
SSc_39
SSc_44
SSc_47
SSc_66
SSc_41
SSc_04
SSc_12
U2
U1
U4
U5
U6
U12
U11
5.8S rRNA
5S rRNA

Figure 3

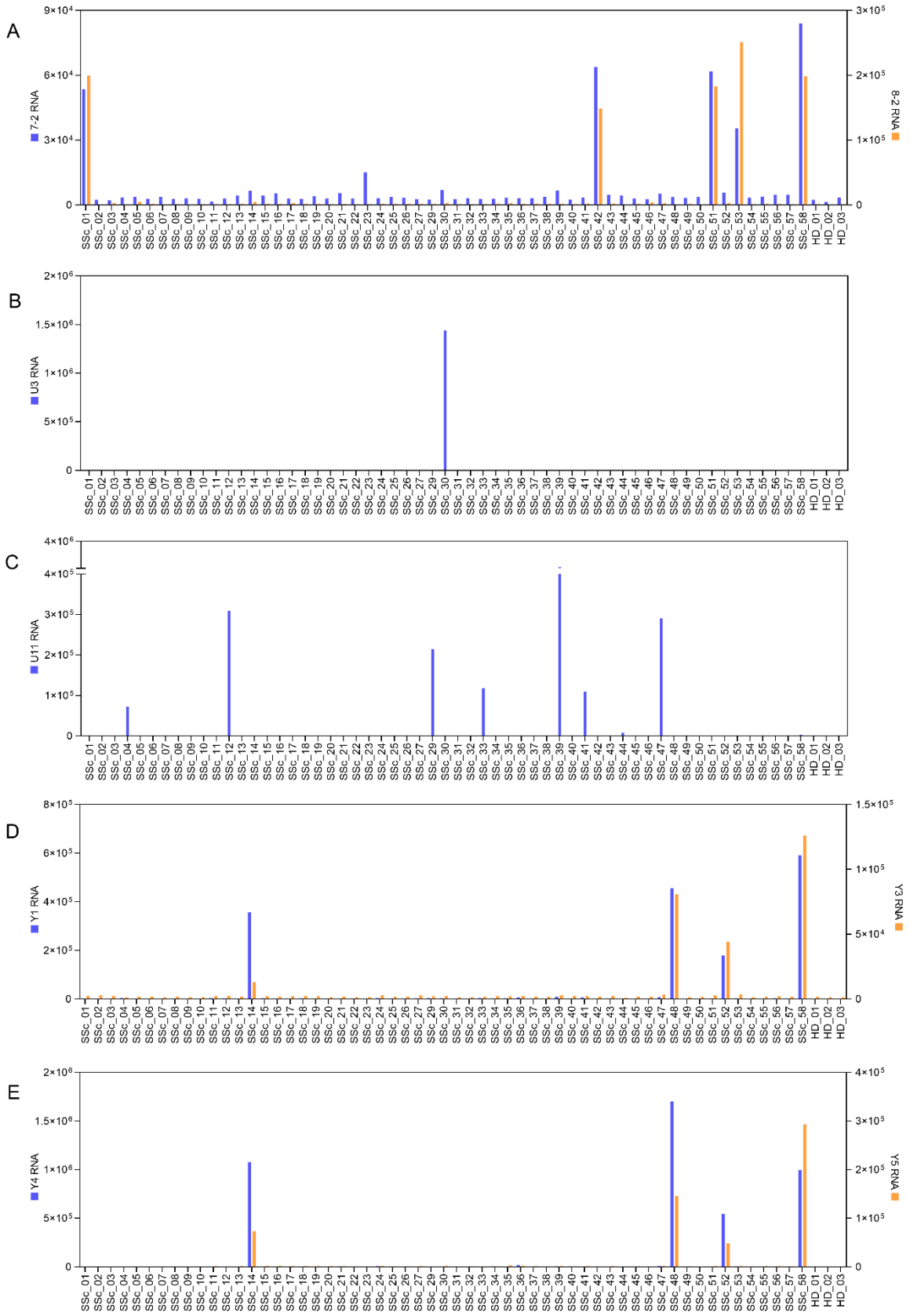
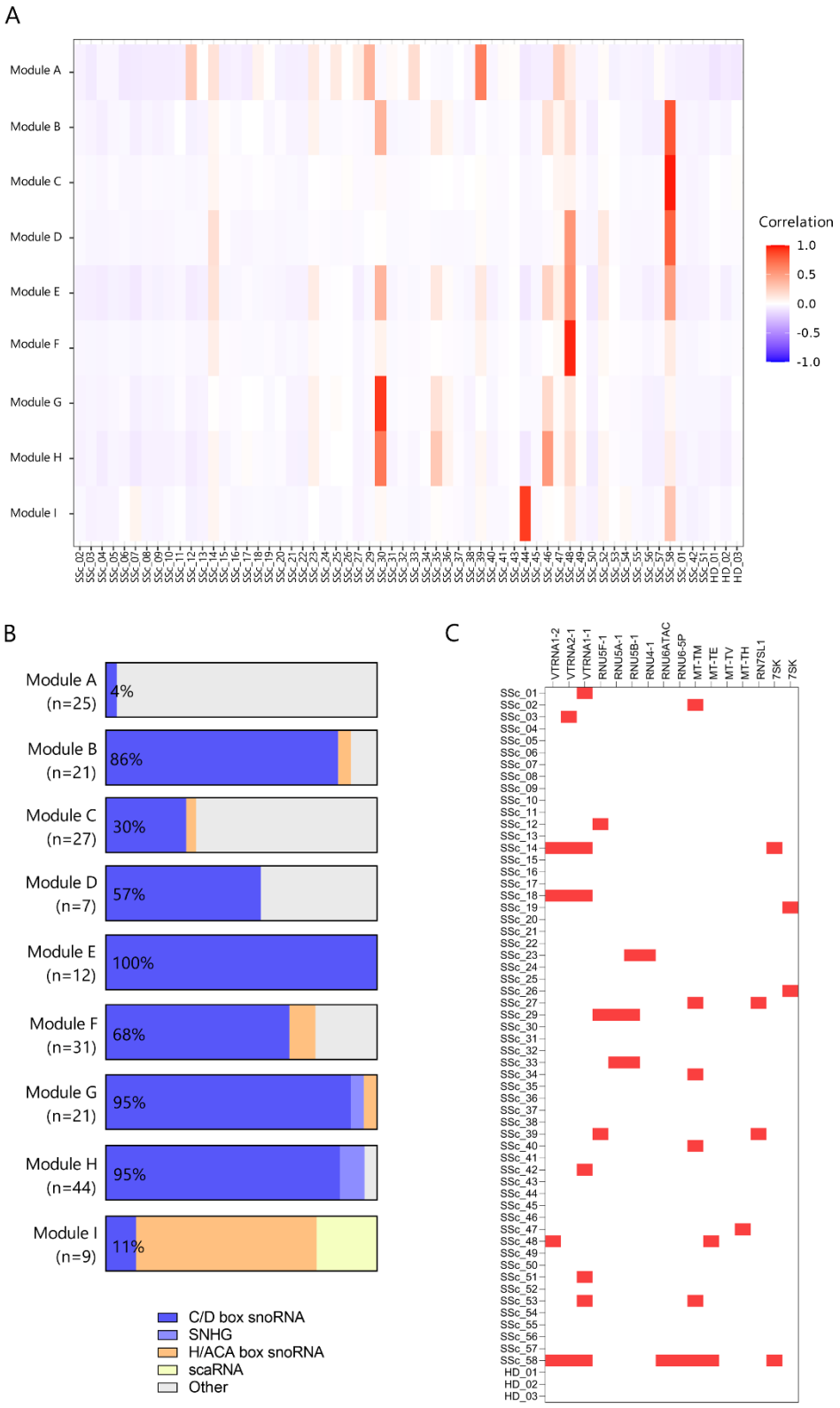


Figure 4


A
Module A
Module B
Module C
Module D
Module E
Module F
Module G
Module H
Module I
Correlation
1.0
0.5
0.0
-0.5
-1.0
B
Module A (n=25)
4%
Module B (n=21)
86%
Module C (n=27)
30%
Module D (n=7)
57%
Module E (n=12)
100%
Module F (n=31)
68%
Module G (n=21)
95%
Module H (n=44)
95%
Module I (n=9)
11%
C/D box snoRNA
SNHG
H/ACA box snoRNA
scaRNA
Other
C
VTRNA1-2
VTRNA2-1
VTRNA1-1
RNU5F-1
RNU6A-1
RNU5B-1
RNU4-1
RNU6ATAC
RNU6-6P
MT-TM
MT-TE
MT-TV
MT-TH
RN7SL1
7SK
7SK

Title: Expanding the landscape of systemic sclerosis-related autoantibodies through RNA immunoprecipitation coupled with massive parallel sequencing

Authors: Janire Perurena-Prieto[a,b,c], María Teresa Sanz-Martínez[a,b], Laura Viñas-Giménez[a,b], Claudia Codina-Clavaguera[d,e], Laura Triginer[e], Fernando Gordillo-González[f,1], Eduardo Andrés-León[f], Laura Batlle-Masó [g,h,i], Javier Martin[f], Albert Selva-O'Callaghan[d,e], Ricardo Pujol[c,j], Neil J McHugh[k], Sarah L Tansley[k], Roger Colobran [a,b,c,l]*, Alfredo Guillen-Del-Castillo [d,e]*, Carmen Pilar Simeón-Aznar[d,e]

SUPPLEMENTARY MATERIAL

Supplementary Patients and Methods

Supplementary Figures S1-S11

Supplementary Tables S1-S4

SUPPLEMENTARY PATIENTS AND METHODS

Patients

Overall, 307 patients with SSc who were visited in the Scleroderma Unit of the Hospital Universitari Vall d'Hebron were included in the study. All patients met the LeRoy and Medsger criteria [1][23], or the 2013 ACR/EULAR classification criteria for systemic sclerosis (SSc) [2]. Clinical and laboratory data were collected from medical records. Age at onset of disease was considered as the moment of the first clinical manifestation attributable to the disease, including Raynaud's phenomenon (RP). We considered the presence of digital vasculopathy if pitting, ulcers or gangrene were present and other causes such as trauma were excluded by treating doctor. The patients were classified considering LeRoy and Medsger's criteria [3]. Limited cutaneous SSc (lcSSc) was considered if the cutaneous thickening was distal to the elbows and knees and face was affected or not; diffuse cutaneous SSc (dcSSc) was considered if skin thickening was also proximal to elbows and knees; sine scleroderma SSc (ssSSc) was defined by RP or peripheral vascular equivalents (pitting scars, typical capillaroscopic alterations), typical scleroderma visceral involvement (oesophageal hipomotility, interstitial lung disease, pulmonary arterial hypertension, scleroderma renal crisis or sclerodermic cardiomyopathy) and antinuclear autoantibody positivity. In relation to capillaroscopic findings, patients were categorized in "early", "active" and "late" taking into account Cutolo's classification. Pulmonary arterial hypertension (PAH) was defined as presence of > 20 mmHg mean pulmonary arterial pressure with ≤ 15 mmHg pulmonary arterial wedge pressure and > 2 Wood units pulmonary vascular resistance measured by right heart catheterization. Interstitial lung disease (ILD) was determined by the presence of interstitial pattern on computerized tomography (CT). Gastrointestinal involvement was considered if oesophageal hypomotility, gastric antral vascular ectasia (GAVE) or bacterial overgrowth was present. Scleroderma renal crisis (SRC) was considered by the combination of sudden onset or worsening of moderate or severe systemic arterial hypertension (>160/90mmHg) as well as features of malignant hypertension, or the presence of a rapid deterioration of renal function in a period of less than one month. Heart involvement was defined as conduction abnormalities established by electrocardiogram or Holter monitoring; pericardial involvement demonstrated by echocardiogram, CT or magnetic resonance imaging (MRI); scleroderma cardiomyopathy or cardiac fibrosis perceptible by MRI; ischemic heart disease in absence of cardiovascular risk factors (CVRF); mitral insufficiency and left ventricular diastolic dysfunction without traditional CVRF proved by echocardiogram, as well as left ventricular

ejection fraction lower than 50% or right ventricular ejection fraction lower than 40% assessed by echocardiogram or MRI. Myositis was established by the presence of skeletal muscle weakness and evidence of muscular inflammation, detected by elevation of muscular enzyme levels, electromyographic or histological findings. The presence of neoplasia was defined by pathological report. One serum sample from each patient was tested for autoantibodies. This study was approved by the ethics Committee for Clinical Research (PG(AG)07/2015) and written informed consent was obtained from all the patients.

Indirect immunofluorescence (IIF)

ANAs and anti-cytoplasmic autoantibodies were detected by immunofluorescence test performed on Hep-2 cells as substrate (INOVA, San Diego, USA). IIF patterns were detected at serum screening dilutions of 1:80, and positive reactions were categorized according to the International Consensus on ANA Patterns (ICAP) classification [4]. Titers were determined by 2-fold endpoint titration. Anti-centromere autoantibodies were determined by IIF.

Immunoblotting (IB)

Autoantibodies against Scl70, RNApol III (RP11, RP155), fibrillarin, NOR-90, Th/To, PM/Scl (PM/Scl-100, PM/Scl-75) and Ku were assessed by a line blot commercial assay (Systemic Sclerosis Profile Euroline® Blot test kit, Euroimmun, Lübeck, Germany) [5] according to the manufacturer's instructions.

Chemiluminescence immunoassay (CLIA)

Anti-U1-RNP autoantibodies were tested by chemiluminescence immunoassay (INOVA, San Diego, EEUU) according to the manufacturer's instructions.

RNA immunoprecipitation (RNA IP)

Sixty-eight sera that were negative for all the autoantibodies assessed by the commercial IB and CLIA, and for anti-centromere by IIF, were tested by RNA IP (**Figure 1**). Protein-A Sepharose beads were incubated with patients' sera at 4ºC for 16h. After three washes, antibody-coated beads were incubated with ribonucleoprotein extracts obtained from $10 \cdot 10^6$ K562 cells at 4ºC for 90'. Antibody-coated beads were then washed three times and a classical phenol:chloroform:isoamyl alcohol RNA extraction protocol was performed. Immunoprecipitated RNA was analysed by 8% or 10% Urea-TBA-PAGE and subsequently visualized by SYBR™ Green II RNA gel stain using a UV transilluminator. Specificities were

verified using reference sera when possible and a healthy donor serum was used in all RNA IP assays.

RNA immunoprecipitation coupled with RNAseq (RIP-Seq)

Fifty-seven sera that were negative for all autoantibodies assessed by the commercial IB and CLIA, and for anti-centromere by IIF, were tested by RIP-Seq together with sera from three healthy donors (**Figure 1**). Immunoprecipitated RNA was obtained by the previously described RNA IP protocol. Isolated RNA was purified by RNeasy MinElute Cleanup kit (Qiagen, Hilden, Germany).

Libraries were prepared using the NEBNext® Small RNA Library Prep Set for Illumina® kit (New England Biolabs, Ipswich, USA) according to the manufacturer's protocol. Briefly, RNA was subjected to adaptor 3´ and 5´ ligation and first strand cDNA synthesis. Subsequently, DNA fragments with adaptor strands on both ends were selectively enriched by PCR. Library amplification was performed by PCR using NEBNext® Multiplex Oligos for Illumina®. All purification steps were performed using AgenCourt AMPure XP beads (Beckman Coulter, Brea, USA). Final libraries were analysed using Agilent Bioanalyzer (Agilent, Santa Clara, USA) to estimate DNA quantity and check size distribution. Samples were pooled and size selection was performed using a 6% Novex™ TBE-PAGE. DNA from the final pool was quantified by qPCR using KAPA Library Quantification Kit (Kapa Biosystems, Wilmington, USA) prior to amplification with cBot (Illumina, San Diego, USA). Libraries were sequenced with single reads of 50bp (8bp index) on HiSeq2500 (Illumina, San Diego, USA).

Obtained reads were mapped using Salmon vs. GRCh38 genome and GENCODE v38 annotation [6,7][354, 355]. Final counts were used for calculating differential expression analysis with DESEQ2. For microRNAseq analysis, reads were trimmed using skewer, and trimmed reads longer than 15bp were mapped using the same genome and annotation using ShortStack. The reads were counted using HTSeq. Reads were also mapped using STAR vs. GRCh38 genome and GENCODE v38 annotation in order to check if samples that were pseudo-mapped to protein coding genes by Salmon were only mapped to small nucleolar RNAs (snoRNAs) codified in the introns of those protein coding genes [8][356].

RNA molecule candidate selection

The analysis was carried out using the R programming language (version 4.1.2) [9][357]. In order to detect RNA molecules with an exacerbated enrichment across different samples, four outlier detection tests were used in combination: a) values with an absolute difference between the first or third quartile greater than 1.5 times the interquartile range (IQR) (Tukey's

rule); b) values whose difference from the mean is greater than 3 absolute deviations from the mean (MAD); c) the Tukey rule criterion corrected for non-normal distributions and multiple testing, using the tukey_mc_up function from the bigutilsr package (v0.3.4) [10]; and d) selecting RNA molecules with a standard deviation higher than Q3 of the distribution of standard deviations of all RNA molecules, calculated with the rowSds function of the genefilter R package (v1.76.0) [11].

A high number of RNA molecules were present at low counts in all samples and were considered as nonspecific immunoprecipitation (i.e. background noise). Considering the distribution of the median counts of protein coding RNA molecules (that should not be specifically immunoprecipitated as they are not part of any ribonucleoprotein in a stable manner), a cut-off of 100 median counts was established. Therefore, these four methods were applied only on genes having median counts greater than 100. On these resulting genes, the ratio between the square of its maximum expression value and its median was calculated. This metric reflects the distance between the highest outlier and the rest of the distribution, so that a gene with a high value of this metric has outliers further from the rest of the values, and therefore, they are more easily detectable. This metric was evaluated by receiver operating characteristics (ROC) curve with the purpose of establishing the positivity cut-off with best sensitivity and specificity.

On RNA molecules above this cut-off a weighted gene co-expression network analysis (WGCNA) was performed using the R WGCNA package (v1.71) [12][358] so as to identify modules formed by RNA molecules with a similar enrichment pattern across samples. Due to the small size of the initial RNA molecule set, a soft-thresholding power of 26, a minimum cluster size of 5 genes and a merge distance cut of 0.10 was set.

Statistical analysis for evaluation of demographical and clinical characteristics

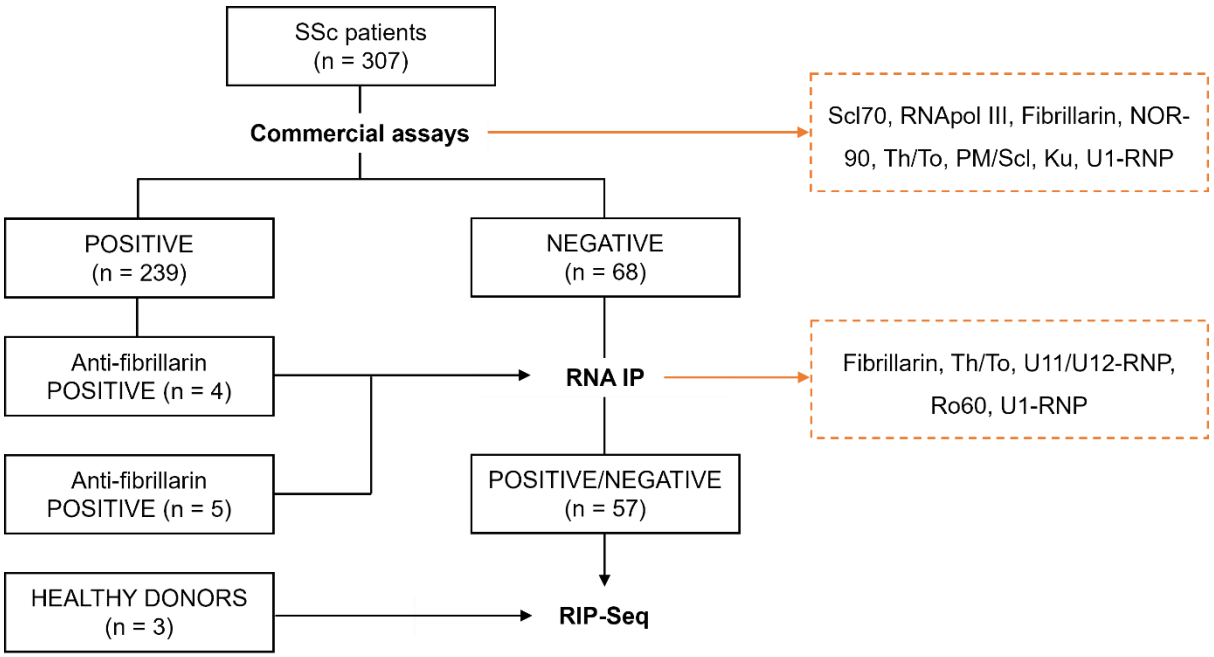
Continuous variables were compared with the Wilcoxon-Mann-Whitney test. Associations between autoantibody status and clinical and demographical features were identified using Fisher's exact test and univariate binary logistic regression models. A 2-sided significance level was set at $p<0.05$. Statistical analyses were performed using IBM-SPSS 20.0 (IBM, Chicago, IL, USA).

REFERENCES

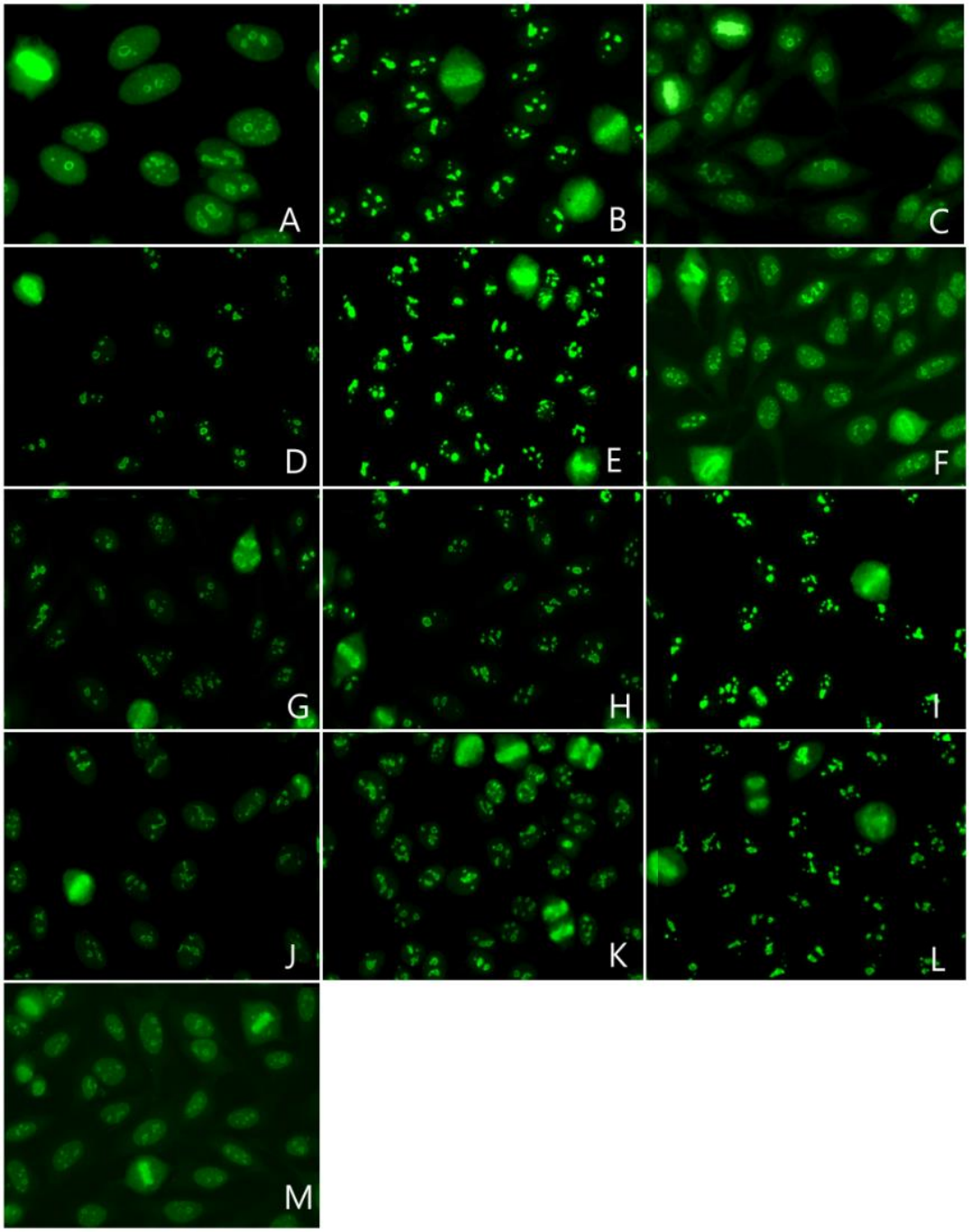
1. LeRoy EC, Medsger TAJ. Criteria for the classification of early systemic sclerosis. J Rheumatol. 2001;28:1573–6.
2. Van Den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: An american college of rheumatology/European league against rheumatism collaborative initiative. Arthritis Rheum. 2013;65:2737–47.
3. Masi AT. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Arthritis Rheum. 1980;23:581–90.
4. Damoiseaux J, Andrade LEC, Carballo OG, Conrad K, Francescantonio PLC, Fritzler MJ, et al. Clinical relevance of HEp-2 indirect immunofluorescent patterns: the International Consensus on ANA patterns (ICAP) perspective. Ann Rheum Dis. 2019;:annrheumdis-2018-214436.
5. Bonroy C, Van Praet J, Smith V, Van Steendam K, Mimori T, Deschepper E, et al. Optimization and diagnostic performance of a single multiparameter lineblot in the serological workup of systemic sclerosis. J Immunol Methods. 2012;379:53–60.
6. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14:417–9.
7. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.
8. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.
9. R Core Team. R: A language and environment for statistical computing (4.1.2) [Computer software]. R Foundation for Statistical Computing. Vienna, Austria. URL https://www.R-project.org/. 2021.
10. Privé F. bigutilsr: Utility Functions for Large-scale Data (0.3.4) [Computer software]. https://github.com/privefl/bigutilsr.
11. Gentleman R, Carey V, Huber W, Hahne F. genefilter: Methods for filtering genes from high-throughput experiments. (R package version 1.86.0.) [Computer software]. DOI: 10.18129/B9.bioc.genefilter. 2024.
12. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.

SUPPLEMENTARY FIGURES

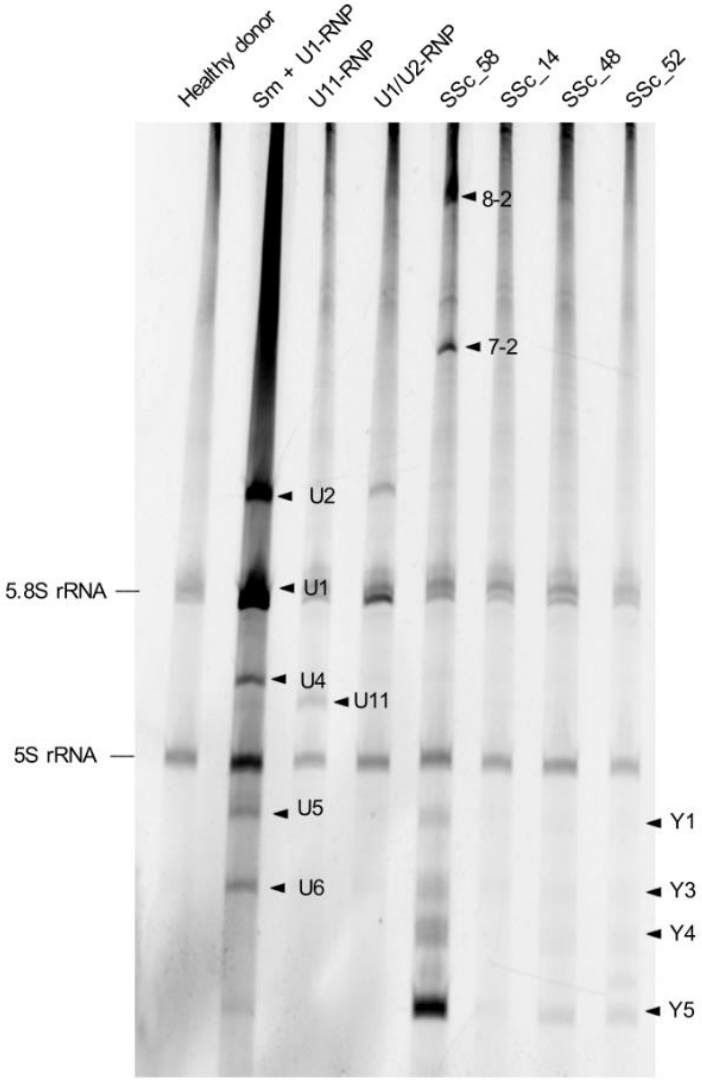
Supplementary Figure S1: Subject cohorts and methods used for different autoantibody detection. Overall, 307 SSc patients were analysed by IIF on Hep-2 cells, IB and CLIA. Sixty-eight patients (22.1%) that were negative for anti-centromere, Scl70, RNApol III, PM/Scl, fibrillarin, Th/To, Ku and U1-RNP autoantibodies, were further analysed by RNA IP. Fifty-seven sera that were analysed by RNA IP together with sera from three healthy donors were also tested by RIP-Seq, Positive and negative samples by RNA IP were included in the RIP-Seq assay.



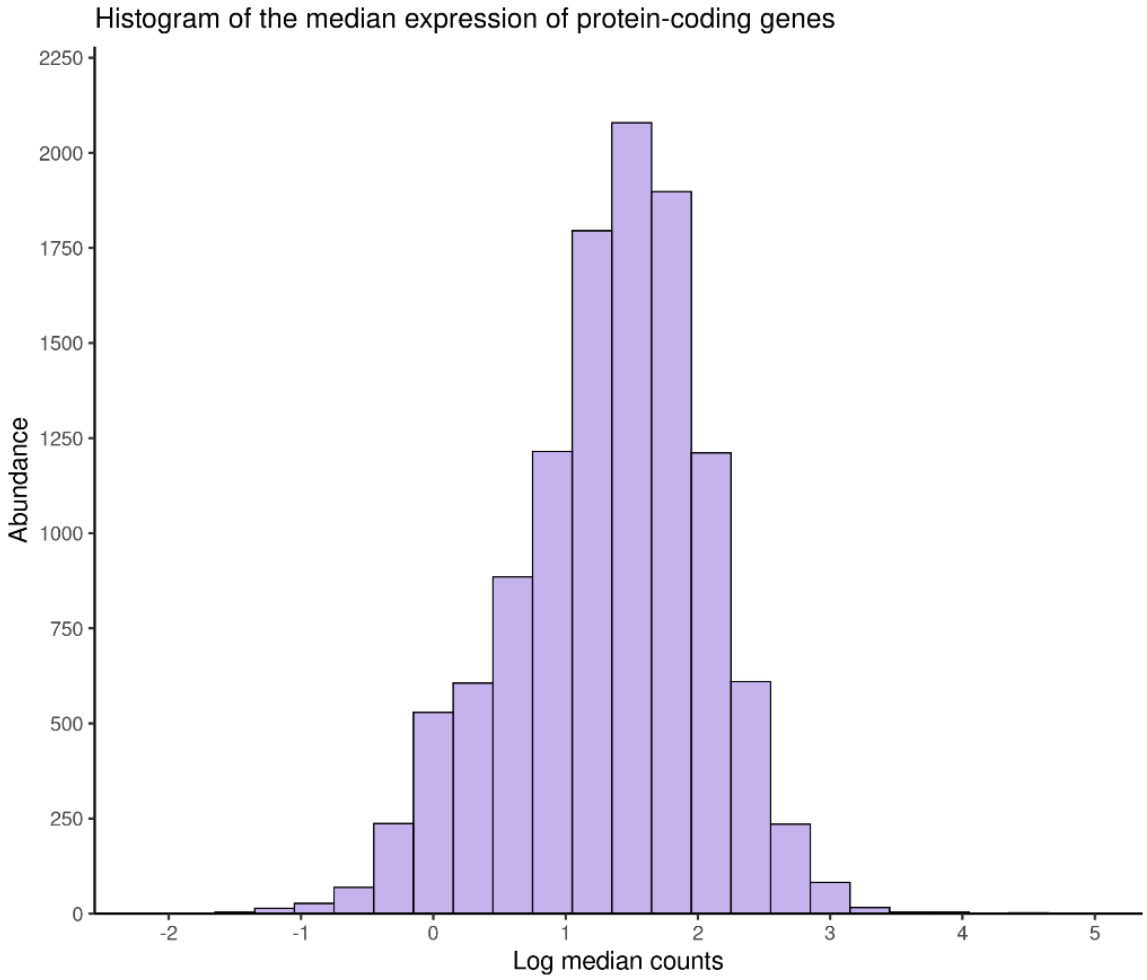
Supplementary Figure S2: HEp-2 staining of selected patients with an AC-9 pattern. IIF pattern observed on HEp-2 cells of samples SSc_30 (A), SSc_46 (B), SSc_48 (C), SSc_78 (D), SSc_79 (E), SSc_80 (F), SSc_81 (G), SSc_82 (H), SSc_83 (I) and SSc_84 (J), SSc_85 (K), SSc_86 (L) and SSc_70 (M).



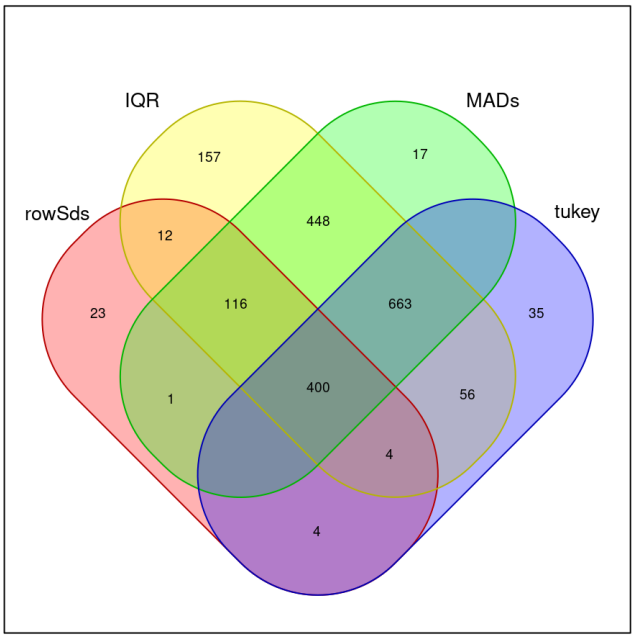
Supplementary Figure S3: traditional RNA IP results of samples detected to be immunoprecipitating Y1, Y3, Y4 and Y5 RNA molecules associated with anti-Ro60 autoantibodies by RIP-Seq. TBE-Urea PAGE obtained by traditional RNA IP of samples SSc_14, SSc_48, SSc_52 and SSc_58. Patient SSc_58, SSc_48 and SSc_52 were positive for anti-Ro60 autoantibodies by RNA IP.



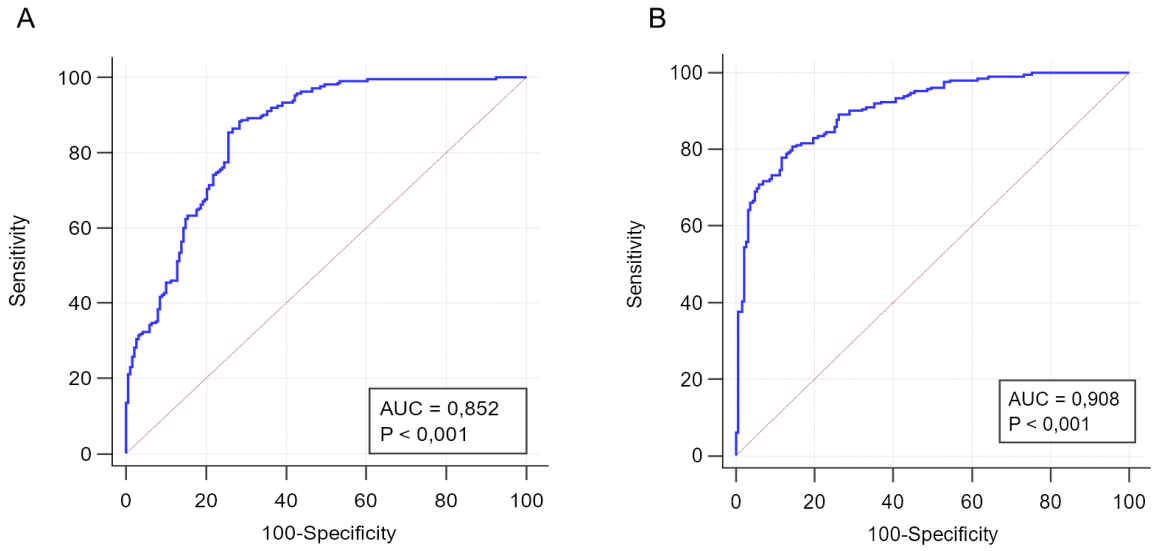
Supplementary Figure S4: distribution of median counts obtained in all protein coding RNA molecules detected by RIP-Seq. Quantity of protein coding genes for different ranges of median counts (log_{10}) obtained along all samples for each gene. Protein coding RNA molecules detected by RIP-Seq rarely presented a median count of more than 100. Those RNA molecules with a median count of 0 were eliminated from the graph as it is not possible to express 0 as a logarithm.



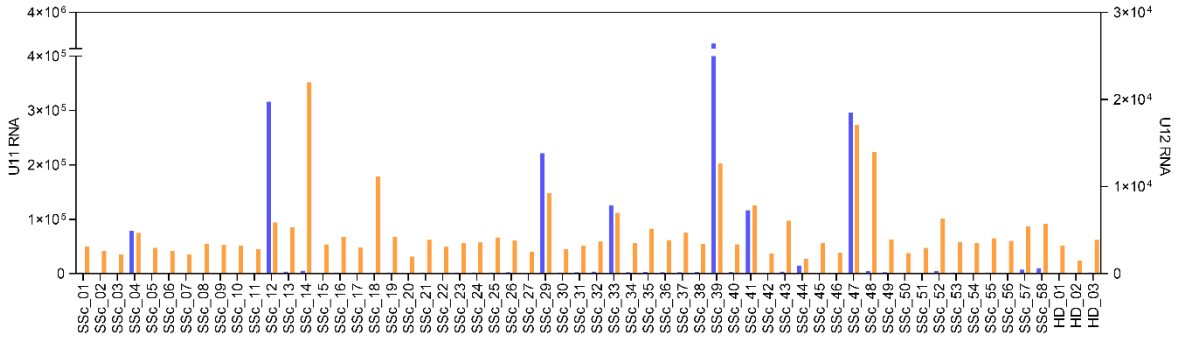
Supplementary Figure S5: Venn diagram showing the results obtained from the four methods used for outlier detection. Four different methods were used for outlier detection. RNA molecules in which outliers were detected by all four methods (n=400) were selected for further analysis.



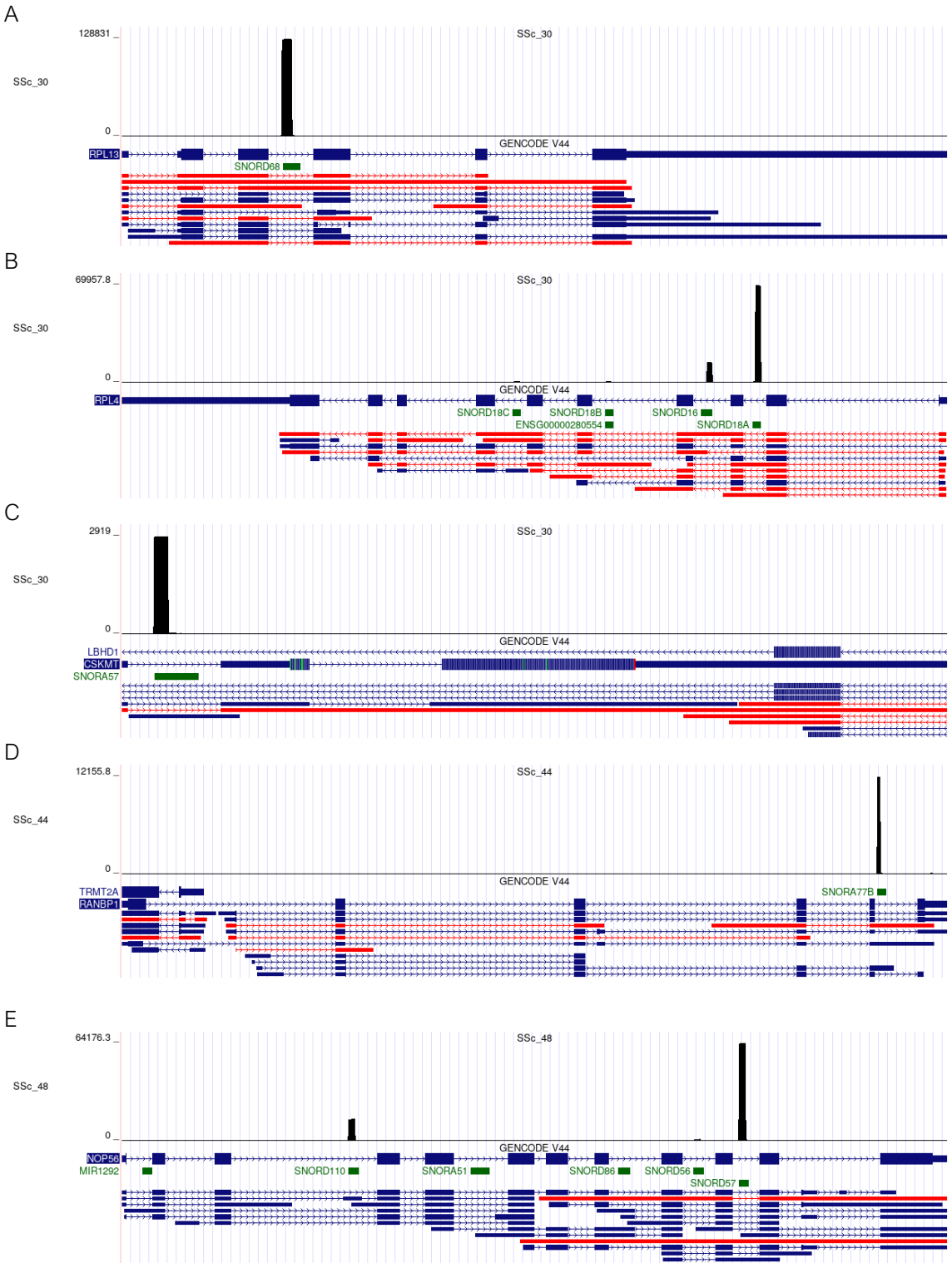
Supplementary Figure S6: ROC curves of the two criteria used for selecting RNA molecules specifically immunoprecipitated by autoantibodies. Two criteria were evaluated by receiver operating characteristics (ROC) curve, considering that all protein coding RNA molecules were false positives. (A) ROC curve obtained using the ratio calculated as the highest outlier count of each specific RNA molecule between the median counts of that same RNA among all samples. (B) ROC curve obtained using the same ratio multiplied by the highest outlier count of each specific RNA molecule.



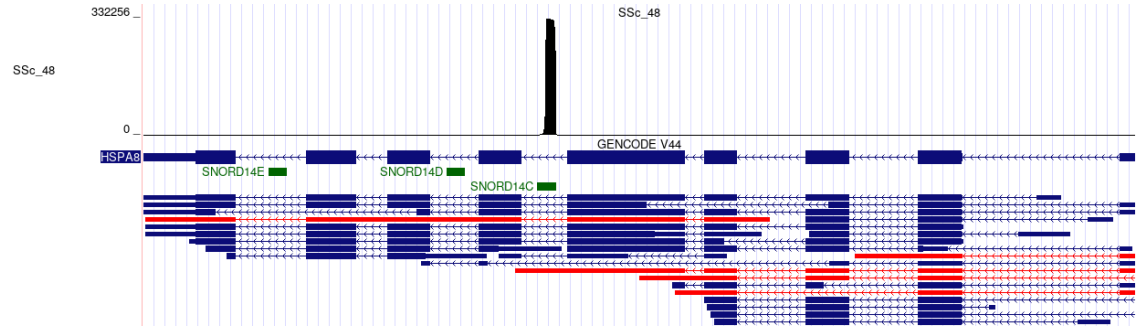
Supplementary Figure S7: RIP-Seq results of U11 RNA and U12 RNA molecules. Counts obtained for U11 RNA (blue) and U12 RNA (orange) for each sample.



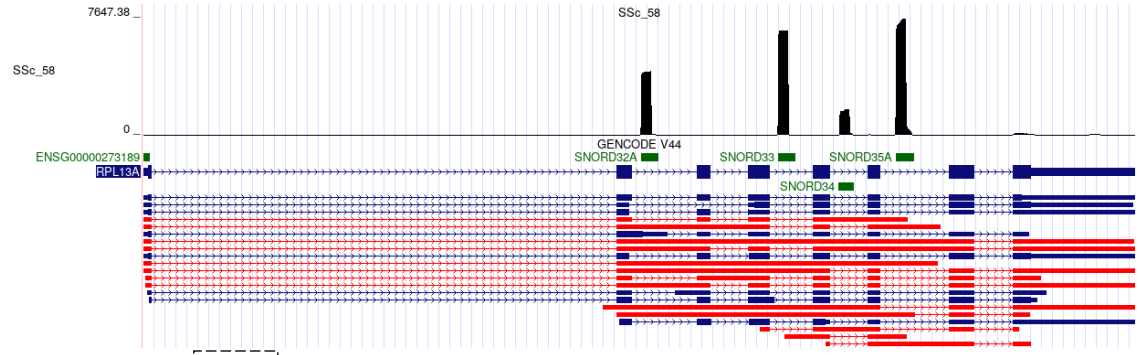
Supplementary Figure S8: read mapping of protein coding RNA molecules selected based on max²/median ratio. Genome browser was used to analyse bam files of samples with the highest read count for each protein coding RNA molecule. If the reads obtained by RNAseq had been mapped to sequences coding for snoRNAs was checked. SSc_30, SSc_44, SSc_48 and SSc_58 samples' reads mapped on RPL13 (A), RPL4 (B) and CSKMT (C); RANBP1 (D); NOP56 (E) and HSPA8 (F); RPL13A (G) and ANK1 (H) genes, respectively.



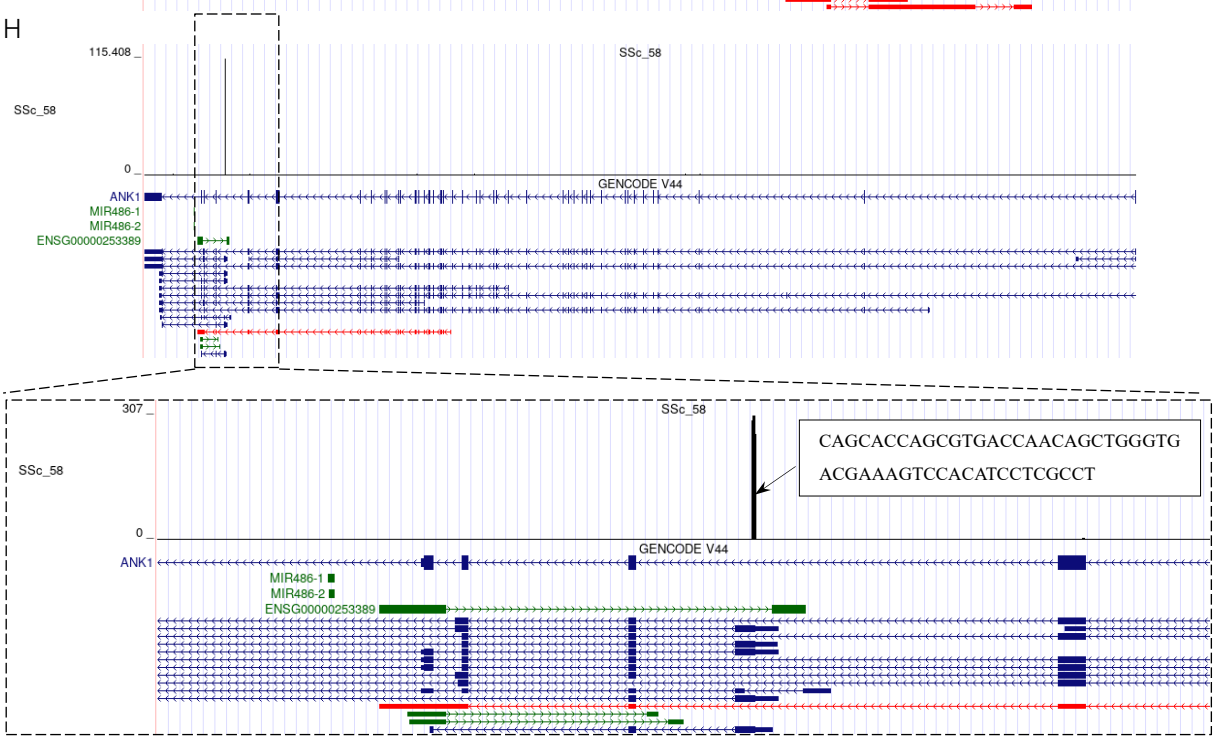
F


332256
SSc_48
SSc_48
0
GENCODE V44
HSPA8
SNORD14E
SNORD14D
SNORD14C

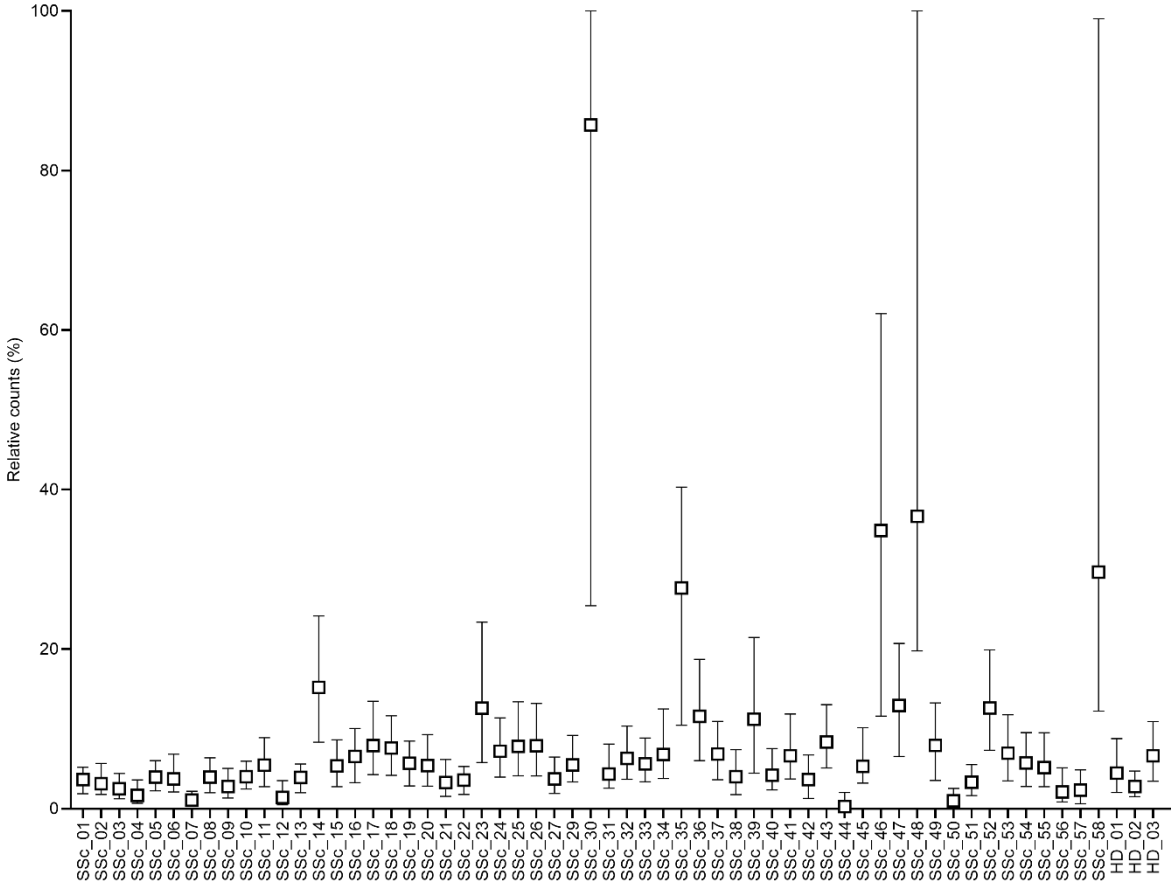
G


7647.38
SSc_58
SSc_58
0
GENCODE V44
ENSG00000273189
RPL13A
SNORD32A
SNORD33
SNORD35A
SNORD34

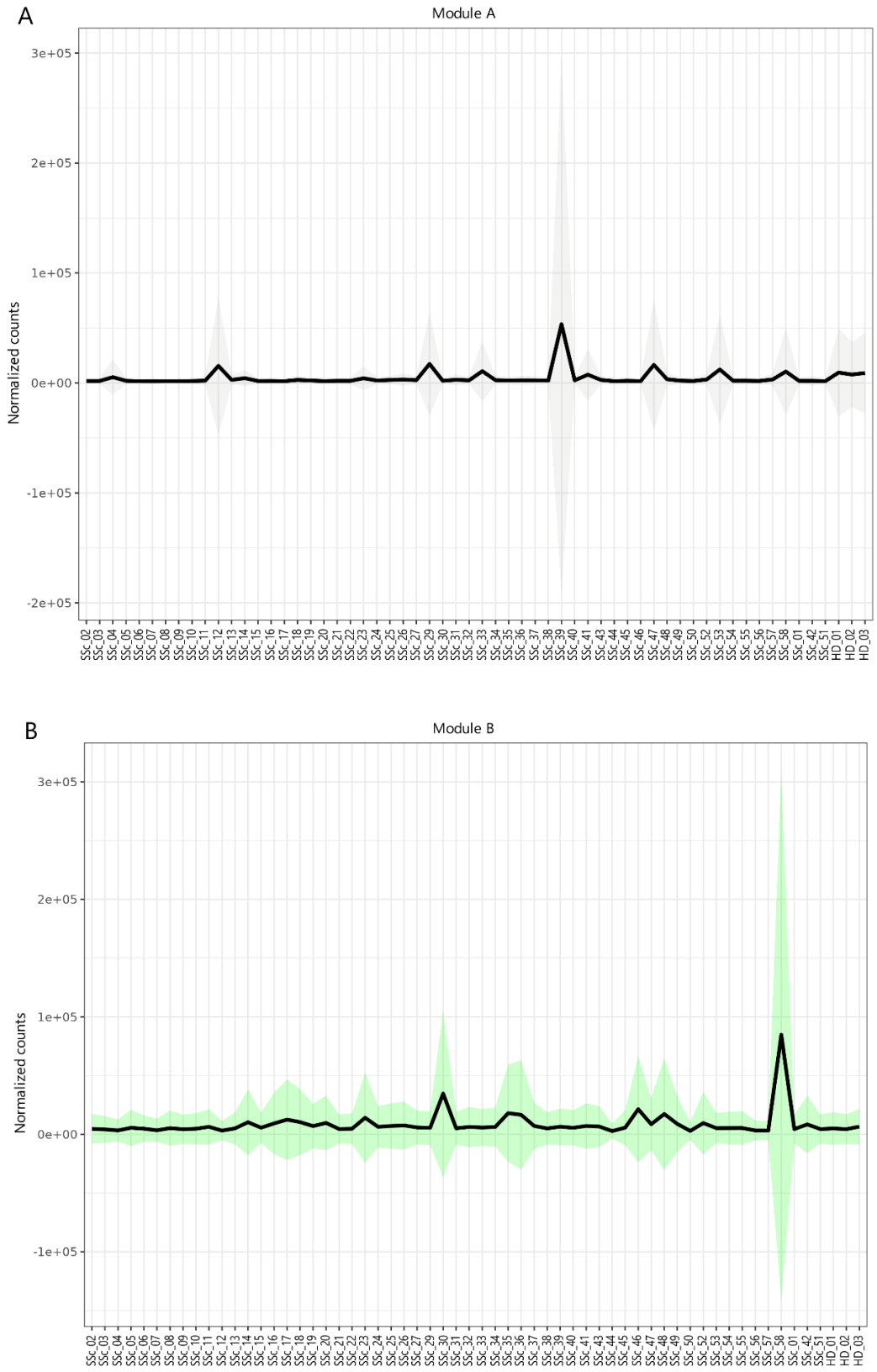
H

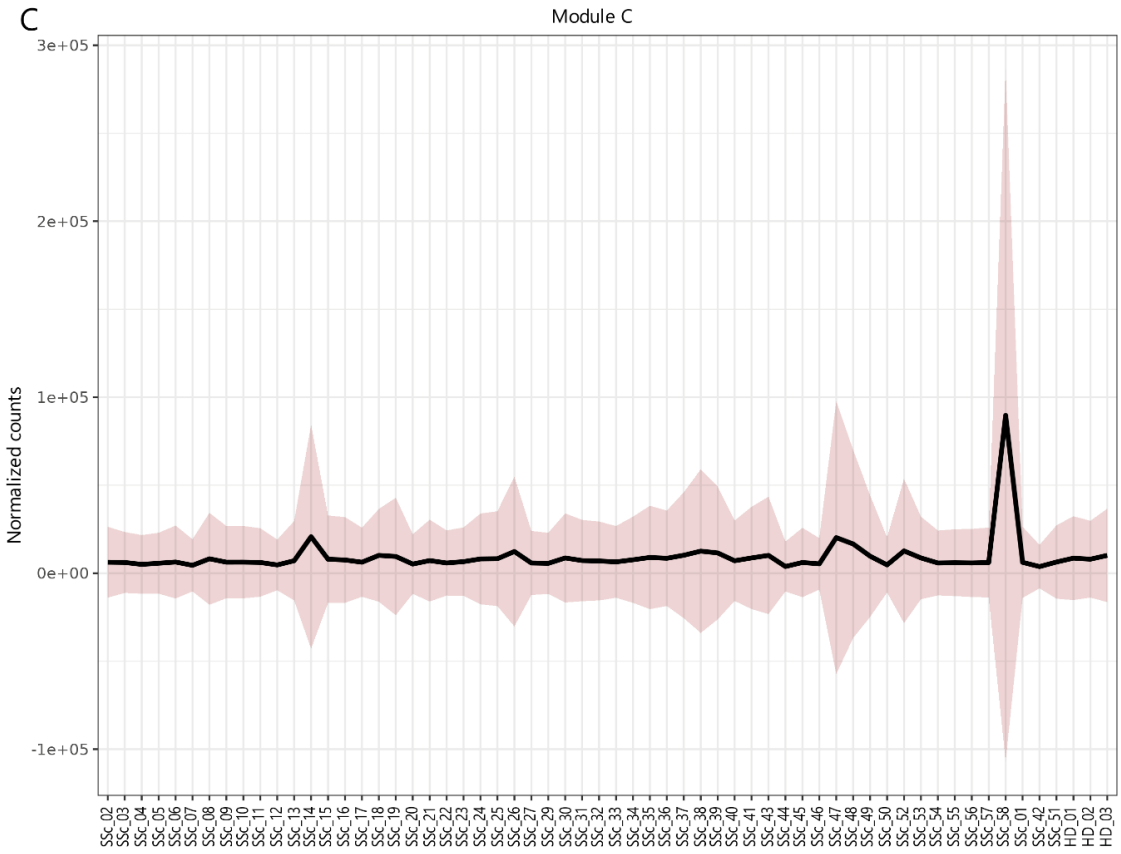

115.408
SSc_58
SSc_58
0
GENCODE V44
ANK1
MIR486-1
MIR486-2
ENSG00000253389
307
SSc_58
SSc_58
CAGCACCAGCGTGACCAACAGCTGGGTG
ACGAAAGTCCACATCCTCGCCT
0
GENCODE V44
ANK1
MIR486-1
MIR486-2
ENSG00000253389

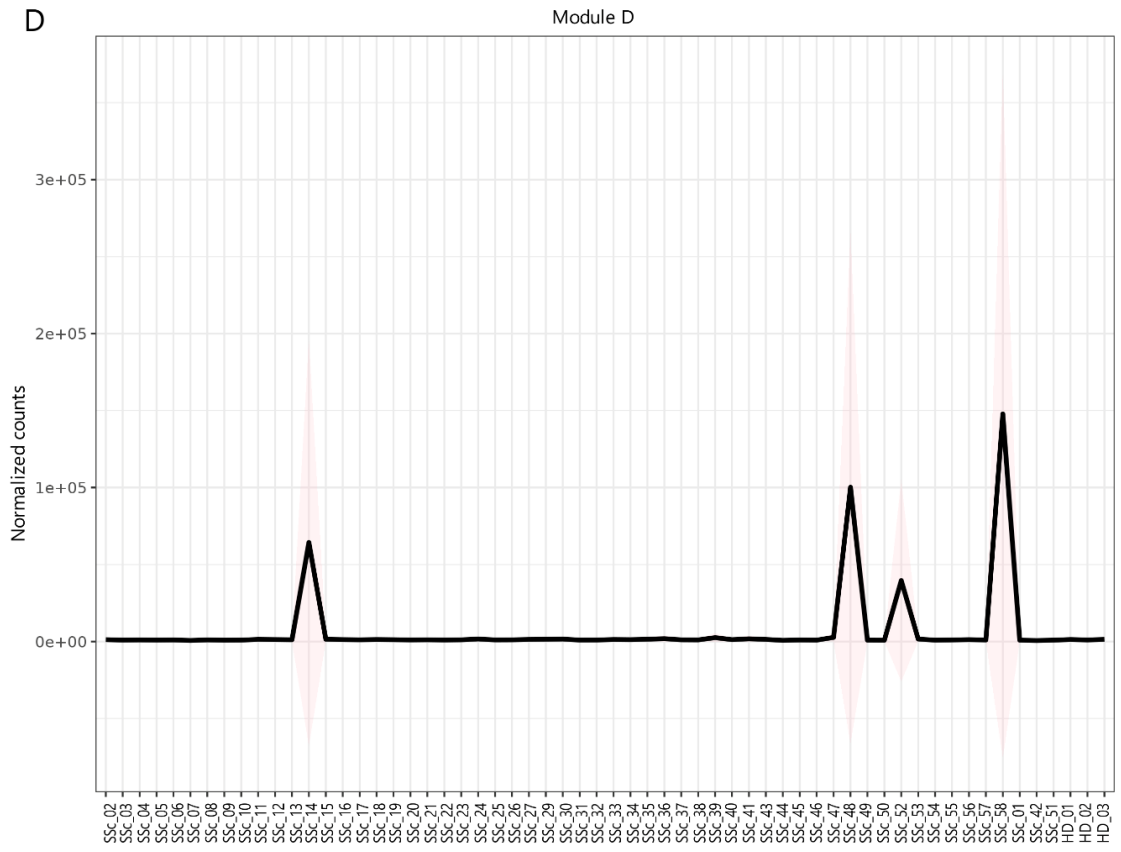
Supplementary Figure S9: graphical distribution of relative counts (%) of C/D box snoRNA molecules among all samples. Due to the broad range of counts of different C/D box snoRNA molecules, to be able to compare them as a group among different samples, counts were normalized in relation to the maximum value obtained for each C/D box snoRNA molecule and expressed as a percentage. Median and interquartile range of relative counts (%) obtained for all the samples on snoRNA molecules.

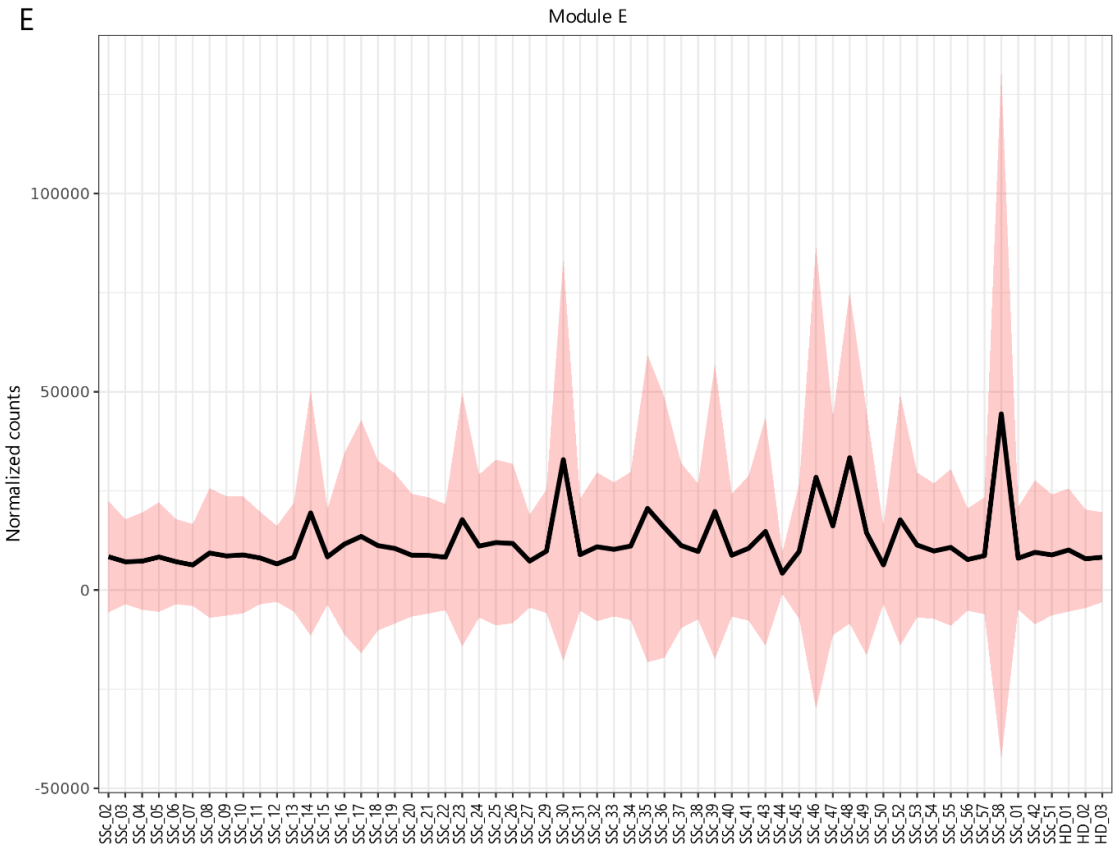


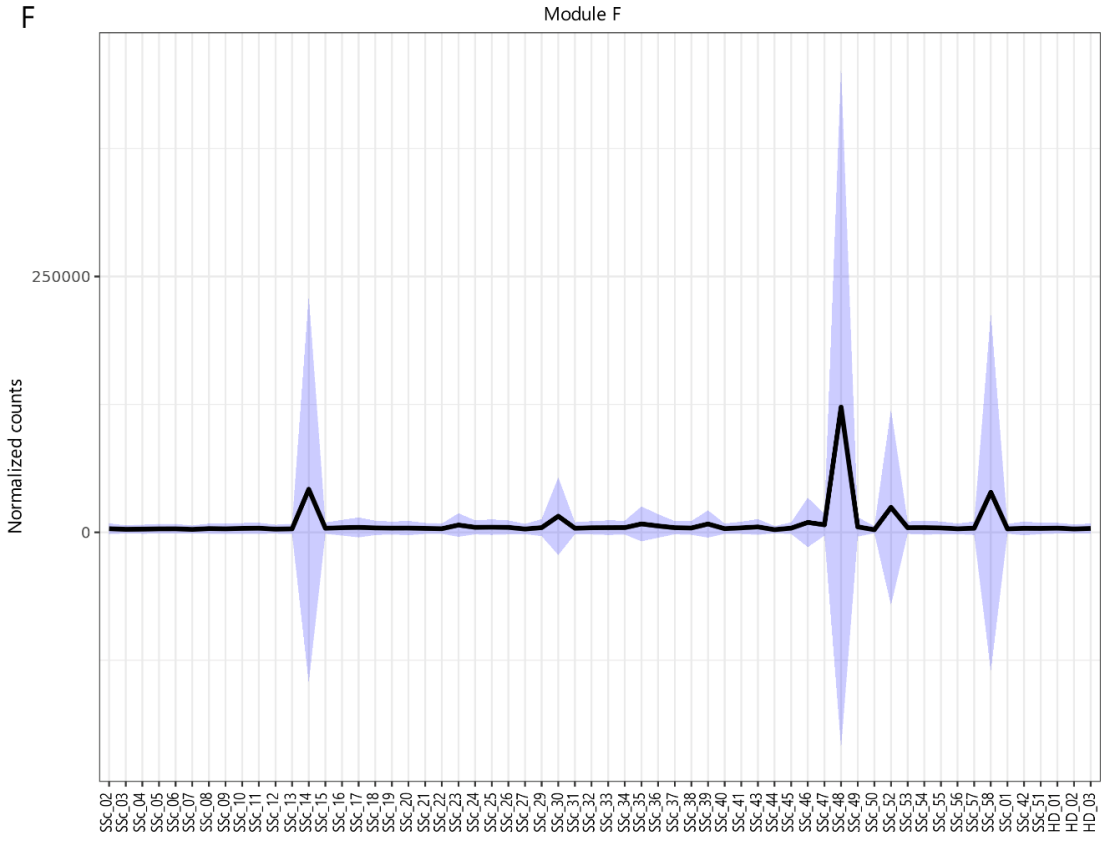
Supplementary Figure S10: WGCNA results. Normalized counts of RNA molecules included in modules A, B, C, D, E, F, G, H and I for each sample.




C
Module C
Normalized counts
3e+05
2e+05
1e+05
0e+00
-1e+05
SSc_02 SSc_03 SSc_04 SSc_05 SSc_06 SSc_07 SSc_08 SSc_09 SSc_10 SSc_11 SSc_12 SSc_13 SSc_14 SSc_15 SSc_16 SSc_17 SSc_18 SSc_19 SSc_20 SSc_21 SSc_22 SSc_23 SSc_24 SSc_25 SSc_26 SSc_27 SSc_29 SSc_30 SSc_31 SSc_32 SSc_33 SSc_34 SSc_35 SSc_36 SSc_37 SSc_38 SSc_39 SSc_40 SSc_41 SSc_43 SSc_44 SSc_45 SSc_46 SSc_47 SSc_48 SSc_49 SSc_50 SSc_52 SSc_53 SSc_54 SSc_55 SSc_56 SSc_57 SSc_58 SSc_01 SSc_42 SSc_51 HD_01 HD_02 HD_03

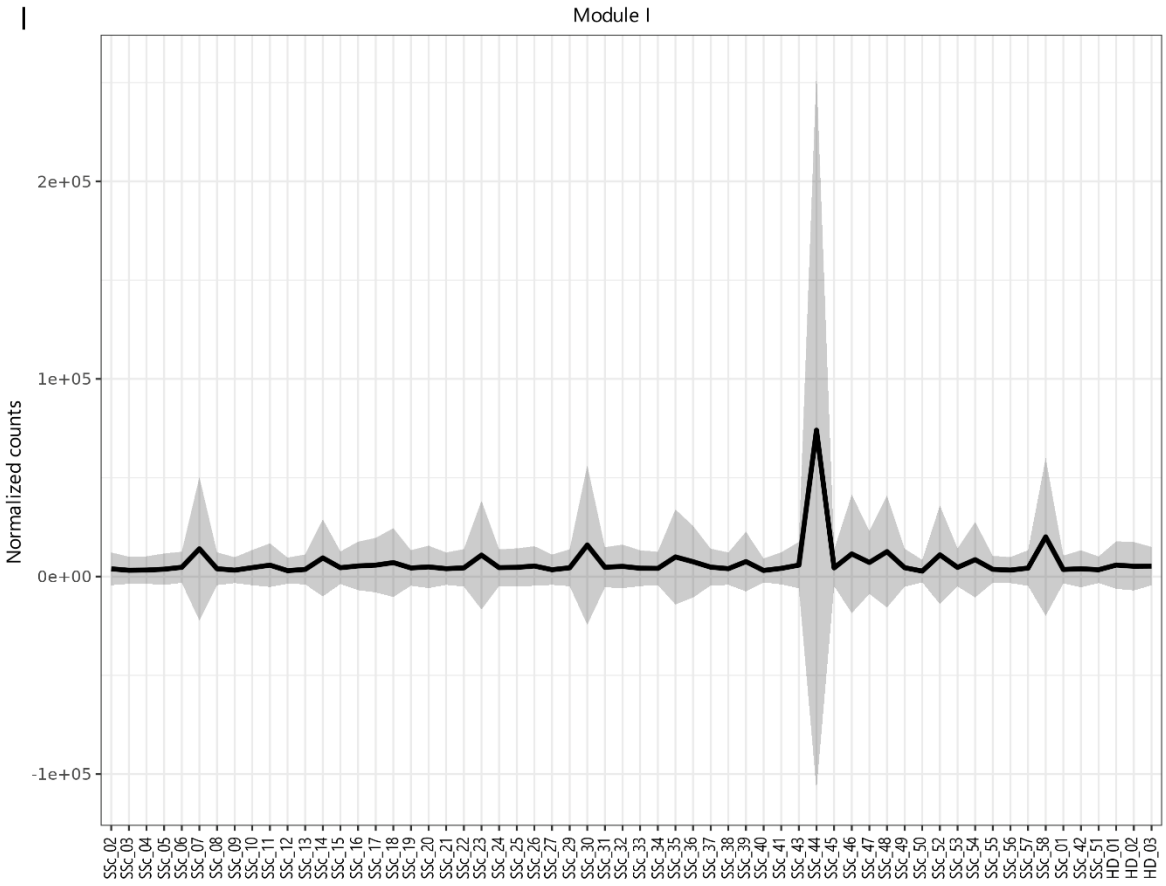

D
Module D
Normalized counts
3e+05
2e+05
1e+05
0e+00
SSc_02 SSc_03 SSc_04 SSc_05 SSc_06 SSc_07 SSc_08 SSc_09 SSc_10 SSc_11 SSc_12 SSc_13 SSc_14 SSc_15 SSc_16 SSc_17 SSc_18 SSc_19 SSc_20 SSc_21 SSc_22 SSc_23 SSc_24 SSc_25 SSc_26 SSc_27 SSc_29 SSc_30 SSc_31 SSc_32 SSc_33 SSc_34 SSc_35 SSc_36 SSc_37 SSc_38 SSc_39 SSc_40 SSc_41 SSc_43 SSc_44 SSc_45 SSc_46 SSc_47 SSc_48 SSc_49 SSc_50 SSc_52 SSc_53 SSc_54 SSc_55 SSc_56 SSc_57 SSc_58 SSc_01 SSc_42 SSc_51 HD_01 HD_02 HD_03


E
Module E
Normalized counts
100000
50000
0
-50000
SSc_02
SSc_03
SSc_04
SSc_05
SSc_06
SSc_07
SSc_08
SSc_09
SSc_10
SSc_11
SSc_12
SSc_13
SSc_14
SSc_15
SSc_16
SSc_17
SSc_18
SSc_19
SSc_20
SSc_21
SSc_22
SSc_23
SSc_24
SSc_25
SSc_26
SSc_27
SSc_29
SSc_30
SSc_31
SSc_32
SSc_33
SSc_34
SSc_35
SSc_36
SSc_37
SSc_38
SSc_39
SSc_40
SSc_41
SSc_43
SSc_44
SSc_45
SSc_46
SSc_47
SSc_48
SSc_49
SSc_50
SSc_52
SSc_53
SSc_54
SSc_55
SSc_56
SSc_57
SSc_58
SSc_01
SSc_42
SSc_51
HD_01
HD_02
HD_03

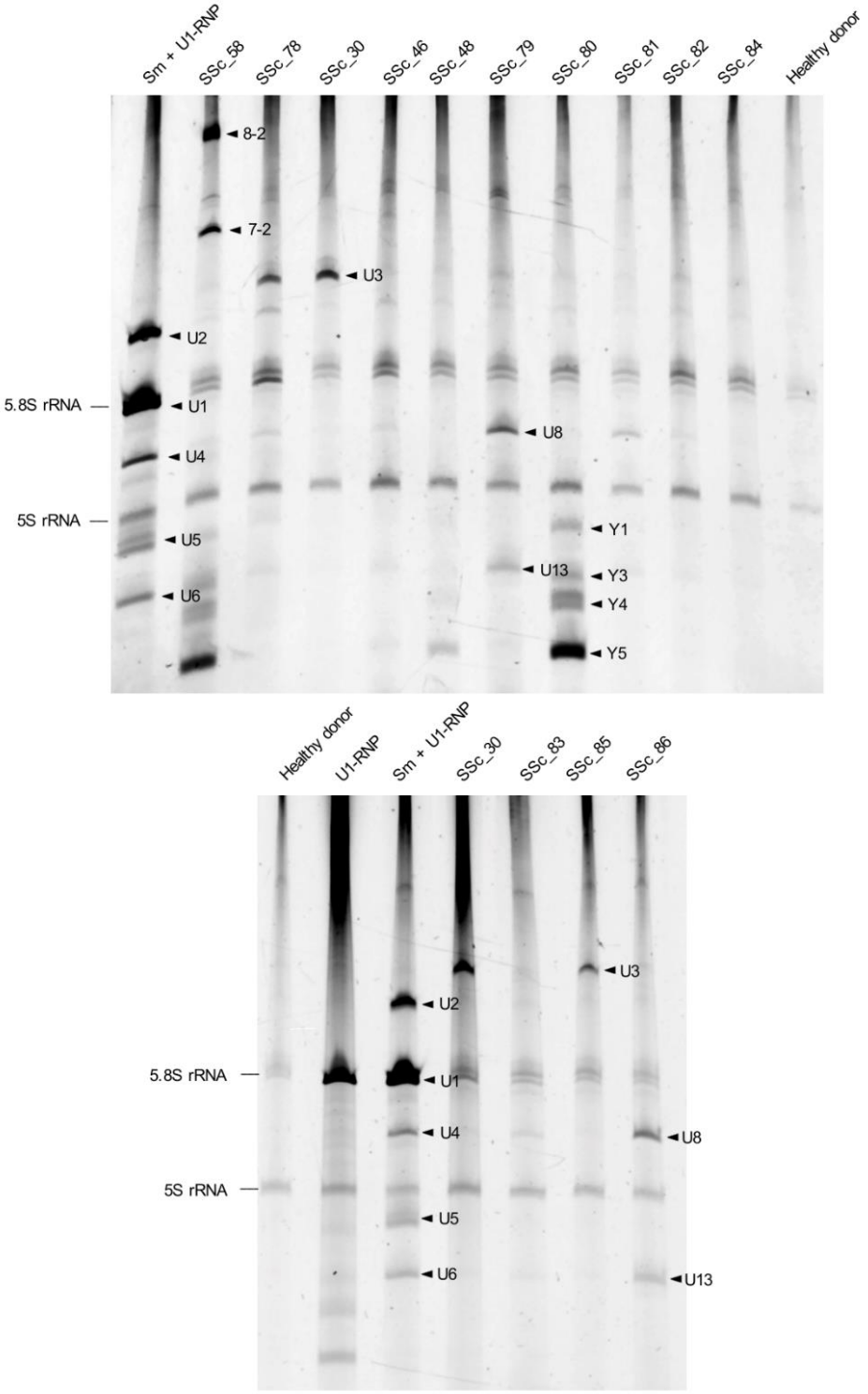

F
Module F
Normalized counts
250000
0
SSc_02
SSc_03
SSc_04
SSc_05
SSc_06
SSc_07
SSc_08
SSc_09
SSc_10
SSc_11
SSc_12
SSc_13
SSc_14
SSc_15
SSc_16
SSc_17
SSc_18
SSc_19
SSc_20
SSc_21
SSc_22
SSc_23
SSc_24
SSc_25
SSc_26
SSc_27
SSc_29
SSc_30
SSc_31
SSc_32
SSc_33
SSc_34
SSc_35
SSc_36
SSc_37
SSc_38
SSc_39
SSc_40
SSc_41
SSc_43
SSc_44
SSc_45
SSc_46
SSc_47
SSc_48
SSc_49
SSc_50
SSc_52
SSc_53
SSc_54
SSc_55
SSc_56
SSc_57
SSc_58
SSc_01
SSc_42
SSc_51
HD_01
HD_02
HD_03


G
Module G
250000
0
Normalized counts
SSc_02
SSc_03
SSc_04
SSc_05
SSc_06
SSc_07
SSc_08
SSc_09
SSc_10
SSc_11
SSc_12
SSc_13
SSc_14
SSc_15
SSc_16
SSc_17
SSc_18
SSc_19
SSc_20
SSc_21
SSc_22
SSc_23
SSc_24
SSc_25
SSc_26
SSc_27
SSc_29
SSc_30
SSc_31
SSc_32
SSc_33
SSc_34
SSc_35
SSc_36
SSc_37
SSc_38
SSc_39
SSc_40
SSc_41
SSc_43
SSc_44
SSc_45
SSc_46
SSc_47
SSc_48
SSc_49
SSc_50
SSc_52
SSc_53
SSc_54
SSc_55
SSc_56
SSc_57
SSc_58
SSc_01
SSc_42
SSc_51
HD_01
HD_02
HD_03


H
Module H
4e+05
2e+05
0e+00
-2e+05
Normalized counts
SSc_02
SSc_03
SSc_04
SSc_05
SSc_06
SSc_07
SSc_08
SSc_09
SSc_10
SSc_11
SSc_12
SSc_13
SSc_14
SSc_15
SSc_16
SSc_17
SSc_18
SSc_19
SSc_20
SSc_21
SSc_22
SSc_23
SSc_24
SSc_25
SSc_26
SSc_27
SSc_29
SSc_30
SSc_31
SSc_32
SSc_33
SSc_34
SSc_35
SSc_36
SSc_37
SSc_38
SSc_39
SSc_40
SSc_41
SSc_43
SSc_44
SSc_45
SSc_46
SSc_47
SSc_48
SSc_49
SSc_50
SSc_52
SSc_53
SSc_54
SSc_55
SSc_56
SSc_57
SSc_58
SSc_01
SSc_42
SSc_51
HD_01
HD_02
HD_03


Module I
Normalized counts
2e+05
1e+05
0e+00
-1e+05
SSc_02
SSc_03
SSc_04
SSc_05
SSc_06
SSc_07
SSc_08
SSc_09
SSc_10
SSc_11
SSc_12
SSc_13
SSc_14
SSc_15
SSc_16
SSc_17
SSc_18
SSc_19
SSc_20
SSc_21
SSc_22
SSc_23
SSc_24
SSc_25
SSc_26
SSc_27
SSc_29
SSc_30
SSc_31
SSc_32
SSc_33
SSc_34
SSc_35
SSc_36
SSc_37
SSc_38
SSc_39
SSc_40
SSc_41
SSc_43
SSc_44
SSc_45
SSc_46
SSc_47
SSc_48
SSc_49
SSc_50
SSc_52
SSc_53
SSc_54
SSc_55
SSc_56
SSc_57
SSc_58
SSc_01
SSc_42
SSc_51
HD_01
HD_02
HD_03

Supplementary Figure S11: traditional RNA IP results of all the samples included in the study showing an AC-9 IIF pattern. TBE-Urea PAGE obtained by traditional RNA IP of samples SSc_30, SSc_46, SSc_48, SSc_78, SSc_79, SSc_80, SSc_81, SSc_82, SSc_83, SSc_84, SSc_85 and SSc_86.



Supplementary Table S1: IIF patterns observed on HEp-2 cells and RNA IP results of the 68 patients that were negative for all the autoantibodies assessed by screening commercial methods. Samples included in the RIP-Seq study are indicated.

No.	Sample	IIF pattern	RNA IP result	RIP-Seq tested
1	SSc_01	Homogeneous nucleolar (AC-8)	Anti-Th/To	Yes
2	SSc_02	Negative (AC-0)	Negative	Yes
3	SSc_03	Cytoplasmatic (Atypical)	Negative	Yes
4	SSc_04	Negative (AC-0)	Anti-U11/U12 RNP	Yes
5	SSc_05	Negative (AC-0)	Negative	Yes
6	SSc_06	Negative (AC-0)	Negative	Yes
7	SSc_07	Negative (AC-0)	Negative	Yes
8	SSc_08	Nuclear fine speckled (AC-4)	Negative	Yes
9	SSc_09	Negative (AC-0)	Negative	Yes
10	SSc_10	Nuclear fine speckled (AC-4)	Negative	Yes
11	SSc_11	Cytoplasmatic reticular (AC-21)	Negative	Yes
12	SSc_12	Negative (AC-0)	Anti-U11/U12 RNP	Yes
13	SSc_13	Negative (AC-0)	Negative	Yes
14	SSc_14	Nuclear speckled (Atypical)	Negative	Yes
15	SSc_15	Homogeneous nucleolar (AC-8)	Negative	Yes
16	SSc_16	Punctate nucleolar (AC-10)	Negative	Yes
17	SSc_17	Negative (AC-0)	Negative	Yes
18	SSc_18	Negative (AC-0)	Negative	Yes
19	SSc_19	Negative (AC-0)	Negative	Yes
20	SSc_20	Negative (AC-0)	Negative	Yes
21	SSc_21	Nuclear homogeneous (AC-1)	Negative	Yes
22	SSc_22	Homogeneous nucleolar (AC-8)	Negative	Yes
23	SSc_23	Negative (AC-0)	Anti-U4/U6 snRNP	Yes
24	SSc_24	Nuclear fine speckled (AC-4)	Negative	Yes
25	SSc_25	Negative (AC-0)	Negative	Yes
26	SSc_26	Negative (AC-0)	Negative	Yes
27	SSc_27	Cytoplasmatic reticular (AC-21)	Negative	Yes
28	SSc_29	Negative (AC-0)	Anti-U11/U12 RNP	Yes

Supplementary Table S1 (continued)

No.	Sample	IIF pattern	RNA IP result	RIP-Seq tested
29	SSc_30	Clumpy nucleolar (AC-9)	Anti-Fibrillarin	Yes
30	SSc_31	Nuclear fine speckled (AC-4)	Negative	Yes
31	SSc_32	Negative (AC-0)	Negative	Yes
32	SSc_33	Nuclear fine speckled (AC-4)	Anti-U11/U12 RNP	Yes
33	SSc_34	Negative (AC-0)	Negative	Yes
34	SSc_35	Negative (AC-0)	Negative	Yes
35	SSc_36	Negative (AC-0)	Negative	Yes
36	SSc_37	Nuclear speckled (Aypical)	Negative	Yes
37	SSc_38	Few nuclear dots (AC-7)	Negative	Yes
38	SSc_39	Negative (AC-0)	Anti-U11/U12 RNP	Yes
39	SSc_40	Negative (AC-0)	Negative	Yes
40	SSc_41	Negative (AC-0)	Anti-U11/U12 RNP	Yes
41	SSc_42	Homogeneous nucleolar (AC-8)	Anti-Th/To	Yes
42	SSc_43	Homogeneous nucleolar (AC-8)	Negative	Yes
43	SSc_44	Nuclear speckled (Atypical)	Anti-U11/U12 RNP	Yes
44	SSc_45	Negative (AC-0)	Negative	Yes
45	SSc_46	Clumpy nucleolar (AC-9)	Anti-U8/U13 RNP	Yes
46	SSc_47	Negative (AC-0)	Anti-U11/U12 RNP	Yes
47	SSc_48	Clumpy nucleolar (AC-9)	Anti-Ro60	Yes
48	SSc_49	Negative (AC-0)	Negative	Yes
49	SSc_50	Negative (AC-0)	Negative	Yes
50	SSc_51	Homogeneous nucleolar (AC-8)	Anti-Th/To	Yes
51	SSc_52	Negative (AC-0)	Anti-Ro60	Yes
52	SSc_53	Homogeneous nucleolar (AC-8)	Anti-Th/To	Yes
53	SSc_54	Nuclear fine speckled (AC-4)	Negative	Yes
54	SSc_55	Negative (AC-0)	Negative	Yes
55	SSc_56	Homogeneous nucleolar (AC-8)	Negative	Yes
56	SSc_57	Negative (AC-0)	Negative	Yes
57	SSc_58	Homogeneous nucleolar (AC-8)	Anti-Th/To + Anti-Ro60	Yes

Supplementary Table S1 (continued)

No.	Sample	IIF pattern	RNA IP result	RIP-Seq tested
58	SSc_66	Homogeneous nucleolar (AC-8)	Anti-U11/U12 RNP	No
59	SSc_68	Negative (AC-0)	Negative	No
60	SSc_69	Negative (AC-0)	Negative	No
61	SSc_70	Clumpy nucleolar (AC-9)	Negative	No
62	SSc_71	Homogeneous nucleolar (AC-8)	Anti-Th/To	No
63	SSc_72	Homogeneous nucleolar (AC-8)	Negative	No
64	SSc_73	Negative (AC-0)	Negative	No
65	SSc_74	Cytoplasmatic (Atypical)	Negative	No
66	SSc_75	Negative (AC-0)	Negative	No
67	SSc_76	Negative (AC-0)	Negative	No
68	SSc_77	Negative (AC-0)	Negative	No

Supplementary Table S2: Two-hundred selected RNA candidates based on max²/median ratio (cut-off ≥21,778). RNA molecules typically detected by traditional RNA IP are marked in blue.

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
SNORD3B-1	ENSG00000265185.6	C/D box snoRNA	4.44E+09	U3 RNA
RNY4	ENSG00000252316.1	Y RNA	1.34E+09	Y4 RNA
RNU11	ENSG00000274978.1	snRNA	7.03E+08	U11 RNA
VTRNA1-2	ENSG00000202111.1	Vault RNA	1.49E+08	-
RNY1	ENSG00000201098.1	Y RNA	1.43E+08	Y1 RNA
RNY5	ENSG00000286171.1	Y RNA	7.62E+07	Y5 RNA
SNORD14C	ENSG00000202252.1	C/D box snoRNA	5.51E+07	-
RMRP	ENSG00000269900.3	lncRNA	3.80E+07	7-2 RNA
VTRNA2-1	ENSG00000270123.4	Vault RNA	3.57E+07	-
SNORD49A	ENSG00000277370.1	C/D box snoRNA	1.84E+07	-
SNORD30	ENSG00000277846.1	C/D box snoRNA	1.46E+07	-
SNHG5	ENSG00000203875.13	SNRHG	1.41E+07	-
SNORD45A	ENSG00000207241.1	C/D box snoRNA	1.07E+07	-
RPL13	ENSG00000167526.14	Protein coding	1.03E+07	-
GAS5	ENSG00000234741.8	SNRHG	1.02E+07	-
SNORD46	ENSG00000200913.1	C/D box snoRNA	7.92E+06	-
SNORD100	ENSG00000221500.1	C/D box snoRNA	7.90E+06	-
SNORD63	ENSG00000206989.1	C/D box snoRNA	7.72E+06	-
RNY3	ENSG00000202354.1	Y RNA	7.69E+06	Y3 RNA
CSKMT	ENSG00000199352.1	rRNA	6.31E+06	-
SNORD2	ENSG00000238942.1	C/D box snoRNA	4.78E+06	-
7SK	ENSG00000202198.1	miscRNA	4.70E+06	-
VTRNA1-1	ENSG00000199990.1	Vault RNA	4.34E+06	-
SNORD21	ENSG00000206680.1	C/D box snoRNA	4.02E+06	-
SNHG29	ENSG00000175061.18	SNRHG	3.36E+06	-
SNORD57	ENSG00000226572.1	C/D box snoRNA	3.10E+06	-
SNORD101	ENSG00000206754.1	C/D box snoRNA	2.87E+06	-
SNORD12	ENSG00000212304.1	C/D box snoRNA	2.72E+06	-
SNORD27	ENSG00000275996.1	C/D box snoRNA	2.69E+06	-
SNHG32	ENSG00000204387.14	SNRHG	2.64E+06	-
RNA5SP376	ENSG00000212251.1	rRNA	2.57E+06	-

Supplementary Table S2 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
SNORD114-1	ENSG00000199575.1	C/D box snoRNA	2.46E+06	-
RPPH1	ENSG00000259001.3	lncRNA	2.08E+06	8-2 RNA
SNORD18C	ENSG00000199574.1	C/D box snoRNA	1.92E+06	-
SNORD91B	ENSG00000275084.4	C/D box snoRNA	1.78E+06	-
SNORD18A	ENSG00000200623.1	C/D box snoRNA	1.62E+06	-
SNORD25	ENSG00000275043.1	C/D box snoRNA	1.55E+06	-
SNORD99	ENSG00000221539.1	C/D box snoRNA	1.38E+06	-
SNORD92	ENSG00000264994.1	C/D box snoRNA	1.37E+06	-
SNORD12B	ENSG00000222365.1	C/D box snoRNA	1.23E+06	-
SNORD3A	ENSG00000263934.5	C/D box snoRNA	1.22E+06	-
ENSG10010134635.1	ENSG00000222185.1	C/D box snoRNA	1.17E+06	-
SNORD63B	ENSG00000222937.1	C/D box snoRNA	1.15E+06	-
SNORD111B	ENSG00000221514.1	C/D box snoRNA	1.00E+06	-
SNORD45C	ENSG00000206620.1	C/D box snoRNA	9.46E+05	-
SNORD114-12	ENSG00000202270.1	C/D box snoRNA	9.37E+05	-
SNORD86	ENSG00000212498.1	C/D box snoRNA	9.00E+05	-
SNORD95	ENSG00000264549.1	C/D box snoRNA	8.13E+05	-
SNORD1B	ENSG00000199961.1	C/D box snoRNA	7.45E+05	-
SNORD6	ENSG00000202314.1	C/D box snoRNA	7.40E+05	-
SNORD82	ENSG00000202400.1	C/D box snoRNA	6.55E+05	-
SNORD43	ENSG00000263764.1	C/D box snoRNA	6.22E+05	-
SNORD58A	ENSG00000206602.1	C/D box snoRNA	5.95E+05	-
MALAT1	ENSG00000251562.8	lncRNA	5.95E+05	-
SNORD24	ENSG00000206611.1	C/D box snoRNA	5.93E+05	-
SNORD114-23	ENSG00000200406.1	C/D box snoRNA	5.76E+05	-
SNORA3B	ENSG00000212607.1	H/ACA box snoRNA	5.72E+05	-
SNORD104	ENSG00000199753.1	C/D box snoRNA	5.58E+05	-
RNU5A-1	ENSG00000199568.1	snRNA	5.47E+05	U5 RNA
RNA5SP289	ENSG00000199202.1	rRNA	5.42E+05	-
SNORD110	ENSG00000221116.1	C/D box snoRNA	5.13E+05	-
SNORD60	ENSG00000206630.1	C/D box snoRNA	4.61E+05	-
SNORD114-22	ENSG00000202293.1	C/D box snoRNA	4.58E+05	-

Supplementary Table S2 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
SNORD102	ENSG00000207500.1	C/D box snoRNA	4.07E+05	-
SNORD52	ENSG00000201754.1	C/D box snoRNA	4.04E+05	-
RNU5B-1	ENSG00000200156.1	snRNA	3.99E+05	U5 RNA
SNORD114-25	ENSG00000200612.1	C/D box snoRNA	3.98E+05	-
SNORA16A	ENSG00000280498.1	H/ACA box snoRNA	3.95E+05	-
SNORD114-26	ENSG00000200413.1	C/D box snoRNA	3.92E+05	-
SNORD114-3	ENSG00000201839.1	C/D box snoRNA	3.84E+05	-
MT-TE	ENSG00000210194.1	Mitochondrial tRNA	3.79E+05	-
SNORD126	ENSG00000238344.1	C/D box snoRNA	3.74E+05	-
RNU5F-1	ENSG00000199377.1	snRNA	3.54E+05	U5 RNA
SNORD33	ENSG00000199631.1	C/D box snoRNA	3.42E+05	-
SNORD37	ENSG00000206775.1	C/D box snoRNA	3.35E+05	-
SNORD113-8	ENSG00000200367.1	C/D box snoRNA	3.21E+05	-
SNORD127	ENSG00000239043.1	C/D box snoRNA	3.17E+05	-
SNORD69	ENSG00000212452.1	C/D box snoRNA	3.00E+05	-
SNORD111	ENSG00000221066.1	C/D box snoRNA	2.97E+05	-
SNORD105B	ENSG00000238531.1	C/D box snoRNA	2.81E+05	-
SNORD42A	ENSG00000238649.1	C/D box snoRNA	2.73E+05	-
G3BP1	ENSG00000145907.16	Protein coding	2.70E+05	-
RPL4	ENSG00000174444.15	Protein coding	2.46E+05	-
RNU6-5P	ENSG00000206965.1	snRNA	2.45E+05	U6 RNA
ENSG00000280494	ENSG00000280494.2	miRNA	2.43E+05	-
SNORD12C	ENSG00000209042.1	C/D box snoRNA	2.25E+05	-
SNORA61	ENSG00000278274.1	H/ACA box snoRNA	2.24E+05	-
SNORA20	ENSG00000207392.1	H/ACA box snoRNA	2.23E+05	-
SNORD34	ENSG00000202503.1	C/D box snoRNA	2.08E+05	-
RPL13A	ENSG00000142541.18	Protein coding	1.90E+05	-
SNORD58C	ENSG00000202093.1	C/D box snoRNA	1.89E+05	-
7SK	ENSG00000271394.1	miscRNA	1.78E+05	-
SNORD65	ENSG00000277512.1	C/D box snoRNA	1.77E+05	-
SNORD72	ENSG00000212296.1	C/D box snoRNA	1.74E+05	-
SNORD28	ENSG00000274544.1	C/D box snoRNA	1.55E+05	-

Supplementary Table S2 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
SNORA3A	ENSG00000200983.1	H/ACA box snoRNA	1.54E+05	-
SNORD66	ENSG00000212158.1	C/D box snoRNA	1.50E+05	-
SNORA18	ENSG00000207145.1	H/ACA box snoRNA	1.42E+05	-
SNORD83A	ENSG00000209482.1	C/D box snoRNA	1.42E+05	-
SNORD61	ENSG00000206979.1	C/D box snoRNA	1.35E+05	-
RNU12	ENSG00000276027.1	snRNA	1.35E+05	U12 RNA
SNORD114-11	ENSG00000200608.1	C/D box snoRNA	1.35E+05	-
SNORD49B	ENSG00000277108.1	C/D box snoRNA	1.29E+05	-
SNORD19C	ENSG00000222345.1	C/D box snoRNA	1.18E+05	-
SNORD41	ENSG00000209702.1	C/D box snoRNA	1.18E+05	-
SNORD53B	ENSG00000265706.1	C/D box snoRNA	1.16E+05	-
SNORD71	ENSG00000223224.1	C/D box snoRNA	1.15E+05	-
SNORA2C	ENSG00000221491.2	H/ACA box snoRNA	1.12E+05	-
SNORD114-14	ENSG00000199593.1	C/D box snoRNA	1.07E+05	-
SNORD38A	ENSG00000202031.1	C/D box snoRNA	1.07E+05	-
SNORD54	ENSG00000238650.1	C/D box snoRNA	1.07E+05	-
SNORA66	ENSG00000207523.1	H/ACA box snoRNA	1.06E+05	-
SNORD51	ENSG00000207047.2	C/D box snoRNA	1.05E+05	-
SCARNA22	ENSG00000249784.1	scaRNA	1.01E+05	-
SNORD19B	ENSG00000238862.1	C/D box snoRNA	9.98E+04	-
SNORD4A	ENSG00000238578.1	H/ACA box snoRNA	9.83E+04	-
SNORD59A	ENSG00000207031.1	C/D box snoRNA	9.81E+04	-
SNORD105	ENSG00000209645.1	C/D box snoRNA	9.71E+04	-
SNORD114-9	ENSG00000201240.1	C/D box snoRNA	9.53E+04	-
SNORD114-17	ENSG00000201569.1	C/D box snoRNA	9.28E+04	-
TTC3	ENSG00000182670.13	Protein coding	8.87E+04	-
SNORD50B	ENSG00000275072.1	C/D box snoRNA	8.84E+04	-
ENSG10010137917.1	ENSG00000280554.1	C/D box snoRNA	8.74E+04	-
SNORD5	ENSG00000239195.1	C/D box snoRNA	8.43E+04	-
MT-CO1	ENSG00000198804.2	Protein coding	8.40E+04	-
SNORD3B-2	ENSG00000262074.7	C/D box snoRNA	8.36E+04	-
SNORD36C	ENSG00000252542.1	C/D box snoRNA	8.28E+04	-

Supplementary Table S2 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
SNORD113-6	ENSG00000200215.3	C/D box snoRNA	7.96E+04	-
CH17-3B23.3	ENSG00000287979.1	lncRNA	7.84E+04	-
SNORD91A	ENSG00000212163.6	C/D box snoRNA	7.67E+04	-
SNORD38B	ENSG00000281859.1	C/D box snoRNA	7.36E+04	-
SNORD87	ENSG00000254341.2	C/D box snoRNA	7.19E+04	-
MT-TV	ENSG00000210077.1	Mitochondrial tRNA	7.02E+04	-
RNU6ATAC	ENSG00000221676.1	snRNA	7.00E+04	U6 RNA
SNORD1A	ENSG00000278261.1	C/D box snoRNA	6.54E+04	-
SNORD84	ENSG00000265236.1	C/D box snoRNA	6.23E+04	-
SNORD114-13	ENSG00000201247.1	C/D box snoRNA	6.18E+04	-
MT-TM	ENSG00000210112.1	Mitochondrial tRNA	6.05E+04	-
SNORD11B	ENSG00000271852.1	C/D box snoRNA	5.97E+04	-
RP11-596C23.6	ENSG00000282885.2	lncRNA	5.93E+04	-
SNORD14B	ENSG00000201403.1	C/D box snoRNA	5.86E+04	-
SNORD45B	ENSG00000201487.1	C/D box snoRNA	5.86E+04	-
SCARNA15	ENSG00000277864.1	scaRNA	5.76E+04	-
FBXL20	ENSG00000108306.13	Protein coding	5.71E+04	-
SNORD7	ENSG00000207297.1	C/D box snoRNA	5.49E+04	-
SNORD114-28	ENSG00000200480.1	C/D box snoRNA	5.39E+04	-
SNORD113-7	ENSG00000200632.1	C/D box snoRNA	5.31E+04	-
LINC02739	ENSG00000255008.3	lncRNA	5.28E+04	-
RANBP2	ENSG00000153201.16	Protein coding	5.24E+04	-
RN7SL1	ENSG00000276168.1	miscRNA	5.15E+04	-
PPP1CB	ENSG00000213639.10	Protein coding	5.04E+04	-
SNORD35A	ENSG00000200259.1	C/D box snoRNA	4.88E+04	-
SNORD94	ENSG00000208772.1	C/D box snoRNA	4.87E+04	-
SNORD70	ENSG00000212534.1	C/D box snoRNA	4.79E+04	-
SNORD1C	ENSG00000274091.1	C/D box snoRNA	4.66E+04	-
SNORA58B	ENSG00000201129.1	H/ACA box snoRNA	4.66E+04	-
RANBP1	ENSG00000099901.17	Protein coding	4.65E+04	-
SNORA46	ENSG00000207493.1	H/ACA box snoRNA	4.51E+04	-
SNORD73A	ENSG00000208797.1	C/D box snoRNA	4.10E+04	-

Supplementary Table S2 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
KMT2A	ENSG00000118058.23	Protein coding	4.08E+04	-
SNORD14D	ENSG00000207118.1	C/D box snoRNA	4.03E+04	-
SNORD98	ENSG00000283551.1	C/D box snoRNA	4.01E+04	-
TUFM	ENSG00000178952.11	Protein coding	4.01E+04	-
AK2	ENSG00000004455.17	Protein coding	3.99E+04	-
SNORD4B	ENSG00000238597.1	C/D box snoRNA	3.92E+04	-
SNORD90	ENSG00000212447.1	C/D box snoRNA	3.86E+04	-
EP300	ENSG00000100393.14	Protein coding	3.82E+04	-
ATRX	ENSG00000085224.23	Protein coding	3.81E+04	-
GAPDH	ENSG00000111640.15	Protein coding	3.68E+04	-
AC084082.3	ENSG00000253190.4	lncRNA	3.59E+04	-
RPS18	ENSG00000231500.7	Protein coding	3.57E+04	-
SNORD83B	ENSG00000209480.1	C/D box snoRNA	3.30E+04	-
SNORD16	ENSG00000199673.1	C/D box snoRNA	3.22E+04	-
ZMYND8	ENSG00000101040.19	Protein coding	3.21E+04	-
CSKMT	ENSG00000214756.8	Protein coding	3.12E+04	-
IST1	ENSG00000182149.21	Protein coding	3.08E+04	-
NOP56	ENSG00000101361.17	Protein coding	3.08E+04	-
SNORD93	ENSG00000221740.1	C/D box snoRNA	3.02E+04	-
ENSG10010138968.2	ENSG00000252787.2	C/D box snoRNA	2.91E+04	-
SNORD114-10	ENSG00000200279.1	C/D box snoRNA	2.85E+04	-
TPT1	ENSG00000133112.17	Protein coding	2.81E+04	-
SNORD89	ENSG00000212283.1	C/D box snoRNA	2.76E+04	-
SNORD62A	ENSG00000235284.1	C/D box snoRNA	2.68E+04	-
SNORD11	ENSG00000238317.2	C/D box snoRNA	2.66E+04	-
SNHG17	ENSG00000196756.13	SNRHG	2.59E+04	-
MT-TH	ENSG00000210176.1	Mitochondrial tRNA	2.59E+04	-
RNA5S9	ENSG00000201321.1	rRNA	2.53E+04	-
SNORD113-5	ENSG00000272474.1	C/D box snoRNA	2.52E+04	-
LPIN1	ENSG00000134324.12	Protein coding	2.51E+04	-
HSPA8	ENSG00000109971.14	Protein coding	2.51E+04	-
RNU4-1	ENSG00000200795.1	snRNA	2.48E+04	U4 RNA

Supplementary Table S2 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Max²/median	Known immunoprecipitated RNA molecule
SNORD118	ENSG00000200463.1	C/D box snoRNA	2.44E+04	-
GAN	ENSG00000261609.8	Protein coding	2.40E+04	-
SNORD14A	ENSG00000272034.1	C/D box snoRNA	2.39E+04	-
ANK1	ENSG00000029534.21	Protein coding	2.28E+04	-
SNORD17	ENSG00000212232.1	C/D box snoRNA	2.28E+04	-
H2AC18	ENSG00000203812.2	Protein coding	2.23E+04	-
PI4KAP2	ENSG00000183506.17	Transcribed unitary pseudogene	2.21E+04	-
SNORD14E	ENSG00000200879.1	C/D box snoRNA	2.18E+04	-
NUP214	ENSG00000126883.17	Protein coding	2.18E+04	-

snoRNA: small nucleolar RNA; snRNA: small nuclear RNA; lncRNA: long non coding RNA; SNRHG: small nucleolar RNA host gene; rRNA: ribosomal RNA; miscRNA: miscellaneous RNA; tRNA: transfer RNA; miRNA: micro RNA.

Supplementary Table S3: Distribution of RNA molecules types among candidate RNA molecules before and after re-analysis of protein coding RNA molecules.

RNA molecule type	Number of RNA molecules before re-analysis (%)	Number of RNA molecules after re-analysis (%)
C/D box snoRNA	120 (60.0)	122 (61.9)
Protein coding	28 (14.0)	20 (10.2)
H/ACA box snoRNA	10 (5.0)	12 (6.1)
snRNA	8 (4.0)	8 (4.1)
SNHG	5 (2.5)	5 (2.5)
lncRNA	7 (3.5)	7 (3.6)
rRNA	4 (2.0)	4 (2.0)
Mitochondrial tRNA	4 (2.0)	4 (2.0)
Y RNA	4 (2.0)	4 (2.0)
Vault RNA	3 (1.5)	3 (1.5)
miscRNA	3 (1.5)	3 (1.5)
scaRNA	2 (1.0)	2 (1.0)
miRNA	1 (0.5)	1 (0.5)
Transcribed unitary pseudogene	1 (0.5)	1 (0.5)
Not identified RNA sequence	0 (0.0)	1 (0.5)
Total	200 (100.0)	197 (100.0)

snoRNA: small nucleolar RNA; snRNA: small nuclear RNA; lncRNA: long non coding RNA; SNHG: small nucleolar RNA host gene; rRNA: ribosomal RNA; miscRNA: miscellaneous RNA; tRNA: transfer RNA; miRNA: micro RNA.

Supplementary Table S4: protein coding RNA molecules that were selected by the max²/median ratio due to immunoprecipitation of snoRNAs codified by their introns. These protein coding RNAs were substituted by the corresponding snoRNA in the final candidate list. snoRNA molecules marked in blue were already present in the candidate list and were not included to avoid redundancy.

Assigned protein coding RNA molecules		Max²/median	Mapped sample	Assigned RNA molecules after re-analysis	
Name	ENSEMBL ID			Name	ENSEMBL ID
RPL13	ENSG00000167526.14	1.03E+07	SSc_30	SNORD68	ENSG00000200084
RPL4	ENSG00000174444.15	2.46E+05	SSc_30	SNORD18A	ENSG00000200623
RPL13A	ENSG00000142541.18	1.90E+05	SSc_58	SNORD35A	ENSG00000200259
RANBP1	ENSG00000099901.17	4.65E+04	SSc_44	SNORA77B	ENSG00000264346
CSKMT	ENSG00000214756.8	3.12E+04	SSc_30	SNORA57	ENSG00000206597
NOP56	ENSG00000101361.17	3.08E+04	SSc_48	SNORD57	ENSG00000226572
HSPA8	ENSG00000109971.14	2.51E+04	SSc_48	SNORD14C	ENSG00000202252

Supplementary Table S5: candidate RNA molecules assignation to different modules by WGCNA.

RNA molecule name	ENSEMBL ID	RNA molecule type	Module
EP300	ENSG00000100393.14	Protein coding	A
FBXL20	ENSG00000108306.13	Protein coding	A
KMT2A	ENSG00000118058.23	Protein coding	A
G3BP1	ENSG00000145907.16	Protein coding	A
IST1	ENSG00000182149.21	Protein coding	A
TTC3	ENSG00000182670.13	Protein coding	A
PI4KAP2	ENSG00000183506.17	Transcribed unitary pseudogene	A
MT-CO1	ENSG00000198804.2	Protein coding	A
RNU5F-1	ENSG00000199377.1	snRNA	A
RNU5A-1	ENSG00000199568.1	snRNA	A
RNU5B-1	ENSG00000200156.1	snRNA	A
H2AC18	ENSG00000203812.2	Protein coding	A
MT-TV	ENSG00000210077.1	Mitochondrial tRNA	A
MT-TM	ENSG00000210112.1	Mitochondrial tRNA	A
RPS18	ENSG00000231500.7	Protein coding	A
AC084082.3	ENSG00000253190.4	lncRNA	A
GAN	ENSG00000261609.8	Protein coding	A
SNORD3B-2	ENSG00000262074.7	C/D box snoRNA	A
RMRP	ENSG00000269900.3	lncRNA	A
7SK	ENSG00000271394.1	miscRNA	A
RNU11	ENSG00000274978.1	snRNA	A
RNU12	ENSG00000276027.1	snRNA	A
RN7SL1	ENSG00000276168.1	miscRNA	A
RP11-596C23.6	ENSG00000282885.2	lncRNA	A
CH17-3B23.3	ENSG00000287979.1	lncRNA	A
SNORD35A	ENSG00000200259.1	C/D box snoRNA	B
RNU4-1	ENSG00000200795.1	snRNA	B
SNORD114-13	ENSG00000201247.1	C/D box snoRNA	B
SNORD52	ENSG00000201754.1	C/D box snoRNA	B
SNORD82	ENSG00000202400.1	C/D box snoRNA	B
SNORD51	ENSG00000207047.2	C/D box snoRNA	B
SNORA18	ENSG00000207145.1	H/ACA box snoRNA	B
SNORD7	ENSG00000207297.1	C/D box snoRNA	B

Supplementary Table S5 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Module
SNORD83B	ENSG00000209480.1	C/D box snoRNA	B
MT-TH	ENSG00000210176.1	Mitochondrial tRNA	B
SNORD91A	ENSG00000212163.6	C/D box snoRNA	B
SNORD100	ENSG00000221500.1	C/D box snoRNA	B
SNORD4A	ENSG00000238578.1	C/D box snoRNA	B
SNORD4B	ENSG00000238597.1	C/D box snoRNA	B
SNORD19B	ENSG00000238862.1	C/D box snoRNA	B
SNORD2	ENSG00000238942.1	C/D box snoRNA	B
ENSG10010138968.2	ENSG00000252787.2	C/D box snoRNA	B
SNORD43	ENSG00000263764.1	C/D box snoRNA	B
SNORD113-5	ENSG00000272474.1	C/D box snoRNA	B
SNORD30	ENSG00000277846.1	C/D box snoRNA	B
SNORD98	ENSG00000283551.1	C/D box snoRNA	B
ZMYND8	ENSG00000101040.19	Protein coding	C
LPIN1	ENSG00000134324.12	Protein coding	C
RANBP2	ENSG00000153201.16	Protein coding	C
RNA5SP289	ENSG00000199202.1	rRNA	C
RNA5S1	ENSG00000199352.1	rRNA	C
SNORD18C	ENSG00000199574.1	C/D box snoRNA	C
VTRNA1-1	ENSG00000199990.1	Vault RNA	C
SNORA3A	ENSG00000200983.1	H/ACA box snoRNA	C
RNA5S9	ENSG00000201321.1	rRNA	C
VTRNA1-2	ENSG00000202111.1	Vault RNA	C
7SK	ENSG00000202198.1	miscRNA	C
RNU6-5P	ENSG00000206965.1	snRNA	C
SNORD66	ENSG00000212158.1	C/D box snoRNA	C
RNA5SP376	ENSG00000212251.1	rRNA	C
PPP1CB	ENSG00000213639.10	Protein coding	C
RNU6ATAC	ENSG00000221676.1	snRNA	C
SNORD12B	ENSG00000222365.1	C/D box snoRNA	C
SNORD36C	ENSG00000252542.1	C/D box snoRNA	C
ENSG00000253389	ENSG00000253389	Not identified RNA molecule	C
LINC02739	ENSG00000255008.3	lncRNA	C

Supplementary Table S5 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Module
RPPH1	ENSG00000259001.3	lncRNA	C
SNORD84	ENSG00000265236.1	C/D box snoRNA	C
VTRNA2-1	ENSG00000270123.4	Vault RNA	C
SNORD11B	ENSG00000271852.1	C/D box snoRNA	C
SNORD1A	ENSG00000278261.1	C/D box snoRNA	C
ENSG00000280494	ENSG00000280494.2	miRNA	C
ENSG10010137917.1	ENSG00000280554.1	C/D box snoRNA	C
SNORD118	ENSG00000200463.1	C/D box snoRNA	D
RNY1	ENSG00000201098.1	Y RNA	D
RNY3	ENSG00000202354.1	Y RNA	D
SNORD72	ENSG00000212296.1	C/D box snoRNA	D
SNORD93	ENSG00000221740.1	C/D box snoRNA	D
SNORD65	ENSG00000277512.1	C/D box snoRNA	D
RNY5	ENSG00000286171.1	Y RNA	D
SNORD18A	ENSG00000200623.1	C/D box snoRNA	E
SNORD14B	ENSG00000201403.1	C/D box snoRNA	E
SNORD45B	ENSG00000201487.1	C/D box snoRNA	E
SNORD58C	ENSG00000202093.1	C/D box snoRNA	E
SNORD60	ENSG00000206630.1	C/D box snoRNA	E
SNORD94	ENSG00000208772.1	C/D box snoRNA	E
SNORD70	ENSG00000212534.1	C/D box snoRNA	E
SNORD71	ENSG00000223224.1	C/D box snoRNA	E
SNORD62A	ENSG00000235284.1	C/D box snoRNA	E
SNORD127	ENSG00000239043.1	C/D box snoRNA	E
SNORD5	ENSG00000239195.1	C/D box snoRNA	E
SNORD53B	ENSG00000265706.1	C/D box snoRNA	E
AK2	ENSG00000004455.17	Protein coding	F
ATRX	ENSG00000085224.23	Protein coding	F
NUP214	ENSG00000126883.17	Protein coding	F
TUFM	ENSG00000178952.11	Protein coding	F
SNORD104	ENSG00000199753.1	C/D box snoRNA	F
SNORD1B	ENSG00000199961.1	C/D box snoRNA	F
SNORD114-11	ENSG00000200608.1	C/D box snoRNA	F

Supplementary Table S5 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Module
SNORA58B	ENSG00000201129.1	H/ACA box snoRNA	F
SNORD14C	ENSG00000202252.1	C/D box snoRNA	F
SNORD34	ENSG00000202503.1	C/D box snoRNA	F
SNORD37	ENSG00000206775.1	C/D box snoRNA	F
SNORD61	ENSG00000206979.1	C/D box snoRNA	F
SNORD63	ENSG00000206989.1	C/D box snoRNA	F
SNORA66	ENSG00000207523.1	H/ACA box snoRNA	F
MT-TE	ENSG00000210194.1	Mitochondrial tRNA	F
SNORA2C	ENSG00000221491.2	H/ACA box snoRNA	F
SNORD111B	ENSG00000221514.1	C/D box snoRNA	F
SNORD99	ENSG00000221539.1	C/D box snoRNA	F
ENSG10010134635.1	ENSG00000222185.1	C/D box snoRNA	F
SNORD63B	ENSG00000222937.1	C/D box snoRNA	F
SNORD57	ENSG00000226572	C/D box snoRNA	F
SNORD57	ENSG00000226572.1	C/D box snoRNA	F
SNORD42A	ENSG00000238649.1	C/D box snoRNA	F
SNORD54	ENSG00000238650.1	C/D box snoRNA	F
MALAT1	ENSG00000251562.8	lncRNA	F
RNY4	ENSG00000252316.1	Y RNA	F
SNORD92	ENSG00000264994.1	C/D box snoRNA	F
SNORD1C	ENSG00000274091.1	C/D box snoRNA	F
SNORD28	ENSG00000274544.1	C/D box snoRNA	F
SNORD91B	ENSG00000275084.4	C/D box snoRNA	F
SNORD38B	ENSG00000281859.1	C/D box snoRNA	F
SNHG17	ENSG00000196756.13	SNRHG	G
SNORD114-1	ENSG00000199575.1	C/D box snoRNA	G
SNORD16	ENSG00000199673.1	C/D box snoRNA	G
SNORD68	ENSG00000200084	C/D box snoRNA	G
SNORD113-6	ENSG00000200215.3	C/D box snoRNA	G
SNORD113-8	ENSG00000200367.1	C/D box snoRNA	G
SNORD114-23	ENSG00000200406.1	C/D box snoRNA	G
SNORD114-28	ENSG00000200480.1	C/D box snoRNA	G
SNORD114-25	ENSG00000200612.1	C/D box snoRNA	G

Supplementary Table S5 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Module
SNORD114-17	ENSG00000201569.1	C/D box snoRNA	G
SNORD114-12	ENSG00000202270.1	C/D box snoRNA	G
SNORD114-22	ENSG00000202293.1	C/D box snoRNA	G
SNORA57	ENSG00000206597	H/ACA box snoRNA	G
SNORD24	ENSG00000206611.1	C/D box snoRNA	G
SNORD45C	ENSG00000206620.1	C/D box snoRNA	G
SNORD105	ENSG00000209645.1	C/D box snoRNA	G
SNORD90	ENSG00000212447.1	C/D box snoRNA	G
SNORD86	ENSG00000212498.1	C/D box snoRNA	G
SNORD3A	ENSG00000263934.5	C/D box snoRNA	G
SNORD95	ENSG00000264549.1	C/D box snoRNA	G
SNORD3B-1	ENSG00000265185.6	C/D box snoRNA	G
GAPDH	ENSG00000111640.15	Protein coding	H
TPT1	ENSG00000133112.17	Protein coding	H
SNHG29	ENSG00000175061.18	SNRHG	H
SNORD114-14	ENSG00000199593.1	C/D box snoRNA	H
SNORD33	ENSG00000199631.1	C/D box snoRNA	H
SNORD114-10	ENSG00000200279.1	C/D box snoRNA	H
SNORD114-26	ENSG00000200413.1	C/D box snoRNA	H
SNORD113-7	ENSG00000200632.1	C/D box snoRNA	H
SNORD14E	ENSG00000200879.1	C/D box snoRNA	H
SNORD46	ENSG00000200913.1	C/D box snoRNA	H
SNORD114-9	ENSG00000201240.1	C/D box snoRNA	H
SNORD114-3	ENSG00000201839.1	C/D box snoRNA	H
SNORD38A	ENSG00000202031.1	C/D box snoRNA	H
SNORD6	ENSG00000202314.1	C/D box snoRNA	H
SNHG5	ENSG00000203875.13	SNRHG	H
SNHG32	ENSG00000204387.14	SNRHG	H
SNORD58A	ENSG00000206602.1	C/D box snoRNA	H
SNORD21	ENSG00000206680.1	C/D box snoRNA	H
SNORD101	ENSG00000206754.1	C/D box snoRNA	H
SNORD59A	ENSG00000207031.1	C/D box snoRNA	H
SNORD14D	ENSG00000207118.1	C/D box snoRNA	H

Supplementary Table S5 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Module
SNORD102	ENSG00000207500.1	C/D box snoRNA	H
SNORD73A	ENSG00000208797.1	C/D box snoRNA	H
SNORD12C	ENSG00000209042.1	C/D box snoRNA	H
SNORD83A	ENSG00000209482.1	C/D box snoRNA	H
SNORD41	ENSG00000209702.1	C/D box snoRNA	H
SNORD17	ENSG00000212232.1	C/D box snoRNA	H
SNORD89	ENSG00000212283.1	C/D box snoRNA	H
SNORD12	ENSG00000212304.1	C/D box snoRNA	H
SNORD69	ENSG00000212452.1	C/D box snoRNA	H
SNORD111	ENSG00000221066.1	C/D box snoRNA	H
SNORD110	ENSG00000221116.1	C/D box snoRNA	H
SNORD19C	ENSG00000222345.1	C/D box snoRNA	H
GAS5	ENSG00000234741.8	SNRHG	H
SNORD11	ENSG00000238317.2	C/D box snoRNA	H
SNORD126	ENSG00000238344.1	C/D box snoRNA	H
SNORD105B	ENSG00000238531.1	C/D box snoRNA	H
SNORD87	ENSG00000254341.2	C/D box snoRNA	H
SNORD14A	ENSG00000272034.1	C/D box snoRNA	H
SNORD25	ENSG00000275043.1	C/D box snoRNA	H
SNORD50B	ENSG00000275072.1	C/D box snoRNA	H
SNORD27	ENSG00000275996.1	C/D box snoRNA	H
SNORD49B	ENSG00000277108.1	C/D box snoRNA	H
SNORD49A	ENSG00000277370.1	C/D box snoRNA	H
SNORD45A	ENSG00000207241.1	C/D box snoRNA	I
SNORA20	ENSG00000207392.1	H/ACA box snoRNA	I
SNORA46	ENSG00000207493.1	H/ACA box snoRNA	I
SNORA3B	ENSG00000212607.1	H/ACA box snoRNA	I
SCARNA22	ENSG00000249784.1	scaRNA	I
SNORA77B	ENSG00000264346	H/ACA box snoRNA	I
SCARNA15	ENSG00000277864.1	scaRNA	I
SNORA61	ENSG00000278274.1	H/ACA box snoRNA	I

Supplementary Table S5 (continued)

RNA molecule name	ENSEMBL ID	RNA molecule type	Module
SNORA16A	ENSG00000280498.1	H/ACA box snoRNA	I

snoRNA: small nucleolar RNA; snRNA: small nuclear RNA; lncRNA: long non coding RNA; SNRHG: small nucleolar RNA host gene; rRNA: ribosomal RNA; miscRNA: miscellaneous RNA; tRNA: transfer RNA; miRNA: micro RNA.

Supplementary Table S6: RNA molecules other than snoRNAs that were shown to be immunoprecipitated by RIP-Seq. The ribonucleoprotein of which is part, and the obtained max^2/median ratio is indicated for each RNA molecule.

RNA molecule	Ribonucleoprotein	Max^2/median
VTRNA1-2	Vault complex	1.49E+08
VTRNA2-1		3.57E+07
VTRNA1-1		4.34E+06
RNU5A-1	U5 snRNP (minor/major spliceosome)	5.47E+05
RNU5B-1		3.99E+05
RNU5F-1		3.54E+05
RNU4-1	U4 snRNP (major spliceosome)	2.48E+04
RNU6ATAC	U6atac snRNP (minor spliceosome)	7.00E+04
RNU6-5P	U6 snRNP (major spliceosome)	2.45E+05
MT-TE	Mitochondrial glutamyl-tRNA synthetase	3.79E+05
MT-TV	Mitochondrial valyl-tRNA synthetase	7.02E+04
MT-TM	Mitochondrial methionyl-tRNA synthetase	6.05E+04
MT-TH	Mitochondrial histidyl-tRNA synthetase	2.59E+04
RN7SL1	SRP	5.15E+04
7SK	7SK snRNP	4.70E+06
7SK		1.78E+05

snRNP: small nuclear ribonucleoprotein; tRNA: transfer RNA; SRP: signal recognition particle.

4.1.6. Anti-nuclear valosin-containing protein-like autoantibody is associated with calcinosis and higher risk of cancer in systemic sclerosis

Anti-nuclear valosin-containing protein-like autoantibody is associated with calcinosis and higher risk of cancer in systemic sclerosis. Perurena-Prieto J, Viñas-Giménez L, Sanz-Martínez MT, Selva-O'Callaghan A, Callejas-Moraga EL, Colobran R, Guillén-Del-Castillo A, Simeón-Aznar CP.

Rheumatology (Oxford). 2024 Aug 1;63(8):2278-2283. PMID: 37769243. DOI: 10.1093/rheumatology/kead520.

Pages: 187 - 200

4.2. Chapter 2

4.2.1. Objective

To develop a novel ELISA assay for anti-IFI16 autoantibody detection.

1- Testing of a cohort of SSc patients in which no specific autoantibodies were detected and analysis of the clinical manifestations associated with anti-IFI16 autoantibodies.
2- Testing of a cohort of SSc patients positive for anti-centromere and analysis of the clinical manifestations associated with the co-positivity of both autoantibodies.

4.2.2. Article

Prognostic value of anti-IFI16 autoantibodies in pulmonary arterial hypertension and mortality in patients with systemic sclerosis. Perurena-Prieto J, Callejas-Moraga EL, Sanz-Martínez MT, Colobran R, Guillén-Del-Castillo A, Simeón-Aznar CP.

Med Clin (Barc). 2024 Apr 26;162(8):370-377. PMID: 38302398. DOI: 10.1016/j.medcli.2023.11.020.

4.2.3. Previous considerations

Anti-IFI16 autoantibodies have been described as the third most prevalent autoantibodies in SSc patients, being detected in 20-30% of SSc patients [67, 336, 337]. Although anti-IFI16 are not specific for SSc, their detection can be of great utility, as they have been associated with lcSSc. Therefore, they are useful prognostic biomarkers in patients in which no SSc-specific autoantibodies are found [67, 336]. Furthermore, it has been shown that anti-centromere positive SSc patients who are also positive for anti-IFI16 autoantibodies had an increased risk of digital vascular events during the course of the disease, reinforcing the idea that these autoantibodies have to be considered as prognostic biomarkers in SSc [338]. However, currently, there is no commercial assay for the detection of anti-IFI16 autoantibodies. Moreover, as IFI16 expression of HEp-2 cells and other human cell lines usually employed for IP is low, anti-IFI16 autoantibodies cannot be detected either by IIF or IP, thus requiring a specific test for their detection [67]. In this line, we

decided that it would be very useful to develop a new ELISA method for detecting anti-IFI16 autoantibodies.

4.2.4. Summary of the results

SSc patients consulting during the 2013 to 2020 period in Vall d'Hebron University Hospital were included in the study. First, all recollected samples were tested by commercial assays. Fifty-eight patients from the cohort that were negative for all SSc-specific autoantibodies and 66 samples positive for anti-centromere autoantibodies by IIF were selected for further study. An in-house ELISA for detection of anti-IFI16 was developed and 52 healthy donors were tested in order to establish a positivity criterion or cut-off value. Overall, 29% of tested SSc patients were positive for anti-IFI16 autoantibodies. Specifically, 17.2% of patients negative for all SSc-specific autoantibodies and 39.4% of anti-centromere positive patients. Although we did not find an association between anti-IFI16 and the lcSSc subset due to the low number of patients with dcSSc included in the study, all anti-IFI16 positive patients were classified as lcSSc. However, anti-IFI16 autoantibody positivity showed a high association with PAH and correlated with an overall worse prognosis and higher risk of mortality.

Although not published in the paper, anti-IFI16 positivity was also analysed in conjunction with results obtained by RNA IP, RIP-Seq and protein IP (Table 7). However, no association between anti-IFI16 autoantibodies and other detected autoantibodies was found.

4.2.5. Prognostic value of anti-IFI16 autoantibodies in pulmonary arterial hypertension and mortality in patients with systemic sclerosis

Prognostic value of anti-IFI16 autoantibodies in pulmonary arterial hypertension and mortality in patients with systemic sclerosis. Perurena-Prieto J, Callejas-Moraga EL, Sanz-Martínez MT, Colobran R, Guillén-Del-Castillo A, Simeón-Aznar CP.

Med Clin (Barc). 2024 Apr 26;162(8):370-377. PMID: 38302398. DOI: 10.1016/j.medcli.2023.11.020.

Pages: 205 - 212

5. DISCUSSION

The presence of serum autoantibodies is a serological hallmark of SSc, as it has been reported that 95% of patients present ANAs by IIF [350]. However, in our cohort of SSc patients, a significant group presenting heterogeneous clinical manifestations, no specific SSc-related autoantibodies were detected by commercial assays. More specifically, from 307 SSc patients consulting during the 2013 to 2020 period in Vall d'Hebron University Hospital, 68 (22.1%) were negative for all tested SSc-specific autoantibodies. Nevertheless, 28 (41.2%) were ANA positive by IIF, indicating that probably this group of patients presented ANAs that were not identified by commercial assays based on detecting specific autoantibodies. Therefore, the main objective of this thesis was to develop novel autoantibody detection strategies to be able to identify specific autoantibodies in the patients that tested negative by commercial assays and to study if their presence could be associated with distinct clinical phenotypes.

5.1. Anti-ribonucleoprotein autoantibodies in SSc

Autoantibodies against ribonucleoprotein complexes were studied by traditional RNA IP and RIP-Seq. When the patients considered negative for specific autoantibodies by commercial assays were tested by traditional RNA IP, 23.5% resulted positive for already known SSc-specific autoantibodies. In particular, six patients were positive for anti-Th/To, nine for anti-U11/U12 snRNP autoantibodies and one for anti-U3 snoRNP. Most samples that were assayed by RNA IP were also tested by a newly developed RIP-Seq methodology. Following this approach, 197 RNA molecules were detected to be possible candidates of forming part of ribonucleoproteins targeted by SSc-related autoantibodies. All RNAs previously found to be immunoprecipitated by known SSc-related autoantibodies (7-2 RNA, 8-2 RNA, U3 RNA, U11 RNA and Y RNAs) were selected as candidate RNAs by RIP-Seq. In addition, a variety of other RNAs not known to be immunoprecipitated by any autoantibody were also detected by RIP-Seq,

5.1.1. Anti-Th/To autoantibodies

Among patients with a homogeneous nucleolar pattern (AC-8) by IIF who were negative for specific autoantibodies by commercial assays, 50.0% were found to be anti-Th/To positive. In the overall cohort (n=307), only 2 patients were classified as anti-Th/To positive by the commercial assay, indicating a sensitivity of 25.0% for detecting this autoantibody,

consistent with previous reports [95]. When considering the results of RNA IP and RIP-Seq together with the commercial test results, anti-Th/To autoantibodies were detected in 2.0% of the patients from the overall cohort, similar to what has been reported (0.8 - 4.6%) [166, 194–197].

Anti-Th/To autoantibodies recognise different subunits of two different essential ribonucleoproteins, RNase P and RNase MRP. Although these ribonucleoproteins share most of their protein components, they are associated with two different RNA molecules, 8-2 RNA and 7-2 RNA, respectively, and perform different functions (see section 1.7.5.2.). In this line, several protein components of RNase P and RNase MRP have been demonstrated to be the target anti-Th/To autoantibodies, including Rpp30, Pop5, Rpp14, Pop4, Rpp21. However, it has been reported that the majority of patients recognise either Rpp38, Pop1 or Rpp25 [165, 187–192]. Therefore, commercial assays using recombinant Pop1 [122, 193] and Rpp25 [192] proteins as antigens have been developed for the detection of anti-Th/To autoantibodies. These assays are specific enough, but their sensibility is not as high as expected because autoantibodies directed against other subunits of the RNase P or RNase MRP complexes are not recognised by these assays (Figure 12) [95, 192].

Although anti-Th/To autoantibodies are associated with lcSSc subsets, when anti-Th/To positive patients are compared to anti-centromere positive patients within the same subset, they more commonly have symptoms seen in dcSSc, such as ILD, as well as reduced survival [166, 194, 198]. As the presence of this autoantibody could indicate a different prognosis and follow-up of SSc patients and considering the low sensitivity of specific commercial assays for its detection, RNA IP analysis in all SSc patients who present an AC-8 IIF pattern and negative testing for SSc-specific autoantibodies by commercial methods should be performed.

5.1.2. Anti-U11/U12 snRNP autoantibodies

Anti-U11/U12 snRNP was the autoantibody most frequently detected by traditional RNA IP: 13.2% of patients who tested negative by commercial assays showed a positive result for this autoantibody. Considering that SSc-specific autoantibodies are generally mutually exclusive, although we have not tested the 307 SSc patient cohort, we can infer that the prevalence of anti-U11/U12 snRNP autoantibody in our cohort was 2.9%, similar to what has been reported (3.2 – 8.0%) [97, 314, 315]. Even if anti-U11/U12 snRNP autoantibodies are strongly linked to severe ILD [97, 314], currently, there is no available commercial test for their detection (see section 1.7.7.2.). Since this autoantibody does not produce a specific

IIF pattern, and considering it is a strong prognostic biomarker, our data indicates that performing RNA IP on all SSc samples testing negative for SSc-specific autoantibodies by commercial assays should be considered.

Interestingly, on RIP-Seq analysis, U12 RNA enrichment appeared to be nonspecific, as it was also detected in samples lacking anti-U11/U12 snRNP autoantibodies according to RNA IP. Anti-U11/U12 snRNP autoantibodies have been reported to recognise subunits of the U12-dependent (minor) spliceosome, responsible for the excision of the U12-type introns. The minor spliceosome is structurally similar to the U2-type (major) spliceosome. Both are composed of five snRNPs and numerous additional non-snRNP proteins. Each snRNP is constituted by Sm or LSm proteins and specific U snRNAs and proteins (see section 1.7.7.). Within the minor spliceosome, four of the five snRNAs are unique, specifically U11, U12, U4atac and U6atac, that are functionally and structurally related to the U1, U2, U4 and U6 snRNAs of the major spliceosome, respectively. Conversely, U5 snRNA is shared between the two spliceosomes [269, 270, 280, 281, 309].

Much of the specificity in the splicing reaction of both spliceosomes is accomplished by pairing with specific snRNAs. Although the overall assembly and catalytic steps of intron removal are very similar between the two spliceosomes [269, 270, 282, 309], there is a significant difference in the intron recognition step, which for minor introns is carried out by a preformed U11/U12 di-snRNP complex, while the major introns are recognised independently by individual U1 and U2 snRNPs (Figure 14) [272, 283–285]. This functional difference is reflected in the composition of the U11/U12 di-snRNP, which, in addition to the two unique U snRNAs, also contains 7 protein species that are unique to the minor spliceosome: RNPC3/65K, PDCD7/59K, 48K, 35K, ZCRB1/31K, 25K and ZMAT5/20K [286–288]. 59K, 48K, 35K, 25K and 20K are considered to be part of the U11 mono-snRNP, while it was first suggested that 65K and 31K constituted the U11 mono sn-RNP. In fact, the U11 snRNP associated 48K recognises the 5'ss together with U11 snRNA and interacts with 59K protein which is further engaged in an interaction with the N-terminal part of 65K. The C-terminal part of 65K binds U12 snRNA and also U6atac snRNA. Therefore, 65K, 59K and 48K form a chain of interactions connecting the U11 and U12 mono-snRNPs into a di-snRNP and are essential for the stability of the di-snRNP [359]. However, recent reports about cryo-EM reconstruction of U11 mono-snRNP suggest the presence of 65K also in this particle [360].

Typically, the band for U12 snRNA is not clearly visible in RNA IP assays of patients positive for anti-U11/U12 snRNP autoantibodies [97, 306]. In fact, U12 snRNA was not detected

through immunoprecipitation when this autoantibody was first described [306]. Nonetheless, it was expected that anti-U11/U12 snRNP autoantibodies also immunoprecipitated U12 snRNA, since U11/U12 snRNP forms part of the minor spliceosome as a stable complex, as previously discussed. Thus, the failure to detect U12 snRNA was first thought to be a consequence of its low abundance in cell extracts, as U11 snRNA is expected to be 5-fold more abundant than U12 snRNA in human cells. In fact, most U12 snRNP is found complexed with U11 snRNP, while a fraction of U11 snRNP, behaves as a mono-snRNP [283, 306]. Moreover, as U12 snRNA migrates in the U1 snRNA region in PAGE, it was also suggested that U12 snRNA's presence could be obscured by the highly abundant U1 snRNA. In the same report in which this autoantibody was first reported, it was noted that the main antigenic protein recognised by the positive sera was a 65kDa protein (RNPC3/65K). Moreover, using glycerol gradients to differentiate U11 mono-snRNP and U11/U12 di-snRNP, it was demonstrated that although the recognised 65K protein associated with both U11 and U12 snRNAs, it was more tightly associated with the U11 snRNA or at least not available for interaction in the U11/U12 di-snRNP complex [306]. Later on, it has been confirmed that 65K or RNPC3 is the major antigenic component recognised by anti-U11/U12 snRNP autoantibodies [97, 310].

Given the high sensitivity of RIP-Seq, our results indicate that the lack of specific U12 snRNA enrichment after immunoprecipitation cannot be explained by its low abundance in human cells. Therefore, although 65K has been demonstrated to specifically bind U12 snRNA, our results suggest that anti-U11/U12 snRNP-denominated autoantibodies specifically recognise only U11 mono-snRNP, as previously reported by *Gilliam et al.* [306]. As 65K has also been detected to form part of U11 mono-snRNP [290], our data could indicate that 65K is only accessible for autoantibody recognition while forming part of U11 mono-snRNP, while its antigenic epitopes remain hidden while forming part of U11/U12 di-snRNP and U12 mono-snRNP complexes.

5.1.3. Novel autoantibodies against snoRNPs

Overall, snoRNAs, the guide RNA molecules of snoRNPs, were the most abundant final RNA candidates obtained by RIP-Seq, comprising 70.5% of the 197 selected RNA molecules. snoRNPs localise in the nucleoli and catalyse pseudouridylation and 2'-O-methylation of ~200nt of pre-rRNA co-transcriptionally and post-transcriptionally [153]. Based on conserved sequence elements, snoRNAs are classified as C/D box snoRNAs or H/ACA box snoRNAs. While methylation of the 2'-hydroxy groups of the riboses is directed by C/D box

snoRNPs, conversion of uridines to pseudouridines is guided by H/ACA box snoRNPs [153]. snoRNAs are usually derived from introns. Normally, after the splicing reaction, introns are excised as lariats, which are then opened by the debranching enzyme and subsequently degraded. However, C/D box snoRNAs escape this degradation by forming a protein complex that consists of at least NHP2L1, NOP56, NOP58 and fibrillarin. Once the C/D box snoRNP is constituted, the antisense boxes of the snoRNA recognise complementary sequences in rRNA by base-pairing, while the 2'-O-methylation reaction is catalysed specifically by fibrillarin, a SAM [155, 156]. On the other hand, H/ACA box snoRNAs escape from degradation by assembling into a protein complex containing the pseudouridine synthetase dyskerin, and the structural proteins GAR1, NHP2 and NOP10. Mature H/ACA box snoRNPs bind to rRNAs which allows recognition of the substrate uridine that is isomerised to pseudouridine by dyskerin (see section 1.7.5.) [157].

As mentioned, a high number of snoRNAs are coded by intronic sequences of protein-coding genes. Consequently, we detected that some of the protein-coding RNA candidates obtained by RIP-Seq were actually mapped to intronic regions that coded for snoRNAs. In total, we identified seven protein-coding RNAs that had been misclassified as such due to the annotation approach used, which prioritised RNA sequence location within protein-coding gene regions over accurate snoRNA identification in their introns. Consequently, these 7 RNAs were reclassified into snoRNAs. Moreover, we detected that a patient's serum immunoprecipitated a specific intronic sequence from the ANK1 protein-coding gene, likely corresponding to a small RNA molecule yet to be identified.

5.1.3.1. Anti-C/D box snoRNPs autoantibodies

Anti-U3 snoRNP autoantibodies were the first described autoantibodies against a C/D box snoRNP. Initially, these autoantibodies were detected by the immunoprecipitation of U3 snoRNA [164]. Later on, it was demonstrated that the majority of these sera targeted a 34kDa protein [202], and due to its localisation in the FC/DFC regions of the nucleolus, it was named "fibrillarin" (see section 1.7.5.3.) [203]. As previously mentioned, fibrillarin is thought to constitute the catalytic subunit of all C/D box snoRNPs [352, 353]. However, although first studies reported that the majority of anti-U3 snoRNP autoantibodies recognised fibrillarin and U3 snoRNA, subsequent studies have shown that these autoantibodies can also immunoprecipitate other protein components [125, 187, 202, 205] and other RNA molecules [206]. Therefore, heterogeneous reactivity against U3 snoRNP has been reported.

In vitro studies of C/D box snoRNPs have suggested an ordered assembly pathway that takes place directly on the snoRNA. NHP2L1 binds to the k-turn motif of the C/D box snoRNA, recognising highly conserved sequences. The interaction formed by the C/D box snoRNA and NHP2L1 mediates the recruitment of the scaffolding proteins Nop56 and Nop58, which are homologous. Fibrillarin is then integrated into the box C/D snoRNP complex via interaction with Nop56 and Nop58 heterodimer. *In vivo*, this process requires several assembly factors, including the RuvBL1/RuvBL2 proteins [157, 361–363]. However, to date, the only structural information available about eukaryotic C/D box snoRNPs has been obtained from U3 snoRNP, which cannot be considered a canonical C/D box snoRNP. In fact, although U3 snoRNP is the most abundant C/D box snoRNP, it does not methylate pre-rRNA molecules during ribosome biogenesis like most of these ribonucleoproteins. Instead, U3 snoRNP guides endoribonucleolytic processing of the 5' ETS of the 47S pre-rRNA and plays a key role in the maturation of 18S rRNA forming part of a bigger complex denominated SSU processome [200]. Furthermore, in the U3 snoRNP complex, two copies of NHP2L1 bind to two k-turn motifs in U3 snoRNA. Conversely, guide RNAs involved in 2'-O-methylation normally contain only one k-turn motif and bind only one copy of NHP2L1. Therefore, whether the architecture of U3 snoRNP can also be assumed for methylation-competent box C/D snoRNPs remains unclear (**Figure 17**) [364].

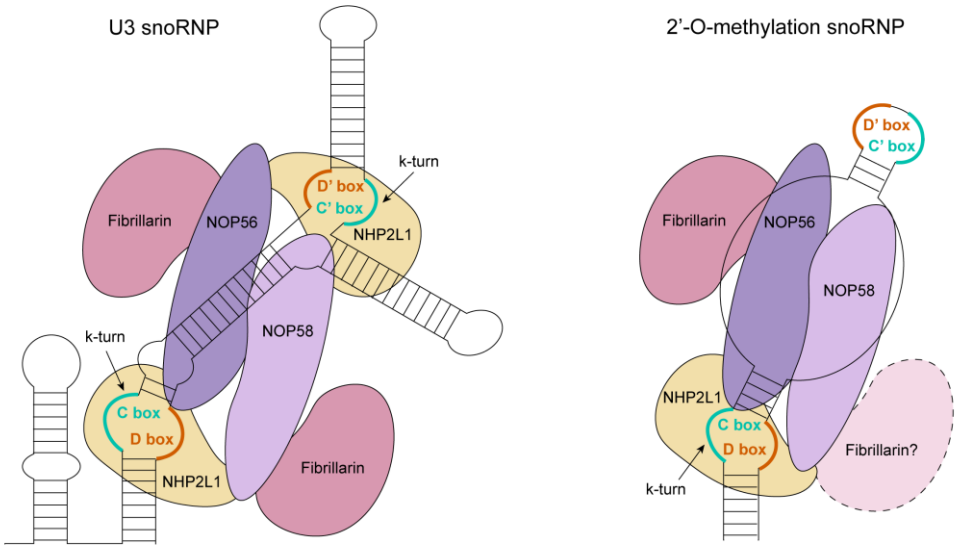


Figure 17. Schematic representation of predicted snoRNP structures. In the U3 snoRNP complex, two copies of NHP2L1 bind to two k-turn motifs in U3 snoRNA by recognising highly conserved sequences. The interaction formed by the C/D box snoRNA and NHP2L1 mediates the recruitment of the scaffolding proteins Nop56 and Nop58, which interact with two fibrillarin molecules. Conversely, guide RNAs involved in 2'-O-methylation normally contain only one k-turn motif and bind only one copy of NHP2L1. Therefore, the architecture of methylation-competent box C/D snoRNPs remains unclear, and whether two copies or one copy of fibrillarin are recruited is yet to be determined.

On the other hand, about half of the known C/D box snoRNP do not have predicted rRNA targets and are considered “orphan C/D box snoRNAs”. Moreover, the association of some C/D box snoRNPs with specific diseases and the demonstration of direct interactions between them and pre-mRNAs, as well as further processing of snoRNAs into miRNAs and shorter RNAs, suggests that C/D box snoRNPs may possess functions in addition to directing the 2’-O-methylation of rRNA. This hypothesis is supported by the demonstration that C/D box snoRNAs are found assembled in protein complexes free from fibrillarin [157, 362, 365–367]. Finally, an asymmetric distribution of core proteins has been demonstrated in yeasts, indicating that different complex proteins can exist associated with C/D box snoRNAs [361]. All these data indicate that different subsets of C/D box snoRNPs, with distinct functions, structure and protein composition, may exist.

The only sample in which U3 snoRNA immunoprecipitation was detected by RIP-Seq was the one presenting the same immunoprecipitation by traditional RNA IP. However, in our study, we found that samples from patients exhibiting nucleolar fibrillar staining pattern on IIF (AC-9) that did not immunoprecipitate U3 snoRNA, neither by traditional RNA IP nor RIP-Seq, immunoprecipitated a wide repertoire of different C/D box snoRNA molecules. Moreover, patients with autoantibodies against C/D box snoRNPs exhibited a heterogeneous profile, as shown by the different RNA immunoprecipitation patterns determined by RNA IP and RIP-Seq. Of particular note, these differing patterns were associated with two distinct clinical phenotypes: some patients exhibited a more severe phenotype similar to that reported for classical anti-U3 snoRNP autoantibodies, with higher rates of heart and lung involvement, as well as arthritis, myositis and dcSSc [210–212], while others had a much milder phenotype. RNA IP was able to differentiate between the two subsets: patients whose samples showed immunoprecipitation of U3, U8, or U13 snoRNA presented the more severe phenotype classically associated with anti-U3 snoRNP autoantibodies when compared to those not immunoprecipitating any RNA molecule. Furthermore, WGCNA of co-immunoprecipitation of RNA molecules detected by RIP-Seq revealed that patients presenting an AC-9 ANA pattern by IIF with a more severe clinical phenotype clustered together, while patients with the same ANA pattern but a milder phenotype did not.

Our data indicate that probably autoantibodies against different subsets of C/D box snoRNPs, composed of different proteins, may arise in SSc patients. More importantly, autoantibodies recognising different C/D box snoRNPs may associated with different clinical phenotypes, and therefore, it is of utmost importance to be able to distinguish those autoantibodies. Additionally, our results demonstrate that available commercial assays

based on the detection of reactivity against fibrillarin were not able to distinguish these different subsets. In this line, ten of the patients from our overall cohort of 307 patients were considered initially positive for anti-U3 snoRNP autoantibody due to a variable positivity of the commercial immunoblot combined with a compatible IIF pattern, but three of them were actually negative when tested by RNA IP. Moreover, these three patients did not present the clinical phenotype associated with anti-U3 snoRNP autoantibodies. On the other hand, two patients who were considered negative by the commercial immunoblot but presented an AC-9 IIF pattern were positive by RNA IP and presented clinical manifestations typically associated with anti-U3 snoRNP autoantibodies. As the immunoprecipitating patterns detected by RNA IP results were highly associated with the different clinical phenotypes, we consider that RNA IP analysis should be performed on all patients with an AC-9 IIF pattern. Actually, RNA IP results could highly contribute to deciding which clinical follow-up approach should be used in SSc patients presenting an AC-9 IIF pattern.

5.1.3.2. Anti-H/ACA box snoRNP autoantibodies

Regarding H/ACA box snoRNAs, the third most represented type of candidate RNAs detected by RIP-Seq, WGCNA analysis showed that patient SSc_44 presented reactivity against various of these ribonucleoproteins. Interestingly, this patient presented an atypical nuclear-speckled pattern by IIF, showing staining of a few nuclear discrete dots and peri-chromosomes. This patient was also positive for anti-U11/U12 snRNP autoantibodies, but this autoantibody is not associated with any specific ANA pattern, and therefore, anti-H/ACA box snoRNP autoantibodies could be accountable for the reactivity detected by IIF. There are few reports on autoantibodies against H/ACA box snoRNPs, and their clinical implications and specific targets remain unclear [329, 353]. Our patient presented the diffuse cutaneous form of the disease and ILD, but due to the co-positivity with anti-U11/U12 snRNP autoantibodies, we could not determine whether anti-H/ACA box snoRNP autoantibodies alone were associated with any specific clinical manifestation.

This patient also presented reactivity against scaRNPs. scaRNPs are specific snoRNPs that localise to the CB and associate with scaRNAs. scaRNAs can either contain a pair of H/ACA box motifs, a pair of C/D box motifs, or the 2 motifs at the same time. scaRNAs presenting C/D box motifs bind to the same proteins as C/D box snoRNPs, while scaRNAs with H/ACA box motifs bind to the same proteins as H/ACA box snoRNPs [298]. The two scaRNAs that were immunoprecipitated by patient SSc_44, present H/ACA box motifs [297] and, therefore,

bind the same proteins as H/ACA box snoRNPs. This finding further confirms that SSc_44 presents autoantibodies against at least one protein component of H/ACA box snoRNPs.

On the other hand, the RNA component of the telomerase holoenzyme requires co-transcriptional assembly as a H/ACA snoRNP with dyskerin, NOP10, NHP2 and GAR1 [330, 332]. Interestingly, autoantibodies targeting telomerase and sheltering proteins have been specifically reported in SSc patients and have been associated with severe lung disease [328, 329]. It could be possible that these autoantibodies also recognise proteins common to H/ACA snoRNPs due to epitope spreading, as reported for other linked sets of autoantibodies in systemic autoimmune diseases.

5.1.4. Novel autoantibodies against 7SK snRNP, mitochondrial tRNA synthetases and vault cytoplasmic complex

As expected, our RIP-Seq results showed that various U5 snRNA and U4 snRNA molecules, which are part of the major and minor spliceosomes, were immunoprecipitated in samples positive for anti-U11/U12 snRNP and anti-U4/U6 snRNP autoantibodies. Unexpectedly, U6 snRNA (major spliceosome) and U6atac snRNA (minor spliceosome) were predominantly immunoprecipitated by SSc_58, a sample that was not known to recognise any protein of the major or minor spliceosome. Interestingly, this sample also immunoprecipitated 7SK RNA. Both U6 and U6atac snRNA, and 7SK present a γ-monomethyl phosphate cap at their 5' ends [269, 368], indicating that sample SSc_58 could present autoantibodies against this structure. Indeed, in SSc autoantibodies against the TMG cap present in the rest of U snRNAs have already been described. These autoantibodies were demonstrated to react with this structure by RNA IP even in the deproteinised extracts [97, 306]. However, autoantibodies against the γ-monomethyl phosphate cap have never been reported before in the sera of a patient or healthy donor. On the other hand, 7SK RNA was also immunoprecipitated by samples SSc_14, SSc_19 and SSc_26, which do not immunoprecipite U6 or U6atac snRNA molecules, suggesting that these patients could present autoantibodies targeting protein components of the 7SK RNP complex. In this line, patients with autoantibodies against TMG caps have also been reported to present autoantibodies against protein components of U snRNPs [97, 306]. Therefore, these RNA structures could become more antigenic in the presence of autoantibodies that recognise protein components to which they are attached. Indeed, T cells responsive to a peptide derived from one component of a supramolecular complex may provide intermolecular and

intra-structural help to B cells producing antibodies to RNA components of the same supramolecular complex [78–80].

Additionally, we identified several patients who presented reactivity against mitochondrial tRNA synthetases but no cytoplasmatic tRNA synthetases by RIP-Seq. Autoantibodies against various cytoplasmatic tRNA synthetases have long been described. These autoantibodies are associated with anti-synthetase syndrome, which is characterised by myositis, ILD, arthritis, RP and mechanic's hands [369]. However, autoantibodies against the mitochondrial tRNA synthetases have never been described before in the literature. Protein components of the tRNA synthetases are not shared between cytoplasmatic and mitochondrial tRNA synthetases [370], and therefore, these newly reported set of autoantibodies are probably not associated with the traditional anti-synthetase syndrome.

On the other hand, we also detected that some patients immunoprecipitated RNA molecules associated with cytoplasmatic vault complexes. These complexes are the biggest ribonucleoprotein complexes of the human cell, with a MW of 13MDa, approximately 3-fold larger than the ribosomes. These big complexes are constituted by several untranslated RNA molecules (vault RNAs) and three high MW proteins, major vault protein (MVP), vault poly(ADP-ribose)polymerase (VPARP) and the telomerase-associated protein 1 (TEP1). The major vault protein (MVP) constitutes over 70% of the molecular mass of the complex, while vault RNAs account for about 5%. Although vaults are not essential to eukaryotic cells, they have been implicated in intracellular transport mechanisms, signal transmission and immune responses [371, 372]. Interestingly, autoantibodies against the MVP have been reported recently in SLE and other systemic autoimmune diseases. This suggests that anti-vault complex autoantibodies may not be specific to SSc but could be associated with it, similar to anti-Ro60 autoantibodies [373]. Interestingly, some of the patients in which reactivity against vault complexes was detected were also positive for anti-Th/To autoantibodies and anti-Ro60. In any case, this study is the first to report these protein complexes as autoantibody targets in SSc. On the other hand, as previously discussed, autoantibodies targeting telomerase and sheltering proteins have been specifically reported in SSc patients [328, 329]. As cytoplasmic vault complexes and the telomerase holoenzyme share TEP1, it could be possible that these autoantibodies recognise these two ribonucleoproteins.

Overall, although we identified several novel autoantibody targets by RIP-Seq, including H/ACA box snoRNPs, mitochondrial tRNA synthetases, 7SK snRNP and the cytoplasmic vault complex, the small number of patients with potential new autoantibodies prevented us

from establishing statistically significant associations between these autoantibodies and specific clinical phenotype or ANA patterns by IIF.

5.2. Autoantibodies detected by protein immunoprecipitation

On the other hand, a newly developed protein IP based on biorthogonal metabolic labelling led to the discovery of a not previously described autoantibody against NVL in six patients. Moreover, this autoantibody was associated with a specific clinical phenotype characterised by a higher prevalence of calcinosis and cancer, specifically synchronous cancer. NVL, a member of the AAA-ATPase type family, is known to exist in two isoforms with N-terminal extensions of different lengths in mammalian cells. To date, many members of the AAA-ATPase family have been identified in a wide range of organisms, ranging from bacteria to mammals. They play important roles in various cellular processes, including proteolysis, membrane fusion, organelle biogenesis, cytoskeletal regulation, protein folding, and DNA replication. Although the biological functions of the various members of this family seem unrelated, it has been proposed that they commonly modulate the assembly and disassembly of macromolecular complexes as energy-dependent molecular machinery. Interestingly, RuvBL1/2 complex, against which autoantibodies have also been detected in SSc, is also an AAA-ATPase [374, 375].

Two alternatively spliced isoforms, NVL1 (a short isoform) and NVL2 (a long isoform) are produced through the utilisation of different methionine and translation initiation sites. NVL1 and NVL2 are differently distributed in the nucleus, but the predominantly expressed NVL2 is mainly localised in the nucleolus. NVL2 is involved in the 60S ribosomal subunit synthesis. Eukaryotic ribosome biogenesis is a highly dynamic multi-step process with more than 200 trans-acting proteins. The assembly and disassembly of these biogenesis factors in the pre-ribosomal particles are strictly coordinated, and many of these interactions are thought to be regulated via energy-consuming mechanisms. Thus, NVL2 likely acts as a molecular chaperone in ribosome biogenesis to facilitate the ordered assembly and remodelling of pre-ribosomal particles. In fact, it has been demonstrated that NVL2 interacts with the human exosome PM/Scl also involved in the biogenesis of ribosomes. Moreover, NVL2 physically interacts with the human telomerase holoenzyme and is required for its assembly [376–378]. Interestingly, autoantibodies against several other components involved in ribosome biogenesis (topoisomerase I, RNApol I, U3 snoRNP, PM/Scl and NOR90) and telomerase holoenzyme have already been described in SSc patients.

Additionally, although not included in the published articles, patients immunoprecipitating other proteins were also found by the newly developed assay. Specifically, one patient was shown to be positive for anti-RuvBL1/2, one positive for anti-Ki/SL and one positive both for anti-RPA and anti-Ki/SL autoantibodies.

Anti-RuvBL1/2 autoantibodies are autoantibodies associated with dcSSc-myositis overlap syndrome. In our cohort, only patient SSc_31 was positive for this autoantibody, who presented dcSSc and myopathy in addition to digital ulcers and recalcitrant calcinosis. Moreover, SSc_31 showed a fine nuclear speckled pattern (AC-4) by IIF as described for these autoantibodies [322]. Considering that SSc-specific autoantibodies are generally mutually exclusive, although we have not tested the 307 SSc patient cohort, we can infer that the prevalence of anti-RuvBL1/2 autoantibodies is 0.3%, indicating that this autoantibody is indeed rare in SSc, as already reported [318].

Anti-Ki/SL autoantibodies, which recognise the proteasome activator subunit 3, are not specific autoantibodies detected in patients with different systemic autoimmune diseases, including SSc. However, they are especially prevalent in SLE, while they have been rarely identified in other rheumatic diseases. Moreover, due to the limited number of studies reporting this autoantibody, that currently cannot be detected by commercial assays, whether these autoantibodies are associated with specific clinical manifestations yet has to be determined [379]. On the other hand, anti-RPA autoantibodies target Replication protein A, a single-stranded-DNA-binding factor involved in DNA replication, repair, and recombination that is constituted by three subunits of 70, 32, and 14kDa. Anti-RPA autoantibodies are not disease specific and have been reported as very rare autoantibodies in SLE and SjS. However, they have not been previously described in SSc patients [380]. In addition, we detected protein bands with a MW not corresponding to any known SSc-related autoantigen in 11 additional samples, indicating that these patients do indeed present ANAs that have not previously been described. However, no more than two patients presented the same immunoprecipitation pattern and were not further studied.

5.3. Anti-IFI16 autoantibodies

Anti-IFI16 autoantibodies are not SSc-specific, but they are considered commonly detected SSc-associated autoantibodies. Moreover, as anti-IFI16 autoantibodies have been associated with the limited cutaneous form of SSc, they are considered of high utility in SSc patients negative for SSc-specific autoantibodies (see section 1.7.12.) [67, 336, 337]. However, no commercial assay for detecting these autoantibodies is yet available. Therefore,

we tested SSc patients negative for all specific autoantibodies by commercial assays (denominated SSc-seronegative henceforth) for anti-IFI16 autoantibodies by an *"in-house"* ELISA.

Up-regulation of IFN-inducible genes, the so-called "IFN signature", has been reported in SSc and other systemic autoimmune diseases [343]**.** IFI16 is a member of the HIN200 gene family, which encodes evolutionarily related nuclear phospho-proteins IFI16, IFIX, MNDA, and AIM-2. These proteins act as innate pattern recognition receptors that sense double-stranded DNA (dsDNA) from invading pathogens in both the nucleus and cytoplasm [344, 345]. Transcriptomic studies performed on skin and oesophagus biopsy specimens have shown that SSc patients fall into one of four intrinsic molecular subsets (inflammatory, fibroproliferative, normal-like and limited) [347, 348, 381, 382]. In the inflammatory molecular subset, IFI16 protein has been found to be up-regulated across several studies [346–348]. Moreover, elevated levels of circulating IFI16 have been found in SSc, SLE, SjS and RA patients [349]. Increased exposure to the protein could explain the production of specific autoantibodies against IFI16. In this line, autoantibodies against IFI16 have been reported with a prevalence ranging from 18.0-29.0% in SSc [67, 336, 337], 23.5%-63.0% in SLE [339–341], 50.0-70.0% in SjS [336, 342], 0.0-13.0% in RA and 3.0% in SSc-PM overlap syndrome [339, 342].

We observed that the prevalence of anti-IFI16 autoantibodies in SS-seronegative patients was 17.2%. The detected frequency of anti-IFI16 was lower than what has been reported in previous assays in which this autoantibody was tested in anti-Scl70 and anti-centromere autoantibody negative patients (25-30%) but higher than was reported in a unique study in which also anti-RNApol III autoantibody negative patients were considered (9%) [67, 336, 337]. On the other hand, 39.4% of anti-centromere positive patients were also positive for anti-IFI16 autoantibodies, indicating that they are indeed a very common autoantibody in SSc. However, a highly variable prevalence of anti-IFI16 autoantibodies within the anti-centromere positive group has been reported in the literature, ranging from 12.7% to 27.5% [67, 337, 338].

Because of the design of our study, few of the included patients were classified as dcSSc. Therefore, no association was found between anti-IFI16 or anti-centromere autoantibodies and any specific cutaneous subset. Nevertheless, all anti-IFI16 positive patients were classified as lcSSc, and anti-IFI16 levels were found to be higher in lcSSc patients than in dcSSc patients, almost reaching statistical significance, indicating the quantitative usefulness of anti-IFI16 levels to distinguish both subsets. Previous studies have reported

conflicting results about anti-IFI16 autoantibodies and their association with the lcSSc subset, and a prevalence of 4% to 14.6% of dcSSc has been reported in anti-IFI16 autoantibody positive patients [67, 336, 337]. However, the prevalence of included dcSSc patients varied from 30.5% to 63.7% in previously studied cohorts, probably accounting for some of the variability observed [67, 336, 337]. Moreover, the discrepancies in overall and dcSSc subset prevalences might also be a consequence of the use of different detection assays and the lack of an international standard reference serum. Therefore, further studies, including a high number of both dcSSc and lcSSc patients, are needed to elucidate the reason for the differences.

Our study found no association between anti-IFI16 autoantibody positivity and a specific ANA IIF pattern, consistent with the fact that IFI16 is not highly expressed by HEp-2 cells [67]. Among all ANA negative SSc patients from our cohort who were tested for anti-IFI16 autoantibodies, 16.7% were anti-IFI16 autoantibody positive, pointing out the usefulness of testing this autoantibody in this group of patients. ANAs tested by IIF in HEp-2 cells is often the screening step for studying SSc-related autoantibodies, but it is important to note that anti-IFI16 autoantibodies could be missed when following this diagnostic flowchart.

When taking into account SSc-seronegative and anti-centromere positive patients, we found higher levels of anti-IFI16 autoantibodies in patients who had RP and digital ulcers than in patients who only had RP. In a matched case-control study, higher levels of anti-IFI16 autoantibodies were found in patients with severe digital ischaemia. A longitudinal study was also performed, showing that anti-IFI16 autoantibody levels indeed varied over time and were higher within six months of a digital ischaemic event [337]. An increased risk of digital vascular events was subsequently found in SSc patients who were both anti-centromere and anti-IFI16 autoantibody positive [338]. In our study, when only anti-centromere autoantibody positive patients were considered, no significant differences were found in anti-IFI16 autoantibody titres when comparing patients who had RP and digital ulcers with patients who only had RP. This different finding may be due to the variability of anti-IFI16 autoantibody levels over time, as the previous study deliberately selected the sera samples closest to the maximum Raynaud's severity score, whereas our study did not [338].

We found that anti-IFI16 autoantibodies were associated with a higher prevalence of isolated PAH and with a lower PAH-free survival. As anti-centromere autoantibodies have traditionally been associated with PAH [383], and more than half of our cohort were anti-centromere autoantibody positive patients, the association of anti-IFI16 autoantibodies with PAH-free survival was also evaluated considering anti-centromere autoantibody status by

multivariate analyses. Moreover, anti-IFI16 autoantibodies have been associated with an older age at the onset of the disease, as we found in the anti-centromere autoantibody positive group. Hence, multivariate analyses, including age at onset of disease, were also performed. Anti-IFI16 autoantibodies remained significantly associated with a higher prevalence of isolated PAH and with a lower PAH-free survival even when adjusting for anti-centromere status and age at onset of disease. Moreover, anti-IFI16 autoantibodies could be used to predict pulmonary involvement in SSc-seronegative patients, as these patients were more likely to have PAH and to have undergone a lung transplant, mainly due to this complication. In any case, when considering only anti-centromere autoantibody positive patients, although anti-IFI16 positive patients were more prone to isolated PAH, anti-IFI16 autoantibodies were not found to be associated with PAH (15.4% *vs.* 7.5%; p=0.420). This lack of association could be explained by the lower heterogeneity of the anti-centromere positive patient group and the low number of patients included in the study.

Interestingly, in our selected cohort of patients, anti-centromere was not an independent predictor of PAH. Overall, it has been described that 10-20% [64] of anti-centromere positive patients present isolated PAH, but anti-centromere has traditionally been associated with this clinical characteristic when compared to anti-Scl70 positive patients, who rarely present isolated PAH. However, some recent studies with large cohorts have not found an increased prevalence of isolated PAH in the anti-centromere positive patients when compared with the frequency of this clinical characteristic in the whole cohort or even when compared with anti-Scl70 positive patients [64]. In our cohort of study, anti-Scl70 positive patients were not included, and while 10.6% of anti-centromere positive patients presented isolated PAH, 6.9% of the SSc-seronegative patients also presented this clinical characteristic, leading to a lack of association of anti-centromere autoantibody to isolated PAH. On the other hand, anti-centromere autoantibody has been negatively associated with ILD [64]. In the present study, this negative association was also found (6.1% vs. 27.6%, p<0.001) when anti-centromere positive patients were compared to SSc-seronegative patients, while anti-IFI16 autoantibodies were not found to be associated with a different prevalence in ILD.

Anti-IFI16 autoantibody positive patients exhibited a statistically significantly poorer overall survival than anti-IFI16 autoantibody negative patients. In fact, anti-IFI16 autoantibody positive patients had 3.2-fold higher risk of death, regardless of anti-centromere status. Even when adjusting for age at onset of disease, anti-IFI16 autoantibody positive patients showed a three-fold higher risk of dying with a trend toward significance (p=0.066). Moreover, in 25%

(n=2) of anti-IFI16 autoantibody positive patients, PAH was the cause of death, indicating that the association of this autoantibody with isolated PAH could partly correlate with the observed overall worse prognosis.

Anti-IFI16 autoantibody results were also analysed in conjunction with results obtained by RNA IP, RIP-Seq and protein IP. However, no association between anti-IFI16 and other detected autoantibodies was found. Therefore, our results confirm previous findings indicating that although anti-IFI16 autoantibodies are SSc-associated autoantibodies, they could be used as a complementary marker of SSc, specifically for the diagnosis of lcSSc in otherwise SSc-seronegative patients. Moreover, we have demonstrated that anti-IFI16 autoantibodies are useful prognostic biomarkers even in patients already positive for anti-centromere autoantibodies, as they are associated with a more severe form of lcSSc. In fact, although SSc-specific autoantibodies such as anti-centromere autoantibodies are closely associated with different SSc clinical phenotypes, significant heterogeneity in clinical outcomes still exists within each of these autoantibody-defined groups. In this line, it has already been described that co-positivity for anti-SMN autoantibodies in MCTD patients positive for anti-U1 snRNP autoantibodies, and co-positivity for anti-Ro60 autoantibodies in SSc patients positive for SSc-specific autoantibodies, are associated with worse overall prognosis and specific clinical manifestations [308, 333–335]. Consequently, screening for co-existing autoantibodies such as anti-IFI16 autoantibodies could be useful for identifying these patient subsets.

5.4. Overall results discussion

Finally, after the application of all the newly developed approaches, only 22 (7.2%) patients of the overall cohort remained negative for specific autoantibodies. Therefore, 46 (67.6%) of the 68 patients who were negative for all SSc-specific autoantibodies detected by commercial assays were actually positive for at least one autoantibody when tested by the newly developed strategies. On the other hand, only 4 of the 22 patients that remained negative were ANA positive by IIF, indicating that 87.5% of patients who tested negative by commercial assays based on specific autoantibody detection but were ANA positive by IIF, resulted positive by the newly developed approaches (Figure 18).

On the other hand, specific autoantibodies were detected by the newly developed strategies in 50.0% of the 36 patients who were ANA negative by IIF (Figure 18). Anti-U11/U12 snRNP and anti-IFI16 autoantibodies were the most common autoantibodies detected in this group, with a prevalence of 16.7% and 16.3%, respectively. Our results indicate that SSc patients

presenting autoantibodies against autoantigens that are not highly expressed by HEp-2 cells could not be detected by IIF. As ANAs tested by IIF in HEp-2 cells is often the screening step for studying SSc-specific and SSc-associated autoantibodies, these autoantibodies could be missed when following this diagnostic flowchart.

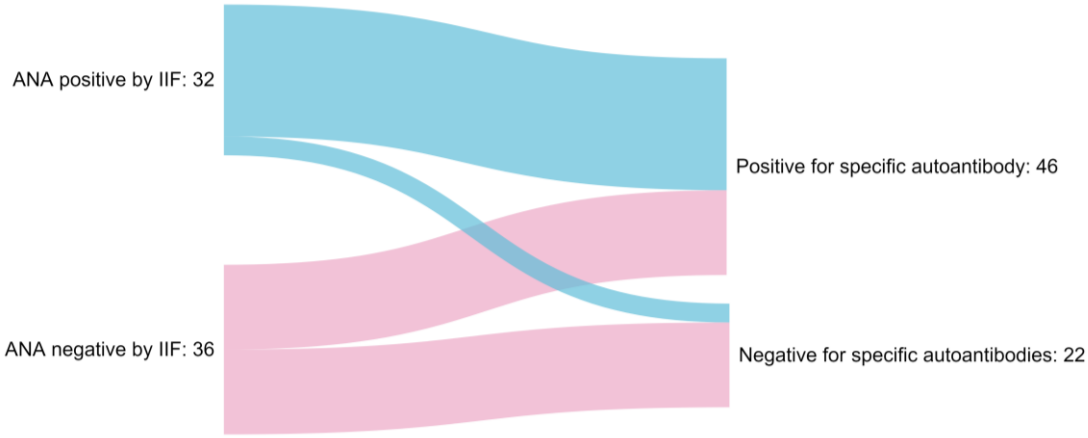


Figure 18. Sankey diagram of IIF patterns and detected specific SS autoantibodies. Graphical summary of the relation between ANA patterns identified by IIF and RNA IP. RIP-Seq, protein IP and ELISA results. Created with *www.SankeyMATIC.com.*

We have detected several autoantibodies that have not been previously reported in SSc patients through the newly developed strategies, including autoantibodies against C/D box snoRNPs, 7SK snRNP, mitochondrial tRNA synthetases, cytoplasmatic vault complexes, NVL and RPA. The small number of patients with potential new autoantibodies prevented us from establishing statistically significant associations between most of these autoantibodies and specific clinical phenotypes. Given the rarity of the disease and the low frequency of these autoantibodies, multicenter studies may be required to achieve a sufficient number of patients. However, we were able to establish that patients presenting reactivity against C/D box snoRNPs distinct to U3, U8 or U13 snoRNP do indeed present mild clinical manifestations, similar to what is observed in anti-centromere positive patients. Interestingly, half of the patients presenting these autoantibodies presented ssSSc. On the other hand, we found that anti-NVL patients were associated with a new clinical phenotype of SSc defined by recalcitrant calcinosis, synchronous cancer and limited cutaneous form of the disease. Furthermore, regarding anti-IFI16 autoantibodies, we confirmed that they are highly prevalent autoantibodies within the lcSSc subset, as previously reported and that their presence is associated with PAH and a poorer overall prognosis in SSc.

Therefore, we have detected three autoantibodies found mainly in the ssSSc and lcSSc subset but with very distinct clinical manifestations (**Figure 19**), further reinforcing the

hypothesis that the detection of autoantibodies enables a better subclassification of SSc patients than the classical approach defined only by cutaneous extension.

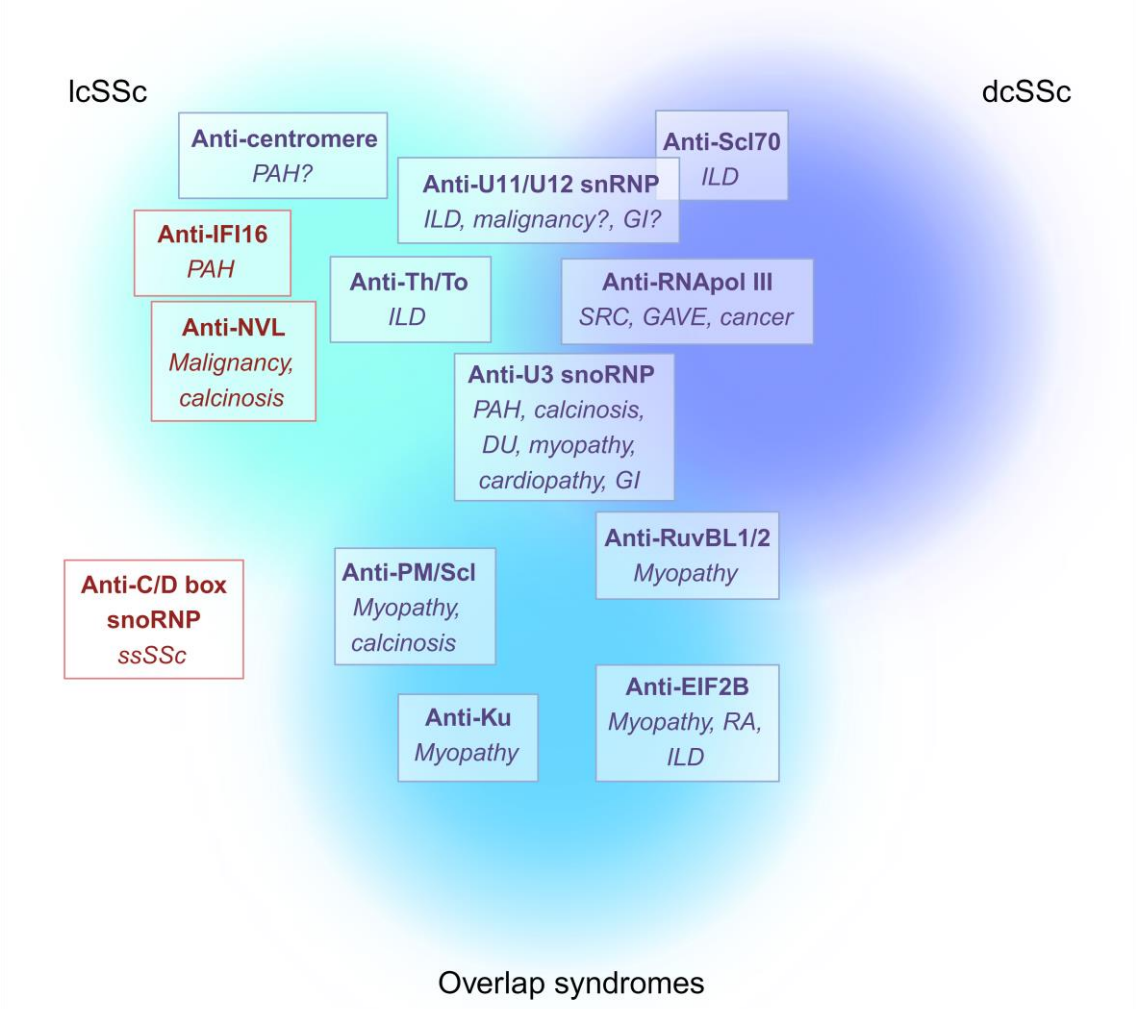


Figure 19. SSc-related autoantibody association with disease subsets and specific clinical manifestations. Previously described SSc-specific autoantibodies are marked in blue, while SSc-related autoantibodies found to be associated with clinical manifestations in this thesis are marked in red. *Created with BioRender.com*

6. CONCLUSIONS

The results obtained in this thesis permit us to propose the following conclusions:

1. Strategies based on RIP-Seq methodology are reliable for detecting autoantibodies in SSc and show higher sensitivity than traditional IP approaches.
2. Protein IP based on bio-orthogonal metabolic labelling coupled with MS, which does not require the use of radioisotopes, is a reliable strategy for new autoantibody identification of SSc-related autoantibodies.
3. A significant group of SSc patients who test negative for specific autoantibodies by commercial assays, especially patients considered ANA positive by IIF, are actually positive when tested by more sensitive assays such as IP and ELISA.
4. Negative ANA results by IIF do not exclude the presence of specific autoantibodies, such as anti-U11/U12 snRNP and anti-IFI16, in patients with SSc.
5. Commercial assays present a low sensitivity for detecting anti-Th/To autoantibodies due to the high complexity of the ribonucleoproteins targeted by these autoantibodies.
6. Anti-U11/U12 snRNP autoantibodies, currently not detectable by commercial assays, are frequently found in SSc patients who test negative for specific autoantibodies and are ANA negative by IIF.
7. Autoantibodies against different subsets of C/D box snoRNPs associated with distinct clinical phenotypes are detected by RNA IP and RIP-Seq in SSc patients.
8. Anti-NVL autoantibodies are associated with an AC-8 ANA IIF pattern and a specific clinical phenotype of SSc defined by limited cutaneous involvement, recalcitrant calcinosis and synchronous cancer.
9. Anti-IFI16 autoantibodies are highly prevalent autoantibodies within the lcSSc subset, and their presence is associated with PAH and a poorer overall prognosis.
10. Overall, not previously reported autoantibodies are very rare in SSc, thus requiring the study of very large cohorts to be able to establish their association with ANA IIF patterns and clinical manifestations.
11. The identification of SSc-related autoantibodies enables a better subclassification of SSc patients than the classical approach defined only by cutaneous extension.

REFERENCES

1. Armando Laborde H, Young P. [History of systemic sclerosis]. Gac Med Mex. 2012;148:201–8.

2. Kowalska-Kępczyńska A. Systemic Scleroderma-Definition, Clinical Picture and Laboratory Diagnostics. J Clin Med. 2022;11.

3. Jerjen R, Nikpour M, Krieg T, Denton CP, Saracino AM. Systemic sclerosis in adults. Part I: Clinical features and pathogenesis. J Am Acad Dermatol. 2022;87:937–54.

4. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TAJ, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol. 1988;15:202–5.

5. Bairkdar M, Rossides M, Westerlind H, Hesselstrand R, Arkema E V, Holmqvist M. Incidence and prevalence of systemic sclerosis globally: a comprehensive systematic review and meta-analysis. Rheumatology (Oxford). 2021;60:3121–33.

6. Arias-Nuñez MC, Llorca J, Vazquez-Rodriguez TR, Gomez-Acebo I, Miranda-Filloy JA, Martin J, et al. Systemic sclerosis in northwestern Spain: A 19-year epidemiologic study. Medicine (Baltimore). 2008;87:272–80.

7. Pattanaik D, Brown M, Postlethwaite BC, Postlethwaite AE. Pathogenesis of Systemic Sclerosis. Front Immunol. 2015;6:272.

8. Asano Y. The Pathogenesis of Systemic Sclerosis: An Understanding Based on a Common Pathologic Cascade across Multiple Organs and Additional Organ-Specific Pathologies. J Clin Med. 2020;9.

9. Fioretto BS, Rosa I, Romano E, Wang Y, Guiducci S, Zhang G, et al. The contribution of epigenetics to the pathogenesis and gender dimorphism of systemic sclerosis: a comprehensive overview. Ther Adv Musculoskelet Dis. 2020;12:1759720X20918456.

10. Hinchcliff M, Garcia-Milian R, Di Donato S, Dill K, Bundschuh E, Galdo F Del. Cellular and Molecular Diversity in Scleroderma. Semin Immunol. 2021;58:101648.

11. Orvain C, Assassi S, Avouac J, Allanore Y. Systemic sclerosis pathogenesis: contribution of recent advances in genetics. Curr Opin Rheumatol. 2020;32:505–14.

12. Ouchene L, Muntyanu A, Lavoué J, Baron M, Litvinov I V, Netchiporouk E. Toward Understanding of Environmental Risk Factors in Systemic Sclerosis [Formula: see text]. J Cutan Med Surg. 2021;25:188–204.

13. Ferri C, Arcangeletti M-C, Caselli E, Zakrzewska K, Maccari C, Calderaro A, et al. Insights into the knowledge of complex diseases: Environmental infectious/toxic agents as potential etiopathogenetic factors of systemic sclerosis. J Autoimmun. 2021;124:102727.

14. Cutolo M, Soldano S, Smith V. Pathophysiology of systemic sclerosis: current understanding and new insights. Expert Rev Clin Immunol. 2019;15:753–64.

15. Rosendahl A-H, Schönborn K, Krieg T. Pathophysiology of systemic sclerosis (scleroderma). Kaohsiung J Med Sci. 2022;38:187–95.

16. Pillai S. T and B lymphocytes in fibrosis and systemic sclerosis. Curr Opin Rheumatol. 2019;31:576–81.

17. Volkmann ER, Andréasson K, Smith V. Systemic sclerosis. Lancet (London, England). 2023;401:304–18.

18. Sticherling M. Systemic sclerosis - the dermatological perspective. J der Dtsch Dermatologischen Gesellschaft = J Ger Soc Dermatology JDDG. 2019;17:716–28.

19. Khanna D, Furst DE, Clements PJ, Allanore Y, Baron M, Czirjak L, et al. Standardization of the modified Rodnan skin score for use in clinical trials of systemic sclerosis. J scleroderma Relat Disord. 2017;2:11–8.

20. 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.

21. Smith V, Herrick AL, Ingegnoli F, Damjanov N, De Angelis R, Denton CP, et al. Standardisation of nailfold capillaroscopy for the assessment of patients with Raynaud's phenomenon and systemic sclerosis. Autoimmun Rev. 2020;19:102458.

22. Cutolo M, Sulli A, Smith V. Assessing microvascular changes in systemic sclerosis diagnosis and management. Nat Rev Rheumatol. 2010;6:578–87.

23. LeRoy EC, Medsger TAJ. Criteria for the classification of early systemic sclerosis. J Rheumatol. 2001;28:1573–6.

24. Pearson DR, Werth VP, Pappas-Taffer L. Systemic sclerosis: Current concepts of skin and systemic manifestations. Clin Dermatol. 2018;36:459–74.

25. Mormile I, Mosella F, Turco P, Napolitano F, de Paulis A, Rossi FW. Calcinosis Cutis and Calciphylaxis in Autoimmune Connective Tissue Diseases. Vaccines. 2023;11.

26. Boulman N, Slobodin G, Rozenbaum M, Rosner I. Calcinosis in rheumatic diseases. Semin Arthritis Rheum. 2005;34:805–12.

27. Valenzuela A, Chung L. Subcutaneous calcinosis: Is it different between systemic sclerosis and dermatomyositis? J scleroderma Relat Disord. 2022;7:7–23.

28. Hsu V, Varga J, Schlesinger N. Calcinosis in scleroderma made crystal clear. Curr Opin Rheumatol. 2019;31:589–94.

29. Denton CP, Wells AU, Coghlan JG. Major lung complications of systemic sclerosis. Nat Rev Rheumatol. 2018;14:511–27.

30. Perelas A, Silver RM, Arrossi A V, Highland KB. Systemic sclerosis-associated interstitial lung disease. Lancet Respir Med. 2020;8:304–20.

31. Scallan CJ, Collins BF, Raghu G. Systemic sclerosis-associated interstitial lung disease. Curr Opin Pulm Med. 2020;26:487–95.

32. Condon DF, Nickel NP, Anderson R, Mirza S, de Jesus Perez VA. The 6th World Symposium on Pulmonary Hypertension: what's old is new. F1000Research. 2019;8.

33. Attanasio U, Cuomo A, Pirozzi F, Loffredo S, Abete P, Petretta M, et al. Pulmonary Hypertension Phenotypes in Systemic Sclerosis: The Right Diagnosis for the Right Treatment. Int J Mol Sci. 2020;21.

34. Oudiz RJ. Classification of Pulmonary Hypertension. Cardiol Clin. 2016;34:359–61.

35. Haque A, Kiely DG, Kovacs G, Thompson AAR, Condliffe R. Pulmonary hypertension phenotypes in patients with systemic sclerosis. Eur Respir Rev an Off J Eur Respir Soc. 2021;30.

36. Lechartier B, Humbert M. Pulmonary arterial hypertension in systemic sclerosis. Presse Med. 2021;50:104062.

37. Chrabaszcz M, Małyszko J, Sikora M, Alda-Malicka R, Stochmal A, Matuszkiewicz-Rowinska J, et al. Renal Involvement in Systemic Sclerosis: An Update. Kidney Blood Press Res. 2020;45:532–48.

38. Reggiani F, Moroni G, Ponticelli C. Kidney Involvement in Systemic Sclerosis. J Pers Med. 2022;12.

39. Hung G, Mercurio V, Hsu S, Mathai SC, Shah AA, Mukherjee M. Progress in Understanding, Diagnosing, and Managing Cardiac Complications of Systemic Sclerosis. Curr Rheumatol Rep. 2019;21:68.

40. Nie L-Y, Wang X-D, Zhang T, Xue J. Cardiac complications in systemic sclerosis: early diagnosis and treatment. Chin Med J (Engl). 2019;132:2865–71.

41. Mani P, Gonzalez D, Chatterjee S, Faulx MD. Cardiovascular complications of systemic sclerosis: What to look for. Cleve Clin J Med. 2019;86:685–95.

42. Miller JB, Gandhi N, Clarke J, McMahan Z. Gastrointestinal Involvement in Systemic Sclerosis: An Update. J Clin Rheumatol Pract reports Rheum Musculoskelet Dis. 2018;24:328–37.

43. Gyger G, Baron M. Gastrointestinal manifestations of scleroderma: recent progress in evaluation, pathogenesis, and management. Curr Rheumatol Rep. 2012;14:22–9.

44. Frech TM, Mar D. Gastrointestinal and Hepatic Disease in Systemic Sclerosis. Rheum Dis Clin North Am. 2018;44:15–28.

45. Sjogren RW. Gastrointestinal features of scleroderma. Curr Opin Rheumatol. 1996;8:569–75.

46. Lepri G, Bellando Randone S, Matucci Cerinic M, Allanore Y. Systemic sclerosis and primary biliary cholangitis: An overlapping entity? J scleroderma Relat Disord. 2019;4:111–7.

47. Lóránd V, Czirják L, Minier T. Musculoskeletal involvement in systemic sclerosis. Presse Med. 2014;43 10 Pt 2:e315-28.

48. Walker UA, Clements PJ, Allanore Y, Distler O, Oddis C V, Khanna D, et al. Muscle involvement in systemic sclerosis: points to consider in clinical trials. Rheumatology (Oxford). 2017;56 suppl_5:v38–44.

49. Giannini M, Ellezam B, Leclair V, Lefebvre F, Troyanov Y, Hudson M, et al. Scleromyositis: A distinct novel entity within the systemic sclerosis and autoimmune myositis spectrum. Implications for care and pathogenesis. Front Immunol. 2022;13:974078.

50. Brown CR, Crouser NJ, Speeckaert AL. Considerations for Hand Surgery in Patients With Scleroderma. Hand (N Y). 2023;18:32–9.

51. Zimmermann F, Robin F, Caillault L, Cazalets C, Llamas-Gutierrez F, Garlantézec R, et al. Sicca syndrome in systemic sclerosis: a narrative review on a neglected issue. Rheumatology (Oxford). 2023;62 SI:SI1–11.

52. Salliot C, Mouthon L, Ardizzone M, Sibilia J, Guillevin L, Gottenberg J-E, et al. Sjogren's syndrome is associated with and not secondary to systemic sclerosis. Rheumatology (Oxford). 2007;46:321–6.

53. Lepri G, Catalano M, Bellando-Randone S, Pillozzi S, Giommoni E, Giorgione R, et al. Systemic Sclerosis Association with Malignancy. Clin Rev Allergy Immunol. 2022;63:398–416.

54. Weeding E, Casciola-Rosen L, Shah AA. Cancer and Scleroderma. Rheum Dis Clin North Am. 2020;46:551–64.

55. Morrisroe K, Nikpour M. Cancer and scleroderma: recent insights. Curr Opin Rheumatol. 2020;32:479–87.

56. Fragoulis GE, Daoussis D, Pagkopoulou E, Garyfallos A, Kitas GD, Dimitroulas T. Cancer risk in systemic sclerosis: identifying risk and managing high-risk patients. Expert Rev Clin Immunol. 2020;16:1105–13.

57. Tyndall AJ, Bannert B, Vonk M, Airò P, Cozzi F, Carreira PE, et al. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. Ann Rheum Dis. 2010;69:1809–15.

58. Rubio-Rivas M, Royo C, Simeón CP, Corbella X, Fonollosa V. Mortality and survival in systemic sclerosis: systematic review and meta-analysis. Semin Arthritis Rheum. 2014;44:208–19.

59. Poudel DR, Jayakumar D, Danve A, Sehra ST, Derk CT. Determinants of mortality in systemic sclerosis: a focused review. Rheumatol Int. 2018;38:1847–58.

60. Poudel DR, Derk CT. Mortality and survival in systemic sclerosis: A review of recent literature. Curr Opin Rheumatol. 2018;30:588–93.

61. Sobanski V, Lescoat A, Launay D. Novel classifications for systemic sclerosis: challenging historical subsets to unlock new doors. Curr Opin Rheumatol. 2020;32:463–71.

62. Masi AT. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Arthritis Rheum. 1980;23:581–90.

63. Sobanski V, Giovannelli J, Allanore Y, Riemekasten G, Airò P, Vettori S, et al. Phenotypes Determined by Cluster Analysis and Their Survival in the Prospective European Scleroderma Trials and Research Cohort of Patients With Systemic Sclerosis. Arthritis Rheumatol (Hoboken, NJ). 2019;71:1553–70.

64. Nihtyanova SI, Sari A, Harvey JC, Leslie A, Derrett-Smith EC, Fonseca C, et al. Using Autoantibodies and

Cutaneous Subset to Develop Outcome-Based Disease Classification in Systemic Sclerosis. Arthritis Rheumatol (Hoboken, NJ). 2020;72:465–76.

65. Stochmal A, Czuwara J, Trojanowska M, Rudnicka L. Antinuclear Antibodies in Systemic Sclerosis: an Update. Clin Rev Allergy Immunol. 2020;58:40–51.

66. Kuwana M. Circulating anti-nuclear antibodies in systemic sclerosis: Utility in diagnosis and disease subsetting. Journal of Nippon Medical School. 2017;84:56–63.

67. Costa S, Mondini M, Caneparo V, Afeltra A, Airò P, Bellisai F, et al. Detection of anti-IFI16 antibodies by ELISA: Clinical and serological associations in systemic sclerosis. Rheumatology. 2011;50:674–81.

68. Casciola-Rosen LA, Miller DK, Anhalt GJ, Rosen A. Specific cleavage of the 70-kDa protein component of the U1 small nuclear ribonucleoprotein is a characteristic biochemical feature of apoptotic cell death. J Biol Chem. 1994;269:30757–60.

69. Casciola-Rosen LA, Anhalt GJ, Rosen A. DNA-dependent protein kinase is one of a subset of autoantigens specifically cleaved early during apoptosis. J Exp Med. 1995;182:1625–34.

70. Casciola-Rosen L, Nicholson DW, Chong T, Rowan KR, Thornberry NA, Miller DK, et al. Apopain/CPP32 cleaves proteins that are essential for cellular repair: a fundamental principle of apoptotic death. J Exp Med. 1996;183:1957–64.

71. Casiano CA, Martin SJ, Green DR, Tan EM. Selective cleavage of nuclear autoantigens during CD95 (Fas/APO-1)-mediated T cell apoptosis. J Exp Med. 1996;184:765–70.

72. McConnell KR, Dynan WS, Hardin JA. The DNA-dependent protein kinase catalytic subunit (p460) is cleaved during Fas-mediated apoptosis in Jurkat cells. J Immunol. 1997;158:2083–9.

73. Takeda Y, Caudell P, Grady G, Wang G, Suwa A, Sharp GC, et al. Human RNA helicase A is a lupus autoantigen that is cleaved during apoptosis. J Immunol. 1999;163:6269–74.

74. Casciola-Rosen L, Andrade F, Ulanet D, Wong WB, Rosen A. Cleavage by granzyme B is strongly predictive of autoantigen status: implications for initiation of autoimmunity. J Exp Med. 1999;190:815–26.

75. Takeda Y, Dynan WS. Autoantibodies against DNA double-strand break repair proteins. Front Biosci. 2001;6:D1412-22.

76. Mattioli M, Reichlin M. Physical association of two nuclear antigens and mutual occurrence of their antibodies: the relationship of the SM and RNAprotein (MO) systems in SLE sera. J Immunol. 1973;110:1318–24.

77. Reeves WH, Fisher DE, Lahita RG, Kunkel HG. Autoimmune sera reactive with Sm antigen contain high levels of RNP-like antibodies. J Clin Invest. 1985;75:580–7.

78. Lake P, Mitchison NA. Regulatory mechanisms in the immune response to cell-surface antigens. Cold Spring Harb Symp Quant Biol. 1977;41 Pt 2:589–95.

79. Russell SM, Liew FY. T cells primed by influenza virion internal components can cooperate in the antibody response to haemagglutinin. Nature. 1979;280:147–8.

80. Scherle PA, Gerhard W. Functional analysis of influenza-specific helper T cell clones in vivo. T cells specific for internal viral proteins provide cognate help for B cell responses to hemagglutinin. J Exp Med. 1986;164:1114–28.

81. Satoh M, Ajmani AK, Stojanov L, Langdon JJ, Ogasawara T, Wang J, et al. Autoantibodies that stabilize the molecular interaction of Ku antigen with DNA-dependent protein kinase catalytic subunit. Clin Exp Immunol. 1996;105:460–7.

82. Ling M, Murali M. Antinuclear Antibody Tests. Clin Lab Med. 2019;39:513–24.

83. Damoiseaux J, von Mühlen CA, Garcia-De La Torre I, Carballo OG, de Melo Cruvinel W, Francescantonio PLC, et al. International consensus on ANA patterns (ICAP): the bumpy road towards a consensus on reporting ANA results. Auto- Immun highlights. 2016;7:1.

84. Wang J, Dong X, Stojanov L, Kimpel D, Satoh M, Reeves WH. Human autoantibodies stabilize the quaternary structure of Ku antigen. Arthritis Rheum. 1997;40:1344–53.

85. Franceschini F, Cavazzana I, Generali D, Quinzanini M, Viardi L, Ghirardello A, et al. Anti-Ku antibodies in connective tissue diseases: Clinical and serological evaluation of 14 patients. J Rheumatol. 2002;29:1393–7.

86. Francoeur AM, Peebles CL, Gompper PT, Tan EM. Identification of Ki (Ku, p70/p80) autoantigens and analysis of anti-Ki autoantibody reactivity. J Immunol. 1986;136:1648–53.

87. Yaneva M, Arnett FC. Antibodies against Ku protein in sera from patients with autoimmune diseases. Clin Exp Immunol. 1989;76:366–72.

88. Reeves WH, Pierani A, Chou CH, Ng T, Nicastri C, Roeder RG, et al. Epitopes of the p70 and p80 (Ku) lupus autoantigens. J Immunol. 1991;146:2678–86.

89. Kuwana M, Kaburaki J, Mimori T, Tojo T, Homma M. Autoantibody reactive with three classes of RNA polymerases in sera from patients with systemic sclerosis. J Clin Invest. 1993;91:1399–404.

90. Mak TW, Saunders ME. Exploiting Antigen–Antibody Interaction. In: The Immune Response. Elsevier; 2006. p. 147–77.

91. James K, Meek G. Evaluation of commercial enzyme immunoassays compared to immunofluorescence and double diffusion for autoantibodies associated with autoimmune diseases. Am J Clin Pathol. 1992;97:559–65.

92. Bridges AJ, Lorden TE, Havighurst TC. Autoantibody testing for connective tissue diseases. Comparison of immunodiffusion, immunoblot, and enzyme immunoassay. Am J Clin Pathol. 1997;108:406–10.

93. Mahler M, Satoh M, Hudson M, Baron M, Chan JYF, Chan EKL, et al. Autoantibodies to the Rpp25 component of the Th/To complex are the most common antibodies in patients with systemic sclerosis without antibodies detectable by widely available commercial tests. J Rheumatol. 2014;41:1334–43.

94. Van Praet JT, Van Steendam K, Smith V, De Bruyne G, Mimori T, Bonroy C, et al. Specific anti-nuclear antibodies in systemic sclerosis patients with and without skin involvement: an extended methodological approach. Rheumatology (Oxford). 2011;50:1302–9.

95. Hamaguchi Y, Kuwana M, Takehara K. Performance evaluation of a commercial line blot assay system for detection of myositis- and systemic sclerosis-related autoantibodies. Clin Rheumatol. 2020;39:3489–97.

96. Gerovac M, Vogel J, Smirnov A. The World of Stable Ribonucleoproteins and Its Mapping With Grad-Seq and Related Approaches. Front Mol Biosci. 2021;8:661448.

97. Fertig N, Domsic RT, Rodriguez-Reyna T, Kuwana M, Lucas M, Medsger TA, et al. Anti-U11/U12 RNP antibodies in systemic sclerosis: A new serologic marker associated with pulmonary fibrosis. Arthritis Care Res. 2009;61:958–65.

98. Perurena-Prieto J, Viñas-Giménez L, Sanz-Martínez MT, Selva-O'Callaghan A, Callejas-Moraga EL, Colobran R, et al. Anti-nuclear valosin-containing protein-like autoantibody is associated with calcinosis and higher risk of cancer in systemic sclerosis. Rheumatology (Oxford). 2023;Online ahead of print.

99. Montzka KA, Steitz JA. Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. Proc Natl Acad Sci U S A. 1988;85:8885–9.

100. Lee B, Craft JE. Molecular structure and function of autoantigens in systemic sclerosis. Int Rev Immunol. 1995;12:129–44.

101. Kajio N, Takeshita M, Suzuki K, Kaneda Y, Yamane H, Ikeura K, et al. Anti-centromere antibodies target centromere-kinetochore macrocomplex: a comprehensive autoantigen profiling. Ann Rheum Dis. 2021;80:651–9.

102. Song G, Hu C, Zhu H, Wang L, Zhang F, Li Y, et al. New centromere autoantigens identified in systemic sclerosis using centromere protein microarrays. J Rheumatol. 2013;40:461–8.

103. Tan EM, Rodnan GP, Garcia I, Moroi Y, Fritzler MJ, Peebles C. Diversity of antinuclear antibodies in progressive systemic sclerosis. Anti-centromere antibody and its relationship to CREST syndrome. Arthritis

Rheum. 1980;23:617–25.

104. Pokeerbux MR, Giovannelli J, Dauchet L, Mouthon L, Agard C, Lega JC, et al. Survival and prognosis factors in systemic sclerosis: data of a French multicenter cohort, systematic review, and meta-analysis of the literature. Arthritis Res Ther. 2019;21:86.

105. Czömpöly T, Simon D, Czirják L, Németh P. Anti-topoisomerase I autoantibodies in systemic sclerosis. Autoimmun Rev. 2009;8:692–6.

106. Andrade LEC, Klotz W, Herold M, Conrad K, Rönnelid J, Fritzler MJ, et al. International consensus on antinuclear antibody patterns: definition of the AC-29 pattern associated with antibodies to DNA topoisomerase I. Clin Chem Lab Med. 2018;56:1783–8.

107. ROEDER RG, RUTTER WJ. Multiple Forms of DNA-dependent RNA Polymerase in Eukaryotic Organisms. Nature. 1969;224:234–7.

108. Roeder RG, Rutter WJ. Specific nucleolar and nucleoplasmic RNA polymerases. Proc Natl Acad Sci U S A. 1970;65:675–82.

109. WILLIS IM. RNA polymerase III: Genes, factors and transcriptional specificity. Eur J Biochem. 1993;212:1–11.

110. Dieci G, Conti A, Pagano A, Carnevali D. Identification of RNA polymerase III-transcribed genes in eukaryotic genomes. Biochim Biophys Acta - Gene Regul Mech. 2013;1829:296–305.

111. Vannini A, Cramer P. Conservation between the RNA Polymerase I, II, and III Transcription Initiation Machineries. Mol Cell. 2012;45:439–46.

112. Cramer P, Armache KJ, Baumli S, Benkert S, Brueckner F, Buchen C, et al. Structure of eukaryotic RNA polymerases. Annu Rev Biophys. 2008;37:337–52.

113. Okano Y, Steen VD, Medsger TA. Autoantibody reactive with RNA polymerase III in systemic sclerosis. Ann Intern Med. 1993;119:1005–13.

114. Harvey GR, Rands AL, Mchugh NJ. Anti-RNA polymerase antibodies in systemic sclerosis (SSc): Association with anti-topoisomerase I antibodies and identification of autoreactive subunits of RNA polymerase II. Clin Exp Immunol. 1996;105:468–74.

115. Hirakata M, Okano Y, Pati U, Suwa A, Medsger TA, Hardin JA, et al. Identification of autoantibodies to RNA polymerase II: Occurrence in systemic sclerosis and association with autoantibodies to RNA polymerases I and III. J Clin Invest. 1993;91:2665–72.

116. Satoh M, Kuwana M, Ogasawara T, Ajmani AK, Langdon JJ, Kimpel D, et al. Association of autoantibodies to topoisomerase I and the phosphorylated (IIO) form of RNA polymerase II in Japanese scleroderma patients. J Immunol. 1994;153:5838–48.

117. Harvey GR, Butts S, Rands AL, Patel Y, McHugh NJ. Clinical and serological associations with anti-RNA polymerase antibodies in systemic sclerosis. Clin Exp Immunol. 1999;117:395–402.

118. Satoh M, Ajmani AK, Ogasawara T, Langdon JJ, Hirakata M, Wang J, et al. Autoantibodies to RNA polymerase II are common in systemic lupus erythematosus and overlap syndrome. Specific recognition of the phosphorylated (IIO) form by a subset of human sera. J Clin Invest. 1994;94:1981–9.

119. Kuwana M, Okano Y, Kaburaki J, Medsger TA, Wright TM. Autoantibodies to RNA polymerases recognize multiple subunits and demonstrate cross-reactivity with rna polymerase complexes. Arthritis Rheum. 1999;42:275–84.

120. Kuwana M, Kimura K, Kawakami Y. Identification of an immunodominant epitope on RNA polymerase III recognized by systemic sclerosis sera: Application to enzyme-linked immunosorbent assay. Arthritis Rheum. 2002;46:2742–7.

121. Kuwana M, Okano Y, Pandey JP, Silver RM, Fertig N, Medsger TA. Enzyme-linked immunosorbent assay for detection of anti-RNA polymerase III antibody: Analytical accuracy and clinical associations in systemic sclerosis. Arthritis Rheum. 2005;52:2425–32.

122. Villalta D, Imbastaro T, Di Giovanni S, Lauriti C, Gabini M, Turi MC, et al. Diagnostic accuracy and predictive value of extended autoantibody profile in systemic sclerosis. Autoimmun Rev. 2012;12:114–20.

123. Yamasaki Y, Honkanen-Scott M, Hernandez L, Ikeda K, Barker T, Bubb MR, et al. Nucleolar staining cannot be used as a screening test for the scleroderma marker anti-RNA Polymerase I/III antibodies. Arthritis Rheum. 2006;54:3051–6.

124. Parker JC, Burlingame RW, Webb TT, Bunn CC. Anti-RNA polymerase III antibodies in patients with systemic sclerosis detected by indirect immunofluorescence and ELISA. Rheumatology. 2008;47:976–9.

125. Reimer G, Steen VD, Penning CA, Medsger TAJ, Tan EM. Correlates between autoantibodies to nucleolar antigens and clinical features in patients with systemic sclerosis (scleroderma). Arthritis Rheum. 1988;31:525–32.

126. Jones E, Kimura H, Vigneron M, Wang Z, Roeder RG, Cook PR. Isolation and characterization of monoclonal antibodies directed against subunits of human RNA polymerases I, II, and III. Exp Cell Res. 2000;254:163–72.

127. Lazzaroni M-G, Cavazzana I, Colombo E, Dobrota R, Hernandez J, Hesselstrand R, et al. Malignancies in Patients with Anti-RNA Polymerase III Antibodies and Systemic Sclerosis: Analysis of the EULAR Scleroderma Trials and Research Cohort and Possible Recommendations for Screening. J Rheumatol. 2017;44:639–47.

128. Lazzaroni M-G, Airò P. Anti-RNA polymerase III antibodies in patients with suspected and definite systemic sclerosis: Why and how to screen. J scleroderma Relat Disord. 2018;3:214–20.

129. Callejas-Moraga EL, Guillén-Del-Castillo A, Marín-Sánchez AM, Roca-Herrera M, Balada E, Tolosa-Vilella C, et al. Clinical features of systemic sclerosis patients with anti-RNA polymerase III antibody in a single centre in Spain. Clin Exp Rheumatol. 2019;37:41–8.

130. Joseph CG, Darrah E, Shah AA, Skora AD, Casciola-Rosen LA, Wigley FM, et al. Association of the autoimmune disease scleroderma with an immunologic response to cancer. Science. 2014;343:152–7.

131. Shah AA, Laiho M, Rosen A, Casciola-Rosen L. Protective Effect Against Cancer of Antibodies to the Large Subunits of Both RNA Polymerases I and III in Scleroderma. Arthritis Rheumatol (Hoboken, NJ). 2019;71:1571–9.

132. Amon A. A decade of Cdc14 - A personal perspectiveDelivered on 9 July 2007 at the 32nd FEBS Congress in Vienna, Austria. FEBS J. 2008;275:5774–84.

133. Boulon S, Westman BJ, Hutten S, Boisvert FM, Lamond AI. The Nucleolus under Stress. Mol Cell. 2010;40:216–27.

134. Salvetti A, Greco A. Viruses and the nucleolus: The fatal attraction. Biochim Biophys Acta - Mol Basis Dis. 2014;1842:840–7.

135. Hetman M, Slomnicki LP. Ribosomal biogenesis as an emerging target of neurodevelopmental pathologies. J Neurochem. 2019;148:325–47.

136. Callé A, Ugrinova I, Epstein AL, Bouvet P, Diaz J-J, Greco A. Nucleolin Is Required for an Efficient Herpes Simplex Virus Type 1 Infection. J Virol. 2008;82:4762–73.

137. Stenström L, Mahdessian D, Gnann C, Cesnik AJ, Ouyang W, Leonetti MD, et al. Mapping the nucleolar proteome reveals a spatiotemporal organization related to intrinsic protein disorder. Mol Syst Biol. 2020;16:1–16.

138. Németh A, Conesa A, Santoyo-Lopez J, Medina I, Montaner D, Péterfia B, et al. Initial genomics of the human nucleolus. PLoS Genet. 2010;6.

139. Lafontaine DLJ, Riback JA, Bascetin R, Brangwynne CP. The nucleolus as a multiphase liquid condensate. Nat Rev Mol Cell Biol. 2021;22:165–82.

140. Kressler D, Linder P, de la Cruz J. Protein trans -Acting Factors Involved in Ribosome Biogenesis in Saccharomyces cerevisiae . Mol Cell Biol. 1999;19:7897–912.

141. Henderson AS, Warburton D, Atwood KC. Location of ribosomal DNA in the human chromosome complement. Proc Natl Acad Sci U S A. 1972;69:3394–8.

142. Worton RG, Sutherland J, Sylvester JE, Willard HF, Bodrug S, Dubé I, et al. Human ribosomal RNA genes: Orientation of the tandem array and conservation of the 5′ end. Science (80-). 1988;239:64–8.

143. Gonzalez IL, Sylvester JE. Complete Sequence of the 43-kb Human Ribosomal DNA Repeat: Analysis of the Intergenic Spacer. Genomics. 1995;27:320–8.

144. Goodfellow SJ, Zomerdijk JCBM. Basic Mechanisms in RNA Polymerase I Transcription of the Ribosomal RNA Genes. In: Epigenetics: Development and Disease,. 2013. p. 211–36.

145. Roussel P, André C, Comai L, Hernandez-Verdun D. The rDNA transcription machinery is assembled during mitosis in active NORs and absent in inactive NORs. J Cell Biol. 1996;133:235–46.

146. Savino TM, Gébrane-Younès J, De Mey J, Sibarita JB, Hernandez-Verdun D. Nucleolar assembly of the rRNA processing machinery in living cells. J Cell Biol. 2001;153:1097–110.

147. Scheer U, Rose KM. Localization of RNA polymerase I in interphase cells and mitotic chromosomes by light and electron microscopic immunocytochemistry. Proc Natl Acad Sci U S A. 1984;81:1431–5.

148. Lafontaine DLJ. Birth of Nucleolar Compartments: Phase Separation-Driven Ribosomal RNA Sorting and Processing. Mol Cell. 2019;76:694–6.

149. Pederson T. The nucleolus. Cold Spring Harb Perspect Biol. 2011;3:1–15.

150. O'Donohue MF, Choesmel V, Faubladier M, Fichant G, Gleizes PE. Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits. J Cell Biol. 2010;190:853–66.

151. Yao RW, Xu G, Wang Y, Shan L, Luan PF, Wang Y, et al. Nascent Pre-rRNA Sorting via Phase Separation Drives the Assembly of Dense Fibrillar Components in the Human Nucleolus. Mol Cell. 2019;76:767-783.e11.

152. Prieto JL, McStay B. Nucleolar biogenesis: The first small steps. Biochem Soc Trans. 2005;33:1441–3.

153. Kiss T. Small nucleolar RNAs: An abundant group of noncoding RNAs with diverse cellular functions. Cell. 2002;109:145–8.

154. Massenet S, Bertrand E, Verheggen C. Assembly and trafficking of box C/D and H/ACA snoRNPs. RNA Biol. 2017;14:680–92.

155. Tollervey D, Lehtonen H, Jansen R, Kern H, Hurt EC. Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. Cell. 1993;72:443–57.

156. Eichler DC, Craig N. Processing of eukaryotic ribosomal RNA. Prog Nucleic Acid Res Mol Biol. 1994;49:197–239.

157. Stamm S, Lodmell JS. C/D box snoRNAs in viral infections: RNA viruses use old dogs for new tricks. Non-coding RNA Res. 2019;4:46–53.

158. Boisvert FM, Van Koningsbruggen S, Navascués J, Lamond AI. The multifunctional nucleolus. Nat Rev Mol Cell Biol. 2007;8:574–85.

159. Iarovaia O V., Minina EP, Sheval E V., Onichtchouk D, Dokudovskaya S, Razin S V., et al. Nucleolus: A Central Hub for Nuclear Functions. Trends Cell Biol. 2019;29:647–59.

160. Pederson T. The plurifunctional nucleolus. Nucleic Acids Res. 1998;26:3871–6.

161. Satoh M, Ceribelli A, Hasegawa T, Tanaka S. Clinical Significance of Antinucleolar Antibodies: Biomarkers for Autoimmune Diseases, Malignancies, and others. Clin Rev Allergy Immunol. 2022;63:210–39.

162. Dima A, Vonk MC, Garaiman A, Kersten BE, Becvar R, Tomcik M, et al. Clinical significance of the anti-Nucleolar Organizer Region 90 antibodies (NOR90) in systemic sclerosis: Analysis of the European Scleroderma Trials and Research (EUSTAR) cohort and a systematic literature review. Eur J Intern Med. 2024;125:104–10.

163. Hardin JA, Rahn DR, Shen C, Lerner MR, Wolin SL, Rosa MD, et al. Antibodies from patients with connective tissue diseases bind specific subsets of cellular RNA-protein particles. J Clin Invest. 1982;70:141–7.

164. Reddy R, Tan EM, Henning D, Nohga K, Busch H. Detection of a nucleolar 7-2 ribonucleoprotein and a cytoplasmic 8-2 ribonucleoprotein with autoantibodies from patients with scleroderma. J Biol Chem. 1983;258:1383–94.

165. Rossmanith W, Karwan R. Definition of the Th/ To ribonucleoprotein by RNase P and RNase MRP. Mol Biol Rep. 1993;18:29–35.

166. Ceribelli A, Cavazzana I, Franceschini F, Airò P, Tincani A, Cattaneo R, et al. Anti-Th/To are common antinucleolar autoantibodies in Italian patients with scleroderma. J Rheumatol. 2010;37:2071–5.

167. Hashimoto C, Steitz JA. Sequential association of nucleolar 7-2 RNA with two different autoantigens. J Biol Chem. 1983;258:1379–82.

168. Gold HA, Craft J, Hardin JA, Bartkiewicz M, Altman S. Antibodies in human serum that precipitate ribonuclease P. Proc Natl Acad Sci U S A. 1988;85:5483–7.

169. Gold HA, Topper JN, Clayton DA, Craft JOE. The RNA processing enzyme RNase MRP is identical to the Th RNP and related to RNase P. Science (80-). 1989;245:1377–80.

170. Robertson HD, Altman S, Smith JD. Purification and properties of a specific Escherichia coli ribonuclease which cleaves a tyrosine transfer ribonucleic acid presursor. J Biol Chem. 1972;247:5243–51.

171. Wu J, Niu S, Tan M, Huang C, Li M, Song Y, et al. Cryo-EM Structure of the Human Ribonuclease P Holoenzyme. Cell. 2018;175:1393-1404.e11.

172. Jacobson MR, Cao LG, Taneja K, Singer RH, Wang YL, Pederson T. Nuclear domains of the RNA subunit of RNase P. J Cell Sci. 1997;110 (Pt 7:829–37.

173. Jarrous N, Wolenski JS, Wesolowski D, Lee C, Altman S. Localization in the nucleolus and coiled bodies of protein subunits of the ribonucleoprotein ribonuclease P. J Cell Biol. 1999;146:559–72.

174. Hopper AK, Pai DA, Engelke DR. Cellular dynamics of tRNAs and their genes. FEBS Lett. 2010;584:310–7.

175. Thompson M, Haeusler RA, Good PD, Engelke DR. Nucleolar clustering of dispersed tRNA genes. Science. 2003;302:1399–401.

176. Esakova O, Krasilnikov AS. Of proteins and RNA: The RNase P/MRP family. Rna. 2010;16:1725–47.

177. Chang DD, Clayton DA. A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication. EMBO J. 1987;6:409–17.

178. Li K, Smagula CS, Parsons WJ, Richardson JA, Gonzalez M, Hagler HK, et al. Subcellular partitioning of MRP RNA assessed by ultrastructural and biochemical analysis. J Cell Biol. 1994;124:871–82.

179. Lu Q, Wierzbicki S, Krasilnikov AS, Schmitt ME. Comparison of mitochondrial and nucleolar RNase MRP reveals identical RNA components with distinct enzymatic activities and protein components. Rna. 2010;16:529–37.

180. Ojha S, Malla S, Lyons SM. snoRNPs: Functions in Ribosome Biogenesis. Biomolecules. 2020;10:783.

181. Schmitt ME, Clayton DA. Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in Saccharomyces cerevisiae. Mol Cell Biol. 1993;13:7935–41.

182. Thiel CT, Mortier G, Kaitila I, Reis A, Rauch A. Type and level of RMRP functional impairment predicts phenotype in the cartilage hair hypoplasia-anauxetic dysplasia spectrum. Am J Hum Genet. 2007;81:519–29.

183. Cai T, Aulds J, Gill T, Cerio M, Schmitt ME. The Saccharomyces cerevisiae RNase mitochondrial RNA processing is critical for cell cycle progression at the end of mitosis. Genetics. 2002;161:1029–42.

184. Gill T, Cai T, Aulds J, Wierzbicki S, Schmitt ME. RNase MRP Cleaves the CLB2 mRNA To Promote Cell Cycle Progression: Novel Method of mRNA Degradation. Mol Cell Biol. 2004;24:945–53.

185. Forster AC, Altman S. Similar cage-shaped structures for the RNA components of all ribonuclease P and ribonuclease MRP enzymes. Cell. 1990;62:407–9.

186. Schmitt ME, Bennett JL, Dairaghi DJ, Clayton DA. Secondary structure of RNase MRP RNA as predicted by phylogenetic comparison. FASEB J. 1993;7:208–13.

187. Kipnis RJ, Craft J, Hardin JA. The analysis of antinuclear and antinucleolar autoantibodies of scleroderma by radioimmunoprecipitation assays. Arthritis Rheum. 1990;33:1431–7.

188. Reimer G, Raska I. Immunolocalization of 7-2-Ribonucleoprotein Granular Component of the Nucleolus ' ENG M . TANS '* in the against a variety of nuclear and nucleolar proteins and RNA-protein complexes scleroderma antibodies stain the fibrillar centers of nucleoli and are d. 1988;176:117–28.

189. Karwan RM. Further characterization of human RNase MRP/RNase P and related autoantibodies. Mol Biol Rep. 1998;25:95–101.

190. Van Eenennaam H, Vogelzangs JHP, Lugtenberg D, Van den Hoogen FHJ, Van Venrooij WJ, Pruijn GJM. Identity of the RNase MRP- and RNase P-associated Th/To autoantigen. Arthritis Rheum. 2002;46:3266–72.

191. Mahler M, Fritzler MJ, Satoh M. Autoantibodies to the mitochondrial RNA processing (MRP) complex also known as Th/To autoantigen. Autoimmun Rev. 2015;14:254–7.

192. Mahler M, Gascon C, Patel S, Ceribelli A, Fritzler MJ, Swart A, et al. Rpp25 is a major target of autoantibodies to the Th/To complex as measured by a novel chemiluminescent assay. Arthritis Res Ther. 2013;15:1–9.

193. Bonroy C, Van Praet J, Smith V, Van Steendam K, Mimori T, Deschepper E, et al. Optimization and diagnostic performance of a single multiparameter lineblot in the serological workup of systemic sclerosis. J Immunol Methods. 2012;379:53–60.

194. Okano Y, Medsger TA. Autoantibody to th ribonucleoprotein (nucleolar 7–2 rna protein particle) in patients with systemic sclerosis. Arthritis Rheum. 1990;33:1822–8.

195. Kuwana M, Kimura K, Hirakata M, Kawakami Y, Ikeda Y. Differences in autoantibody response to Th/To between systemic sclerosis and other autoimmune diseases. Ann Rheum Dis. 2002;61:842–6.

196. Meyer OC, Fertig N, Lucas M, Somogyi N, Medsger TA. Disease subsets, antinuclear antibody profile, and clinical features in 127 French and 247 US adult patients with systemic sclerosis. J Rheumatol. 2007;34:104–9.

197. Hamaguchi Y, Hasegawa M, Fujimoto M, Matsushita T, Komura K, Kaji K, et al. The clinical relevance of serum antinuclear antibodies in Japanese patients with systemic sclerosis. Br J Dermatol. 2008;158:487–95.

198. Mitri GM, Lucas M, Fertig N, Steen VD, Medsger TA. A comparison between anti-TH/To- and anticentromere antibody-positive systemic sclerosis patients with limited cutaneous involvement. Arthritis Rheum. 2003;48:203–9.

199. Yamane K, Ihn H, Kubo M, Kuwana M, Asano Y, Yazawa N, et al. Antibodies to Th/To ribonucleoprotein in patients with localized scleroderma. Rheumatology. 2001;40:683–6.

200. Dragon F, Gallagher JEG, Compagnone-Post PA, Mitchell BM, Porwancher KA, Wehner KA, et al. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. Nature. 2002;417:967–70.

201. Bernstein KA, Gallagher JEG, Mitchell BM, Granneman S, Baserga SJ. The small-subunit processome is a ribosome assembly intermediate. Eukaryot Cell. 2004;3:1619–26.

202. Lischwe MA, Ochs RL, Reddy R, Cook RG, Yeoman LC, Tan EM, et al. Purification and partial characterization of a nucleolar scleroderma antigen (Mr = 34,000; pI, 8.5) rich in NG,NG-dimethylarginine. J Biol Chem. 1985;260:14304–10.

203. Ochs RL, Lischwe MA, Spohn WH, Busch H. Fibrillarin: a new protein of the nucleolus identified by autoimmune sera. Biol cell. 1985;54:123–33.

204. Benyamine A, Bertin D, Heim X, Bermudez J, Launay D, Dubucquoi S, et al. Quantification of Antifibrillarin (anti-U3 RNP) Antibodies : A New Insight for Patients with Systemic Sclerosis. 2021;:1–10.

205. Okano Y, Medsger TA. Novel human autoantibodies reactive with 5'-terminal trimethylguanosine cap structures of U small nuclear RNA. J Immunol. 1992;149:1093–8.

206. Tyc K, Steitz JA. U3, U8 and U13 comprise a new class of mammalian snRNPs localized in the cell nucleolus. EMBO J. 1989;8:3113–9.

207. Parker JC, Bunn CC. Sensitivity of the Phadia EliA connective tissue disease screen for less common disease-specific autoantibodies. 2010; table 1:631–4.

208. Peterson LK, Jaskowski TD, Mayes MD, Tebo AE. Detection of anti-U3-RNP / fibrillarin IgG antibodies by line immunoblot assay has comparable clinical significance to immunoprecipitation testing in systemic sclerosis. Immunol Res. 2015. https://doi.org/10.1007/s12026-015-8710-9.

209. Mahler M, Kim G, Roup F, Bentow C, Fabien N, Goncalves D, et al. Evaluation of a novel particle - based multi - analyte technology for the detection of anti - fibrillarin antibodies. Immunol Res. 2021;:239–48.

210. Okano Y, Steen VD, Medsger TAJ. Autoantibody to U3 nucleolar ribonucleoprotein (fibrillarin) in patients with systemic sclerosis. Arthritis Rheum. 1992;35:95–100.

211. Aggarwal R, Lucas M, Fertig N, Oddis C V, Medsger TA. Anti – U3 RNP Autoantibodies in Systemic Sclerosis. 2009;60:1112–8.

212. Steen V, Domsic RT, Lucas M, Fertig N, Jr TAM. A Clinical and Serologic Comparison of African American and Caucasian Patients With Systemic Sclerosis. 2012;64:2986–94.

213. Kuwana M, Okano Y, Kaburaki J, Tojo T, Medsger TAJ. Racial differences in the distribution of systemic sclerosis-related serum antinuclear antibodies. Arthritis Rheum. 1994;37:902–6.

214. Jacobsen S, Halberg P, Ullman S, Venrooij WJVAN. CLINICAL FEATURES AND SERUM ANTINUCLEAR ANTIBODIES IN 230 DANISH PATIENTS WITH SYSTEMIC SCLEROSIS. 1998;:39–45.

215. Tormey VJ, Bunn CC, Denton CP, Black CM. Anti-fibrillarin antibodies in systemic sclerosis. Rheumatology. 2001;40:1157–62.

216. Reveille JD, Fischbach M, Mcnearney T, Friedman AW, Aguilar MB, Lisse J, et al. Systemic Sclerosis in 3 US Ethnic Groups: A Comparison of Clinical, Sociodemographic, Serologic, and Immunogenetic Determinants. 2001;30:332–46.

217. Nandiwada SL, Peterson LK, Mayes MD, Jaskowski TD, Assassi S, Satoh M, et al. Ethnic Differences in Autoantibody Diversity and Hierarchy : More Clues from a US Cohort of Patients with Systemic Sclerosis Ethnic Differences in Autoantibody Diversity and Hierarchy : More Clues from a US Cohort of Patients with Systemic Sclerosis. 2016. https://doi.org/10.3899/jrheum.160106.

218. Wolfe JF, Adelstein E, Sharp GC. Antinuclear antibody with distinct specificity for polymyositis. J Clin Invest. 1977;59:176–8.

219. Treadwell EL, Alspaugh MA, Wolfe JF, Sharp GC. Clinical relevance of PM-1 antibody and physiochemical characterization of PM-1 antigen. J Rheumatol. 1984;11:658–62.

220. Reichlin M, Maddison PJ, Targoff I, Bunch T, Arnett F, Sharp G, et al. Antibodies to a nuclear/nucleolar antigen in patients with polymyositis overlap syndromes. J Clin Immunol. 1984;4:40–4.

221. Gelpi C, Alguero A, Angeles Martinez M, Vidal S, Juarez C, Rodriguez-Sanchez JL. Identification of protein components reactive with anti-PM/Scl autoantibodies. Clin Exp Immunol. 1990;81:59–64.

222. Reimer G, Scheer U, Peters JM, Tan EM. Immunolocalization and partial characterization of a nucleolar autoantigen (PM-Scl) associated with polymyositis/scleroderma overlap syndromes. J Immunol. 1986;137:3802–8.

223. Brouwer R, Pruijn GJM, van Venrooij WJ. The human exosome: An autoantigenic complex of exoribonucleases in myositis and scleroderma. Arthritis Res. 2001;3:102–6.

224. Allmang C, Petfalski E, Podtelejnikov A, Mann M, Tollervey D, Mitchell P. The yeast exosome and human PM-Scl are related complexes of 3' → 5' exonucleases. Genes Dev. 1999;13:2148–58.

225. Briggs MW, Burkard KTD, Butler JS. Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation. J Biol Chem. 1998;273:13255–63.

226. Mian IS. Comparative sequence analysis of ribonucleases HII, III, II PH and D. Nucleic Acids Res. 1997;25:3187–95.

227. Schaeffer D, Tsanova B, Barbas A, Reis FP, Dastidar EG, Sanchez-Rotunno M, et al. The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. Nat Struct Mol Biol. 2009;16:56–62.

228. Coller J, Parker R. Eukaryotic mRNA decapping. Annu Rev Biochem. 2004;73:861–90.

229. Hilleren P, Parker R. Mechanisms of mRNA surveillance in eukaryotes. Annu Rev Genet. 1999;33:229–60.

230. Wang Z, Kiledjian M. Functional link between the mammalian exosome and mRNA decapping. Cell. 2001;107:751–62.

231. Mitchell P, Petfalski E, Shevchenko A, Mann M, Tollervey D. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. Cell. 1997;91:457–66.

232. Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, Tollervey D. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. EMBO J. 1999;18:5399–410.

233. van Hoof A, Lennertz P, Parker R. Yeast Exosome Mutants Accumulate 3′-Extended Polyadenylated Forms of U4 Small Nuclear RNA and Small Nucleolar RNAs. Mol Cell Biol. 2000;20:441–52.

234. Grosshans H, Deinert K, Hurt E, Simos G. Biogenesis of the signal recognition particle (SRP) involves import of SRP proteins into the nucleolus, assembly with the SRP-RNA, and Xpo1p-mediated export. J Cell Biol. 2001;153:745–61.

235. Van Hoof A, Lennertz P, Parker R. Three conserved members of the RNase D family have unique and overlapping functions in the processing of 5S, 5.8S, U4, U5. RNase MRP and RNase P RNAs in yeast. EMBO J. 2000;19:1357–65.

236. Liu Q, Greimann JC, Lima CD. Reconstitution, Activities, and Structure of the Eukaryotic RNA Exosome. Cell. 2006;127:1223–37.

237. Makino DL, Baumgärtner M, Conti E. Crystal structure of an rna-bound 11-subunit eukaryotic exosome complex. Nature. 2013;495:70–5.

238. Kowalinski E, Kögel A, Ebert J, Reichelt P, Stegmann E, Habermann B, et al. Structure of a Cytoplasmic 11-Subunit RNA Exosome Complex. Mol Cell. 2016;63:125–34.

239. Targoff IN, Reichlin M. Nucleolar localization of the PM-Scl antigen. Arthritis Rheum. 1985;28:226–30.

240. Raijmakers R, Schilders G, Pruijn GJM. The exosome, a molecular machine for controlled RNA degradation in both nucleus and cytoplasm. Eur J Cell Biol. 2004;83:175–83.

241. Alderuccio F, Chan EK, Tan EM. Molecular characterization of an autoantigen of PM-Scl in the polymyositis/scleroderma overlap syndrome: a unique and complete human cDNA encoding an apparent 75-kD acidic protein of the nucleolar complex. J Exp Med. 1991;173:941–52.

242. Ge Q, Wu Y, Trieu EP, Targoff IN. Analysis of the specificity of anti-pm-scl autoantibodies. Arthritis Rheum. 1994;37:1445–52.

243. Ge Q, Frank MB, O'Brien C, Targoff IN. Cloning of a complementary DNA coding for the 100-kD antigenic protein of the PM-Scl autoantigen. J Clin Invest. 1992;90:559–70.

244. Ge Q, Frank MB, O'Brien C, Targoff IN. Cloning of a complementary DNA coding for the 100-kD antigenic protein of the PM-Scl autoantigen. J Clin Invest. 1992;90:559–70.

245. Blüthner M, Bautz FA. Cloning and characterization of the cDNA coding for a polymyositis-scleroderma overlap syndrome-related nucleolar 100-kD protein. J Exp Med. 1992;176:973–80.

246. Raijmakers R, Renz M, Wiemann C, Egberts WV, Seelig HP, Van Venrooij WJ, et al. PM-Scl-75 Is the Main

Autoantigen in Patients with the Polymyositis/Scleroderma Overlap Syndrome. Arthritis Rheum. 2004;50:565–9.

247. Oddis C V, Okano Y, Rudert WA, Trucco M, Duquesnoy RJ, Medsger TAJ. Serum autoantibody to the nucleolar antigen PM-Scl. Clinical and immunogenetic associations. Arthritis Rheum. 1992;35:1211–7.

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

249. Targoff IN. Autoantibodies in polymyositis. Rheum Dis Clin North Am. 1992;18:455–82.

250. Selva-O'Callaghan A, Labrador-Horrillo M, Solans-Laque R, Simeon-Aznar CP, Martínez-Gómez X, Vilardell-Tarrés M. Myositis-specific and myositis-associated antibodies in a series of eighty-eight mediterranean patients with idiopathic inflammatory myopathy. Arthritis Care Res. 2006;55:791–8.

251. Brouwer R, Hengstman GJD, Vree Egberts W, Ehrfeld H, Bozic B, Ghirardello A, et al. Autoantibody profiles in the sera of European patients with myositis. Ann Rheum Dis. 2001;60:116–23.

252. Jablonska S, Blaszczyk M. Scleroderma Overlap Syndromes. In: International Journal of Rheumatic Diseases. 1999. p. 85–92.

253. Marguerie C, Bunn CC, Copier J, Bernstein RM, Gilroy JM, Black CM, et al. The clinical and immunogenetic features of patients with autoantibodies to the nucleolar antigen PM-Scl. Medicine (Baltimore). 1992;71:327–36.

254. Targoff IN. Idiopathic inflammatory myopathy: Autoantibody update. Curr Rheumatol Rep. 2002;4:434–41.

255. Kuwana M, Kaburaki J, Okano Y, Tojo T, Homma M. Clinical and Prognostic Associations Based on Serum Antinuclear Antibodies in Japanese Patients with Systemic Sclerosis. Arthritis Rheum. 1994;37:75–83.

256. Mimori T, Akizuki M, Yamagata H, Inada S, Yoshida S, Homma M. Characterization of a high molecular weight acidic nuclear protein recognized by autoantibodies in sera from patients with polymyositis-scleroderma overlap. J Clin Invest. 1981;68:611–20.

257. Mimori T, Hardin JA, Steitz JA. Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. J Biol Chem. 1986;261:2274–8.

258. Parodi A, Rebora A. Anti-Ku antibodies in connective tissue diseases. Report of three cases. J Am Acad Dermatol. 1989;21 2 Pt 2:433–5.

259. Nitta Y, Muramatsu M. A juvenile case of overlap syndrome of systemic lupus erythematosus and polymyositis, later accompanied by systemic sclerosis with the development of anti-Scl 70 and anti-Ku antibodies. Pediatr Dermatol. 2000;17:381–3.

260. Reeves WH. Use of monoclonal antibodies for the characterization of novel DNA-binding proteins recognized by human autoimmune sera. J Exp Med. 1985;161:18–39.

261. Cooley HM, Melny BJ, Gleeson R, Greco T, Kay TW. Clinical and serological associations of anti-Ku antibody. J Rheumatol. 1999;26:563–7.

262. Hirakata M, Mimori T, Akizuki M, Craft J, Hardin JA, Homma M. Autoantibodies to small nuclear and cytoplasmic ribonucleoproteins in Japanese patients with inflammatory muscle disease. Arthritis Rheum. 1992;35:449–56.

263. Hirakata M. Humoral aspects of polymyositis/dermatomyositis. Mod Rheumatol. 2000;10:199–206.

264. Kragelund BB, Weterings E, Hartmann-Petersen R, Keijzers G. The Ku70/80 ring in Non-Homologous End-Joining: easy to slip on, hard to remove. Front Biosci (Landmark Ed. 2016;21:514–27.

265. Fell VL, Schild-Poulter C. The Ku heterodimer: function in DNA repair and beyond. Mutat Res Rev Mutat Res. 2015;763:15–29.

266. Suwa A, Hirakata M, Takeda Y, Okano Y, Mimori T, Inada S, et al. Autoantibodies to DNA-dependent protein kinase. Probes for the catalytic subunit. J Clin Invest. 1996;97:1417–21.

267. Schild-Poulter C, Su A, Shih A, Kelly OP, Fritzler MJ, Goldstein R, et al. Association of autoantibodies with

Ku and DNA repair proteins in connective tissue diseases. Rheumatology (Oxford). 2008;47:165–71.

268. Nilsen TW. The spliceosome: The most complex macromolecular machine in the cell? BioEssays. 2003;25:1147–9.

269. Tarn W-Y, Steitz JA. Highly Diverged U4 and U6 Small Nuclear RNAs Required for Splicing Rare AT-AC Introns. Science (80-). 1996;273:1824–32.

270. Tarn WY, Steitz JA. A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. Cell. 1996;84:801–11.

271. Will CL, Lührmann R. Splicing of a rare class of introns by the U12-dependent spliceosome. Biol Chem. 2005;386:713–24.

272. Turunen JJ, Niemelä EH, Verma B, Frilander MJ. The significant other: Splicing by the minor spliceosome. Wiley Interdiscip Rev RNA. 2013;4:61–76.

273. Bhat P, Honson D, Guttman M. Nuclear compartmentalization as a mechanism of quantitative control of gene expression. Nat Rev Mol Cell Biol. 2021;22:653–70.

274. Zhou Z, Licklider LJ, Gygi SP, Reed R. Comprehensive proteomic analysis of the human spliceosome. Nature. 2002;419:182–5.

275. Van Der Feltz C, Anthony K, Brilot A, Pomeranz Krummel DA. Architecture of the spliceosome. Biochemistry. 2012;51:3321–33.

276. Hamm J, Darzynkiewicz E, Tahara SM, Mattaj IW. The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal. Cell. 1990;62:569–77.

277. Raker VA, Plessel G, Lührmann R. The snRNP core assembly pathway: Identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. EMBO J. 1996;15:2256–69.

278. Pomeranz Krummel DA, Oubridge C, Leung AKW, Li J, Nagai K. Crystal structure of human spliceosomal U1 snRNP at 5.5 resolution. Nature. 2009;458:475–80.

279. Achsel T, Brahms H, Kastner B, Bachi A, Wilm M, Lührmann R. A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. EMBO J. 1999;18:5789–802.

280. Yu YT, Steitz JA. Site-specific crosslinking of mammalian U11 and U6atac to the 5′ splice site of an AT-AC intron. Proc Natl Acad Sci U S A. 1997;94:6030–5.

281. Irina Kolossova and Richard A. Padgett. U11snRNA interact in- vivo with the 5'ss of the U12 INTRONS.pdf. 1997;:227–33.

282. Frilander MJ, Steitz JA. Dynamic exchanges of RNA interactions leading to catalytic core formation in the U12-dependent spliceosome. Mol Cell. 2001;7:217–26.

283. Wassarman KM, Steitz JA. The low-abundance U11 and U12 small nuclear ribonucleoproteins (snRNPs) interact to form a two-snRNP complex. Mol Cell Biol. 1992;12:1276–85.

284. Golas MM, Sander B, Will CL, Lührmann R, Stark H. Major conformational change in the complex SF3b upon integration into the spliceosomal U11/U12 di-snRNP as revealed by electron cryomicroscopy. Mol Cell. 2005;17:869–83.

285. Frilander MJ, Steitz JA. Initial recognition of U12-dependent introns requires both U11/5' splice-site and U12/branchpoint interactions. Genes Dev. 1999;13:851–63.

286. Will CL, Schneider C, Reed R, Lührmann R. Identification of both shared and distinct proteins in the major and minor spliceosomes. Science (80-). 1999;284:2003–5.

287. Will CL, Schneider C, Hossbach M, Urlaub H, Rauhut R, Elbashir S, et al. The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. Rna. 2004;10:929–41.

288. Marabti E El, Malek J, Younis I. Minor intron splicing from basic science to disease. Int J Mol Sci. 2021;22.

289. Bai R, Wan R, Wang L, Xu K, Zhang Q, Lei J, et al. Structure of the activated human minor spliceosome. Science. 2021;371.

290. Norppa AJ, Chowdhury I, van Rooijen LE, Ravantti JJ, Snel B, Varjosalo M, et al. Distinct functions for the paralogous RBM41 and U11/U12-65K proteins in the minor spliceosome. Nucleic Acids Res. 2024;52:4037–52.

291. El Marabti E, Malek J, Younis I. Minor Intron Splicing from Basic Science to Disease. Int J Mol Sci. 2021;22.

292. Sleeman JE, Lamond AI. Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. Curr Biol. 1999;9:1065–74.

293. Kiss T. NEW EMBO MEMBER'S REVIEW: Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. EMBO J. 2001;20:3617–22.

294. Neugebauer KM. Special focus on the Cajal Body. RNA biology. 2017;14:669–70.

295. Raimer AC, Gray KM, Matera AG. SMN - A chaperone for nuclear RNP social occasions? RNA Biol. 2017;14:701–11.

296. Staněk D. Cajal bodies and snRNPs - friends with benefits. RNA Biol. 2017;14:671–9.

297. Meier UT. RNA modification in Cajal bodies. RNA Biology. 2017;14:693–700.

298. Cao T, Rajasingh S, Samanta S, Dawn B, Bittel DC, Rajasingh J. Biology and clinical relevance of noncoding sno/scaRNAs. Trends Cardiovasc Med. 2018;28:81–90.

299. Satoh M, Fritzler MJ, Chan EKL. Antihistone and Antispliceosomal Antibodies. In: Systemic Lupus Erythematosus. Elsevier; 2011. p. 275–92.

300. Fujii T, Mimori T, Hama N, Suwa A, Akizuki M, Tojo T, et al. Characterization of autoantibodies that recognize U4/U6 small ribonucleoprotein particles in serum from a patient with primary Sjogren's syndrome. J Biol Chem. 1992;267:16412–6.

301. Okano Y, Medsger TA. Newly identified U4/U6 snRNP-binding proteins by serum autoantibodies from a patient with systemic sclerosis. J Immunol. 1991;146:535–42.

302. Okano Y, Targoff IN, Oddis C V., Fujii T, Kuwana M, Tsuzaka K, et al. Anti-U5 small nuclear ribonucleoprotein (snRNP) antibodies: A rare anti-U snRNP specificity. Clin Immunol Immunopathol. 1996;81:41–7.

303. Kubo M, Ihn H, Kuwana M, Asano Y, Tamaki T, Yamane K, et al. Anti-U5 snRNP antibody as a possible serological marker for scleroderma-polymyositis overlap. Rheumatology. 2002;41:531–4.

304. Mimori T, Hinterberger M, Pettersson I, Steitz JA. Autoantibodies to the U2 small nuclear ribonucleoprotein in a patient with scleroderma-polymyositis overlap syndrome. J Biol Chem. 1984;259:560–5.

305. Habets W, Hoet M, Bringmann P, Lührmann R, van Venrooij W. Autoantibodies to ribonucleoprotein particles containing U2 small nuclear RNA. EMBO J. 1985;4:1545–50.

306. Gilliam AC, Steitz JA. Rare scleroderma autoantibodies to the U11 small nuclear ribonucleoprotein and to the trimethylguanosine cap of U small nuclear RNAs. Proc Natl Acad Sci U S A. 1993;90:6781–5.

307. Satoh M, Chan JYF, Ross SJ, Ceribelli A, Cavazzana I, Franceschini F, et al. Autoantibodies to survival of motor neuron complex in patients with polymyositis: Immunoprecipitation of D, E, F, and G proteins without other components of small nuclear ribonucleoproteins. Arthritis Rheum. 2011;63:1972–8.

308. Todoroki Y, Satoh M, Kubo S, Kosaka S, Fukuyo S, Nakatsuka K, et al. Anti-survival motor neuron complex antibodies as a novel biomarker for pulmonary arterial hypertension and interstitial lung disease in mixed connective tissue disease. Rheumatology. 2024;63:1068–75.

309. Verma B, Akinyi M V, Norppa AJ, Frilander MJ. Minor spliceosome and disease. Semin Cell Dev Biol. 2018;79:103–12.

310. Xu GJ, Shah AA, Li MZ, Xu Q, Rosen A, Casciola-Rosen L, et al. Systematic autoantigen analysis identifies

a distinct subtype of scleroderma with coincident cancer. Proc Natl Acad Sci U S A. 2016;113:E7526–34.

311. Fertig N, Domsic RT, Rodriguez-Reyna T, Kuwana M, Lucas M, Medsger TA, et al. Anti-U11/U12 RNP antibodies in systemic sclerosis: A new serologic marker associated with pulmonary fibrosis. Arthritis Care Res. 2009;61:958–65.

312. Shah AA, Xu G, Rosen A, Hummers LK, Wigley FM, Elledge SJ, et al. Brief Report: Anti–RNPC-3 Antibodies As a Marker of Cancer-Associated Scleroderma. Arthritis Rheumatol. 2017;69:1306–12.

313. McMahan ZH, Domsic RT, Zhu L, Medsger TA, Casciola-Rosen L, Shah AA. Anti– <scp>RNPC</scp> -3 (U11/U12) Antibodies in Systemic Sclerosis in Patients With Moderate-to-Severe Gastrointestinal Dysmotility. Arthritis Care Res (Hoboken). 2019;71:1164–70.

314. Callejas-Moraga EL, Guillén-Del-Castillo A, Perurena-Prieto J, Sanz-Martínez MT, Fonollosa-Pla V, Lorite-Gomez K, et al. Anti-RNPC-3 antibody predicts poor prognosis in patients with interstitial lung disease associated to systemic sclerosis. Rheumatol (United Kingdom). 2022;61:154–62.

315. Fritzler MJ, Bentow C, Beretta L, Palterer B, Perurena-Prieto J, Sanz-Martínez MT, et al. Anti-U11/U12 Antibodies as a Rare but Important Biomarker in Patients with Systemic Sclerosis: A Narrative Review. Diagnostics (Basel, Switzerland). 2023;13.

316. Betteridge ZE, Woodhead F, Lu H, Shaddick G, Bunn CC, Denton CP, et al. Brief Report: Anti–Eukaryotic Initiation Factor 2B Autoantibodies Are Associated With Interstitial Lung Disease in Patients With Systemic Sclerosis. Arthritis Rheumatol. 2016;68:2778–83.

317. Ceribelli A, Isailovic N, De Santis M, Gorlino C, Satoh M, Selmi C. Autoantibodies as biomarkers for interstitial lung disease in idiopathic inflammatory myositis and systemic sclerosis: The case of anti-eIF2B antibodies. J Transl Autoimmun. 2020;3:100049.

318. Pauling JD, Salazar G, Lu H, Betteridge ZE, Assassi S, Mayes MD, et al. Presence of anti-eukaryotic initiation factor-2B, anti-RuvBL1/2 and anti-synthetase antibodies in patients with anti-nuclear antibody negative systemic sclerosis. Rheumatology (Oxford). 2018;57:712–7.

319. Vulsteke J-B, Coutant F, Goncalves D, Nespola B, De Haes P, Wuyts WA, et al. Detection of anti-eIF2B autoantibodies in systemic sclerosis by immunoprecipitation-mass spectrometry. Rheumatology (Oxford). 2023;62:e216–8.

320. Moon SL, Parker R. EIF2B2 mutations in vanishing white matter disease hypersuppress translation and delay recovery during the integrated stress response. RNA. 2018;24:841–52.

321. Moon SL, Sonenberg N, Parker R. Neuronal Regulation of eIF2α Function in Health and Neurological Disorders. Trends Mol Med. 2018;24:575–89.

322. Kaji K, Fertig N, Medsger TA, Satoh T, Hoshino K, Hamaguchi Y, et al. Autoantibodies to RuvBL1 and RuvBL2: A novel systemic sclerosis-related antibody associated with diffuse cutaneous and skeletal muscle involvement. Arthritis Care Res. 2014;66:575–84.

323. Takahashi T, Nakanishi T, Hamaguchi Y, Tanaka T, Fujimoto N. Case of anti-RuvBL1/2 antibody-positive morphea and polymyositis. The Journal of dermatology. 2017;44:1188–90.

324. Vulsteke J-B, Piette Y, Bonroy C, Verschueren P, Blockmans D, Vanderschueren S, et al. Anti-RuvBL1/2 autoantibodies in patients with systemic sclerosis or idiopathic inflammatory myopathy and a nuclear speckled pattern. Ann Rheum Dis. 2022;81:742–4.

325. Di Pietro L, Chiccoli F, Salvati L, Vivarelli E, Vultaggio A, Matucci A, et al. Anti-RuvBL1/2 Autoantibodies Detection in a Patient with Overlap Systemic Sclerosis and Polymyositis. Antibodies (Basel, Switzerland). 2023;12.

326. Nomura Y, Ueda-Hayakawa I, Yamazaki F, Ozaki Y, Hamaguchi Y, Takehara K, et al. A case of anti-RuvBL1/2 antibody-positive systemic sclerosis overlapping with myositis. European journal of dermatology : EJD. 2020;30:52–3.

327. Jha S, Dutta A. RVB1/RVB2: running rings around molecular biology. Mol Cell. 2009;34:521–33.

328. Adler BL, Boin F, Wolters PJ, Bingham CO, Shah AA, Greider C, et al. Autoantibodies targeting telomere-associated proteins in systemic sclerosis. Ann Rheum Dis. 2021;80:912–9.

329. Vulsteke J-B, Smith V, Bonroy C, Derua R, Blockmans D, De Haes P, et al. Identification of new telomere- and telomerase-associated autoantigens in systemic sclerosis. J Autoimmun. 2023;135:102988.

330. Vogan JM, Collins K. Dynamics of Human Telomerase Holoenzyme Assembly and Subunit Exchange across the Cell Cycle. J Biol Chem. 2015;290:21320–35.

331. Smith EM, Pendlebury DF, Nandakumar J. Structural biology of telomeres and telomerase. Cell Mol Life Sci. 2020;77:61–79.

332. Lee JH, Lee YS, Jeong SA, Khadka P, Roth J, Chung IK. Catalytically active telomerase holoenzyme is assembled in the dense fibrillar component of the nucleolus during S phase. Histochem Cell Biol. 2014;141:137–52.

333. Hal Scofield R. Ro/SSA Autoantibodies. In: Autoantibodies. Elsevier; 2014. p. 233–8.

334. Kapsogeorgou EK, Tzioufas AG. SS-B (La) Autoantibodies. In: Autoantibodies. Elsevier; 2014. p. 247–53.

335. Watanabe T, Ototake Y, Akita A, Suzuki M, Kanaoka M, Tamura J, et al. Clinical features of patients with systemic sclerosis positive for anti-SS-A antibody: a cohort study of 156 patients. Arthritis Res Ther. 2024;26:93.

336. Mondini M, Vidali M, De Andrea M, Azzimonti B, Airò P, D'Ambrosio R, et al. A novel autoantigen to differentiate limited cutaneous systemic sclerosis from diffuse cutaneous systemic sclerosis: The interferon-inducible gene IFI16. Arthritis Rheum. 2006;54:3939–44.

337. McMahan ZH, Shah AA, Vaidya D, Wigley FM, Rosen A, Casciola-Rosen L. Anti-Interferon-Inducible Protein 16 Antibodies Associate with Digital Gangrene in Patients with Scleroderma. Arthritis Rheumatol. 2016;68:1262–71.

338. McMahan ZH, Wigley FM, Casciola-Rosen L. Risk of Digital Vascular Events in Scleroderma Patients Who Have Both Anticentromere and Anti–Interferon-Inducible Protein 16 Antibodies. Arthritis Care Res. 2017;69:922–6.

339. Seelig HP, Ehrfeld H, Renz M. Interferon-γ–inducible protein p16. a new target of antinuclear antibodies in patients with systemic lupus erythematosus. Arthritis Rheum. 1994;37:1672–83.

340. Caneparo V, Cena T, De Andrea M, Dell'Oste V, Stratta P, Quaglia M, et al. Anti-IFI16 antibodies and their relation to disease characteristics in systemic lupus erythematosus. Lupus. 2013;22:607–13.

341. Baer AN, Petri M, Sohn J, Rosen A, Casciola-Rosen L. Association of Antibodies to Interferon-Inducible Protein-16 with Markers of More Severe Disease in Primary Sjögren's Syndrome. Arthritis Care Res. 2016;68:254–60.

342. Uchida K, Akita Y, Matsuo K, Fujiwara S, Nakagawa A, Kazaoka Y, et al. Identification of specific autoantigens in Sjögren's syndrome by SEREX. Immunology. 2005;116:53–63.

343. Rönnblom L, Eloranta ML. The interferon signature in autoimmune diseases. Curr Opin Rheumatol. 2013;25:248–53.

344. Ludlow LEA, Johnstone RW, Clarke CJP. The HIN-200 family: More than interferon-inducible genes? Exp Cell Res. 2005;308:1–17.

345. Unterholzner L, Keating SE, Baran M, Horan KA, Søren B, Sharma S, et al. IFI16 is an innate immune sensor for intracellular DNA. 2011;11:997–1004.

346. Hinchcliff M, Huang CC, Wood TA, Matthew Mahoney J, Martyanov V, Bhattacharyya S, et al. Molecular signatures in skin associated with clinical improvement during mycophenolate treatment in systemic sclerosis. J Invest Dermatol. 2013;133:1979–89.

347. Mahoney JM, Taroni J, Martyanov V, Wood TA, Greene CS, Pioli PA, et al. Systems Level Analysis of Systemic Sclerosis Shows a Network of Immune and Profibrotic Pathways Connected with Genetic Polymorphisms. PLoS Comput Biol. 2015;11.

348. Taroni JN, Martyanov V, Huang CC, Mahoney JM, Hirano I, Shetuni B, et al. Molecular characterization of systemic sclerosis esophageal pathology identifies inflammatory and proliferative signatures. Arthritis Res Ther. 2015;17:1–14.

349. Gugliesi F, Bawadekar M, de Andrea M, Dell'Oste V, Caneparo V, Tincani A, et al. Nuclear DNA Sensor IFI16 as Circulating Protein in Autoimmune Diseases Is a Signal of Damage that Impairs Endothelial Cells through High-Affinity Membrane Binding. PLoS One. 2013;8:1–11.

350. Denton CP, Khanna D. Systemic sclerosis. Lancet. 2017;390:1685–99.

351. Moore KS, 't Hoen PAC. Computational approaches for the analysis of RNA-protein interactions: A primer for biologists. J Biol Chem. 2019;294:1–9.

352. Shah S, Denton CP. Scleroderma autoantibodies in guiding monitoring and treatment decisions. Curr Opin Rheumatol. 2022;34:302–10.

353. Van Eenennaam H, Vogelzangs JHP, Bisschops L, Te Boome LCJ, Seelig HP, Renz M, et al. Autoantibodies against small nucleolar ribonucleoprotein complexes and their clinical associations. Clin Exp Immunol. 2002;130:532–40.

354. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14:417–9.

355. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.

356. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.

357. R Core Team. R: A language and environment for statistical computing (4.1.2) [Computer software]. R Foundation for Statistical Computing. Vienna, Austria. URL https://www.R-project.org/. 2021.

358. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.

359. Benecke H, Lührmann R, Will CL. The U11/U12 snRNP 65K protein acts as a molecular bridge, binding the U12 snRNA and U11-59K protein. EMBO J. 2005;24:3057–69.

360. Zhao J, Peter D, Brandina I, Liu X, Galej WP. Structure of the minor spliceosomal U11 snRNP. bioRxiv. 2023;:2023.12.22.573053.

361. van Nues RW, Granneman S, Kudla G, Sloan KE, Chicken M, Tollervey D, et al. Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. EMBO J. 2011;30:2420–30.

362. Falaleeva M, Pages A, Matuszek Z, Hidmi S, Agranat-Tamir L, Korotkov K, et al. Dual function of C/D box small nucleolar RNAs in rRNA modification and alternative pre-mRNA splicing. Proc Natl Acad Sci U S A. 2016;113:E1625-34.

363. Bizarro J, Charron C, Boulon S, Westman B, Pradet-Balade B, Vandermoere F, et al. Proteomic and 3D structure analyses highlight the C/D box snoRNP assembly mechanism and its control. J Cell Biol. 2014;207:463–80.

364. Höfler S, Lukat P, Blankenfeldt W, Carlomagno T. High-resolution structure of eukaryotic Fibrillarin interacting with Nop56 amino-terminal domain. RNA. 2021;27:496–512.

365. Huang L, Ashraf S, Wang J, Lilley DM. Control of box C/D snoRNP assembly by N(6)-methylation of adenine. EMBO Rep. 2017;18:1631–45.

366. Kishore S, Khanna A, Zhang Z, Hui J, Balwierz PJ, Stefan M, et al. The snoRNA MBII-52 (SNORD 115) is processed into smaller RNAs and regulates alternative splicing. Hum Mol Genet. 2010;19:1153–64.

367. Soeno Y, Taya Y, Stasyk T, Huber LA, Aoba T, Hüttenhofer A. Identification of novel ribonucleo-protein complexes from the brain-specific snoRNA MBII-52. RNA (New York, N.Y.). 2010;16:1293–300.

368. Gupta S, Busch RK, Singh R, Reddy R. Characterization of U6 small nuclear RNA cap-specific antibodies. Identification of gamma-monomethyl-GTP cap structure in 7SK and several other human small RNAs. J Biol Chem. 1990;265:19137–42.

369. Witt LJ, Curran JJ, Strek ME. The Diagnosis and Treatment of Antisynthetase Syndrome. Clin Pulm Med. 2016;23:218–26.

370. Sissler M, González-Serrano LE, Westhof E. Recent Advances in Mitochondrial Aminoacyl-tRNA Synthetases and Disease. Trends Mol Med. 2017;23:693–708.

371. van Zon A, Mossink MH, Scheper RJ, Sonneveld P, Wiemer EAC. The vault complex. Cell Mol Life Sci. 2003;60:1828–37.

372. Berger W, Steiner E, Grusch M, Elbling L, Micksche M. Vaults and the major vault protein: novel roles in signal pathway regulation and immunity. Cell Mol Life Sci. 2009;66:43–61.

373. Moadab F, Wang X, Najjar R, Ukadike KC, Hu S, Hulett T, et al. Argonaute, Vault, and Ribosomal Proteins Targeted by Autoantibodies in Systemic Lupus Erythematosus. J Rheumatol. 2023;50:1136–44.

374. Yoshikatsu Y, Ishida Y, Sudo H, Yuasa K, Tsuji A, Nagahama M. NVL2, a nucleolar AAA-ATPase, is associated with the nuclear exosome and is involved in pre-rRNA processing. Biochem Biophys Res Commun. 2015;464:780–6.

375. Nagahama M, Hara Y, Seki A, Yamazoe T, Kawate Y, Shinohara T, et al. NVL2 is a nucleolar AAA-ATPase that interacts with ribosomal protein L5 through its nucleolar localization sequence. Mol Biol Cell. 2004;15:5712–23.

376. Her J, Chung IK. The AAA-ATPase NVL2 is a telomerase component essential for holoenzyme assembly. Biochem Biophys Res Commun. 2012;417:1086–92.

377. Hiraishi N, Ishida Y, Nagahama M. AAA-ATPase NVL2 acts on MTR4-exosome complex to dissociate the nucleolar protein WDR74. Biochem Biophys Res Commun. 2015;467:534–40.

378. Ishida Y-I, Miyao S, Saito M, Hiraishi N, Nagahama M. Interactome analysis of the Tudor domain-containing protein SPF30 which associates with the MTR4-exosome RNA-decay machinery under the regulation of AAA-ATPase NVL2. Int J Biochem Cell Biol. 2021;132:105919.

379. Mahler M, Bentow C, Aure M-A, Fritzler MJ, Satoh M. Significance of Autoantibodies to Ki/SL as Biomarkers for Systemic Lupus Erythematosus and Sicca Syndrome. J Clin Med. 2022;11.

380. Yamasaki Y, Narain S, Hernandez L, Barker T, Ikeda K, Segal MS, et al. Autoantibodies against the replication protein A complex in systemic lupus erythematosus and other autoimmune diseases. Arthritis Res Ther. 2006;8:R111.

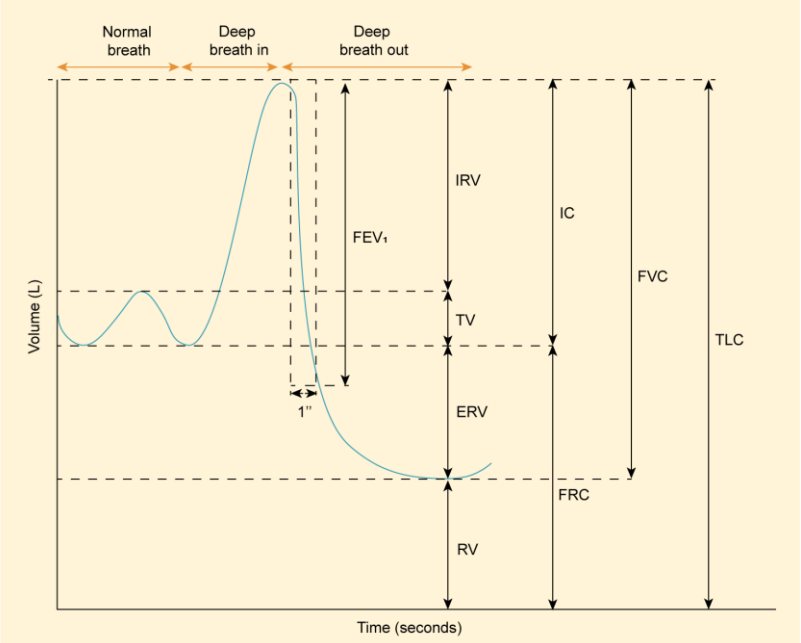
381. Franks JM, Martyanov V, Cai G, Wang Y, Li Z, Wood TA, et al. A Machine Learning Classifier for Assigning Individual Patients With Systemic Sclerosis to Intrinsic Molecular Subsets. Arthritis Rheumatol. 2019;71:1701–10.

382. Taroni JN, Mahoney JM, Whitfield ML. The Mechanistic Implications of Gene Expression Studies in SSc: Insights From Systems Biology. Curr Treat Options Rheumatol. 2017;3:181–92.

383. Nihtyanova SI, Denton CP. Autoantibodies as predictive tools in systemic sclerosis. Nat Rev Rheumatol. 2010;6:112–6.

ANNEXE 1 – Pulmonary function tests, basic concepts

- Spirometry allows to determine different parameters, being the most important ones the forced vital capacity (FVC), the forced expiratory volume in one second (FEV_1) and the FEV_1/FVC ratio.



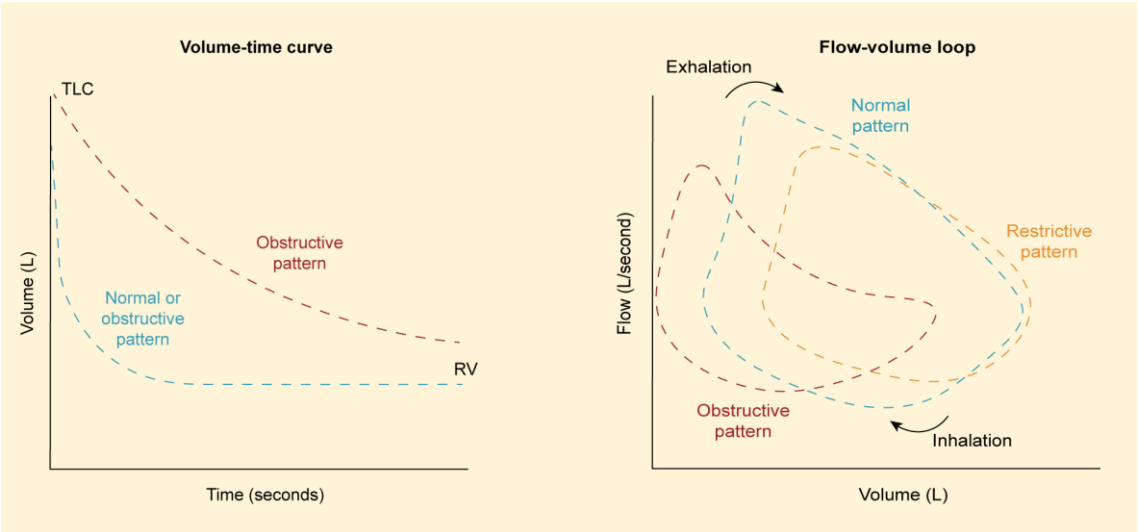
FEV_1: forced expiratory volume in 1 second
IRV: inspiratory reserve volume
TV: tidal volume
ERV: expiratory reserve volume
RV: residual volume
IC: inspiratory capacity
FRC: functional residual capacity
FVC: forced vital capacity
TLC: total lung capacity

- Diffusing capacity for carbon monoxide (DLco): the diffusing capacity is a measure of the ability of the lungs to transfer gas. Diffusion in the lungs is most efficient when the surface area for gas transfer is high and the blood is readily able to accept the gas being transferred. It is thus decreased in:
    - Conditions that minimize the ability of the blood to accept and bind the gas that is diffusing, like anaemia.
    - Conditions that decrease the surface area of the alveolar-capillary membrane, like emphysema or pulmonary embolism.
    - Conditions that alter the membrane's permeability or increase its thickness, like endothelial cell proliferation.
- Residual volume and total lung capacity: even after one exhales as long and as hard as possible, some air remains in the lungs; this is called the residual volume. The residual volume plus the FVC equals the total lung capacity (TLC). The residual volume cannot be measured by spirometry and requires special tests:
    - The patient breathes an inert gas such as helium, and the concentration of it is measured in the expired air, from which the residual volume is calculated.
    - The patient sits in an airtight booth in which the pressure is measured as he or she breathes.
- Obstructive/restrictive pattern:
    - In obstructive lung diseases, airway obstruction causes an increase in resistance. During normal breathing, there is no need for a high difference in pressure, and therefore, the pressure/volume relationship is no different from a normal lung. However, for rapid breathing, a greater difference in pressure is needed to be able to overcome the resistance to flow, and consequently, the volume of each breath gets smaller in obstructive disease.

- In restrictive lung diseases, the compliance of the lung is reduced, which increases the stiffness of the lung and limits its expansion. In these cases, a greater difference in pressure than normal is required to give the same increase in volume even in normal breathing.
- For recognizing the obstructive or restrictive pattern, the first step is to assess the FEV_1/FVC ratio. If this ratio is less than the lower limit of normal for the patient, an obstructive defect is present. If this ratio is greater than the lower limit of normal, then either the spirometry test is normal, or a restrictive defect is present. If the FVC is less than the lower limit of normal, a restrictive defect is suggested. The total lung capacity can confirm the presence of restriction if this value is less than the predicted lower limit of normal.

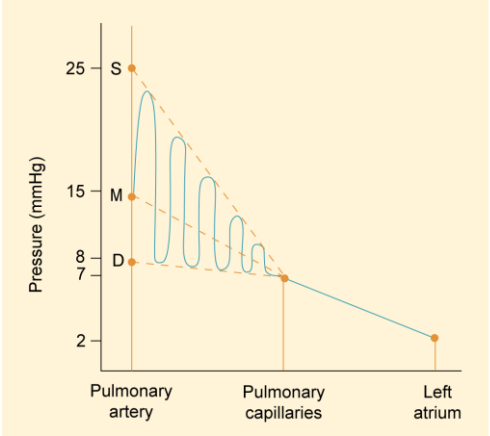
Measurement	Obstructive pattern	Restrictive pattern
FVC	Decreased or normal	Decreased
FEV1	Decreased	Decreased or normal
FEV_1/FVC	Decreased	Normal
TLC	Normal or increased	Decreased

- A volume-time curve and flow-volume loop are usually included with the report of spirometry which allows a quick recognition of restrictive and obstructive patterns. In a proper FVC manoeuvre, volume decline should be at its highest near the start of exhalation. This is graphed in the volume-time curve. In the case of an obstructive pattern, the volume decline will be constant during the exhalation, and a high flow rate at the start will not be observed. Similarly, flow rates are highest near the total lung capacity, where the FVC manoeuvre begins, and the decline until the residual volume (where the FVC manoeuvre ends). This is graphed as the flow-volume loop. In an obstructive pattern, the flow rates will be diminished in comparison with a healthy person, while in a restrictive pattern, the flow rates and also the volumes will be diminished.



ANNEXE 2 – Haemodynamics, basic concepts

- Right ventricle afterload: the “load” against which the right ventricle ejects blood, which corresponds with pressure in the pulmonary artery and the right ventricle systolic pressure. There is an inverse relationship between afterload and systolic performance, meaning that cardiac output decreases as the afterload on the heart increases.
- Pressures in the pulmonary artery: during systole, the pressure in the pulmonary artery is essentially equal to the pressure in the right ventricle. However, after the pulmonary valve closes at the end of systole, the ventricular pressure falls precipitously, whereas the pulmonary arterial pressure falls more slowly as blood flows through the lungs. The systolic pulmonary arterial pressure normally averages about 25mmHg, and the diastolic pulmonary arterial pressure is about 8mmHg, while the mean pulmonary arterial pressure (mPAP) is 15mmHg.



- Pulmonary arterial wedge pressure (PAWP): it usually is not feasible to measure someone’s left atrial pressure using a direct measuring device because it is difficult to pass a catheter through the heart chambers into the left atrium. However, the left atrial pressure can be estimated with moderate accuracy by measuring the so-called pulmonary arterial wedge pressure. This is measured by inserting a catheter first through a peripheral vein to the right atrium, then through the right side of the heart and through the pulmonary artery into one of the small branches of the pulmonary artery, and finally pushing the catheter until it wedges tightly in the small branch. The pressure measured through the catheter, called the “wedge pressure” is about 5-6mmHg. Because all blood flow has been stopped in the small wedged artery, and because the blood vessels extending beyond this artery make a direct connection with the pulmonary capillaries, this wedge pressure is usually only 2-3mmHg higher than the left atrial pressure. When the left atrial pressure rises to high values, the pulmonary wedge pressure also rises. Therefore, wedge pressure measurements can be used to estimate changes in pulmonary capillary pressure and left atrial pressure in patients with congestive heart failure.
- Cardiac output: the amount of blood pumped by the heart each minute.
- Total pulmonary vascular resistance (PVR): is defined as the resistance against blood flow from the pulmonary artery to the left atrium. It is most commonly modelled using a modification of Ohm’s law described below, and a value of 0.25-1.6mmHg·min/L or WU (Wood Units) is considered as normal.

$$PVR = \frac{\text{Input pressure (mPAP) - Out pressure (PAWP)}}{\text{Total blood flow (cardiac output)}}$$