XIST ribonucleoproteins promote female sex-biased autoimmunity

Diana R. Dou[1], Yanding Zhao[1], Julia A. Belk[1], Yang Zhao[1], Kerriann M. Casey[2], Derek C. Chen[1], Rui Li[1], Bingfei Yu[1], Brian T. Abe[1], Katerina Kraft[1], Cecilia Hellström[3], Ronald Sjöberg[3], Sarah Chang[4], Allan Feng[4], Daniel W. Goldman[5], Ami A. Shah[5], Michelle Petri[5], Frederick M. Wigley[5], Lorinda S. Chung[4], David F. Fiorentino[6], Emma K. Lundberg[3,7], Anton Wutz[8], Paul J. Utz[4], Howard Y. Chang[1,9,*]

1 Center for Personal Dynamic Regulomes and Program in Epithelial Biology, Department of Dermatology, Stanford University School of Medicine, Stanford, CA
2 Department of Comparative Medicine, Stanford University, Stanford, CA
3 Autoimmunity and Serology Profiling, Division of Affinity Proteomics, Department of Protein Science, KTH Royal Institute of Technology, SciLifeLab, Stockholm, Sweden
4 Department of Medicine, Division of Immunology and Rheumatology, Stanford University School of Medicine, Stanford, CA
5 Johns Hopkins University School of Medicine, Division of Rheumatology, Baltimore, MD
6 Department of Dermatology, Stanford University School of Medicine, Redwood City, CA
7 Departments of Bioengineering and Pathology, Stanford University, Stanford, CA
8 Department of Biology, Institute of Molecular Health Sciences, Swiss Federal Institute of Technology, ETH Hönggerberg, Zurich, Switzerland
9 Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305, USA.

*Correspondence to: H.Y.C. at howchang@stanford.edu

SUMMARY

Autoimmune diseases disproportionately affect females more than males. The XX sex chromosome complement is strongly associated with susceptibility to autoimmunity. XIST long noncoding RNA (lncRNA) is expressed only in females to randomly inactivate one of the two X chromosomes to achieve gene dosage compensation. Here, we show that the XIST ribonucleoprotein (RNP) complex, comprised of numerous autoantigenic components, is an important driver of sex-biased autoimmunity. Inducible transgenic expression of a non-silencing form of *Xist* in male mice introduced Xist RNP complexes and sufficed to produce autoantibodies. Male SJL/J mice expressing transgenic Xist developed more severe multiorgan pathology in pristane-induced model of lupus than wild-type males. Xist expression in males reprogrammed T and B cell population and chromatin states to more resemble wild type females. Human patients with autoimmune diseases displayed significant autoantibodies to multiple components of XIST RNP. Thus, a sex-specific lncRNA scaffolds ubiquitous RNP components to drive sex-biased immunity.

INTRODUCTION

Autoimmune diseases are the third-most prevalent disease category, outpaced only by cancer and heart disease[1]. Four out of five patients with autoimmune diseases are female. For instance, in systemic lupus erythematosus (SLE), the disproportionate ratio of affected females to males is as high as 9:1; the ratio of female to male patients in Sjogren's disease is 19:1 [2,3] In-depth studies have identified estrogen[4] and specific X-linked genes, such as *TLR7* in SLE[5-8], as contributors to autoimmunity in specific diseases. However, targeting female hormones and individual genes have limited efficacy and/or are clinically risky[9] and the risk factors underlying autoimmune diseases in aggregate remain unresolved. Identical twin studies have also shown varying degrees of autoimmune disease penetrance, suggesting a genetic disposition that is also reliant on environmental factors[10]. Hence, adjuvant triggers[11] rather than, or in addition to, gene expression may be initiators of autoimmune disease development. Evidence against hormones

as principal drivers of autoimmunity has emerged from studies with Klinefelter syndrome (XXY) males and sex chromosome complement studies in mice, which demonstrate that autoimmunity risk correlates with the XX complement irrespective of phenotypic sex and hormonal background.[12,13]

Individuals with the XX sex complement express the lncRNA, XIST, on one of the X chromosomes. XIST is critical in random X inactivation, spreading from the X-inactivation center and recruiting SPEN to activate silencing of the X chromosome. In mESCs, Xist associates with 81 unique binding proteins to form an RNP complex[14]. Over a third of the proteins complexing with the Xist lncRNA are also known autoantigens associated with autoimmune diseases. Studies in SLE patients and mice demonstrated that DNA-autoantibody and RNA-autoantigen immune complexes, such as Sm/RNP and U1A, activate the TLR7, TLR8 and TLR9 pathways of the innate immune system[15-17]. The XIST RNP, comprised of a lncRNA, bound autoantigenic proteins, and tethered to pieces of genomic DNA, presents qualities resembling nucleic acid-autoantigen immune complexes.

In order to study the impact of the XIST RNP, independent of hormonal background, in autoimmune predilection, we utilized an inducible, non-silencing form of Xist introduced into an autosome in the autoimmune-resistant C57BL/6J and autoimmune-prone SJL/J strain backgrounds. Inducing transgenic Xist RNP formation in male animals allowed the study of this female-specific lncRNA in a male background in the context of a chemically-induced SLE model. Both increased disease severity and elevated autoreactive lymphocyte pathway signatures were observed in the mouse models of pristane-induced SLE. Concurrently, we designed an antigen array to test autoimmune patient seroactivity to XIST-associating proteins and detected significant reactivity towards multiple components of the XIST RNP. Altogether, our data point to a significant role for the *Xist* RNP as a driver for autoimmunity that may underly the sex-biased female preponderance for developing autoimmune diseases.

RESULTS

Developing the TetOP-DRepA-Xist transgenic mouse to model autoimmune diseases

Our study is the first to link *Xist* ribonucleoprotein particles as potential triggers underlying the greater prevalence of autoimmune diseases in females. While it is a well-documented fact that females are much more prone to autoimmune diseases than males, previous studies have examined differences in gene contribution and hormonal background rather than investigate the lncRNA, *Xist*, as a stimulant or adjuvant. *Xist* is a long noncoding RNA (lncRNA), expressed only in females that is required for X chromosome silencing and forms ribonucleoprotein particles (RNPs) containing many autoantigens that are associated with autoimmune diseases[14], suggestive of a role for *Xist* RNPs in female-biased autoimmunity.

In collaboration with the Wutz lab, we developed a TetOP-**D**RepA-Xist transgenic mouse that enables inducible expression of *Xist* in male animals. A truncation of Xist removes the A-repeat (RepA) element required for the silencing activity of Xist[18], but does not ablate chromosome coating or particle formation. Expression of the **D**RepA mutant Xist is controlled through the Tet-operon (TetOP) promoter. The insert is located in the Col1A1 locus on chromosome 11 (**Figure 1A**). Since Xist is expressed on only 1 of the 2 X chromosomes, we used mice heterozygous for TetOP-**D**RepA-Xist (denoted as tgXist onwards) in our studies. After only 2 weeks of doxycycline administration in heterozygous tgXist male mice, expression of *tgXist* was detectable through *Xist* qPCR and FISH (**Figure 1B,C**). Backcrossing of the inducible tgXist transgene into common mouse strains enables the study of *Xist* in male animals in multiple autoimmune disease modles.

To test our hypothesis that expression of *XIST* ribonucleoprotein particles (RNPs) creates a higher autoimmune background and increased susceptibility to autoimmune diseases evinced

by females[2,3] and Klinefelter syndrome males[12], male and female C57BL/6J mice heterozygous for the DRepA-Xist transgene were injected with pristane to chemically induce systemic lupus erythematosus (SLE)[19-22] and evaluated for disease 1 year post-treatment (**Figure 1E**). Since C57BL/6J mouse[19,23], the most common and widely used murine experimental background, are autoimmune resistant and male mice are not expected to develop pristane-induced SLE[19], we first tested the effect of expressing *tgXist* compared to wild-type males to determine if expression of *Xist* alone is sufficient to induce disease.

The pristane-injected transgenic cohorts consisted of a female control untreated for doxycycline (female), male control untreated for doxycycline (tg male), and the test group of male mice continuously treated with doxycycline to induce *tgXist* expression (male+dox) (**Figure 1F**). Measurement of anti-nuclear antibody (ANAb) production showed elevation of 24 weeks post-pristane injection in all transgenic cohorts tested, including the transgenic male mice not administered doxycycline (**Figure 1G**). Comparison of *tgXist* levels through qPCR between control transgenic male mice and wild-type C57BL/6 male mice suggests transgenic male mice already express a basal level of *tgXist* due to promoter leakage (**Figure 1D**). This level is low compared to the transgenic female control mice (**Figure 1B**), but may be sufficient to affect changes in autoimmunity. Since C57BL/6J mice are autoimmune-resistant, 52 weeks post-induction was the terminal timepoint for pristane-induced SLE[19]. However, the resulting advanced age of the mice combined with their natural autoimmune resistance made detection of disease-specific pathophysiological changes difficult to determine (data not shown).

Despite the lack of significant physiological SLE, we detected elevated presence of autoantibodies to several SLE-associated antigens (smRNP, SMITH, U1-A, and U1-68) 52-weeks post-pristane treatment with the highest median elevation in the tgXist-induced male cohort. Since elevated antigen levels were detected across all three transgenic cohorts (female, male, and male+dox), we also injected wild-type C57BL/6J male mice with pristane (WT Male) or a mock injection of PBS (WT Male +PBS) to test for pristane effects on a wild-type background without basal levels of *tgXist*. While some elevated autoantibodies were also detected in one of four wild-type pristane-injected male mice, it does not approach the higher frequency and level of elevation observed in the tgXist-induced male group (**Figure 1H**).

tgXist-expressing male mice correlate to a more female-like splenic CD4+ T-cell profile

While pristane-induced SLE in C57BL/6J mice did not physically manifest disease (data not shown), splenic $CD4^+$ T-cells (a critically important cell type for balancing and driving autoimmunity[24-26]) from *tgXist*-expressing male mice correlated closer to females than to control (tgXist-non-induced/wild-type) males at several levels. Changes in the transcriptional regulation and gene expression of CD4+ splenic T-cells were assessed using ATAC-seq and RNA-seq of multiple animals per cohort (**Figure 2A**). Global comparison of ATAC-seq differential peaks showed several differences between non-tgXist/Xist-expressing animals compared to tgXist-expressing males and wild-type Xist-expressing females, but very few differential peaks between females compared to males expressing tgXist (**Figure 2B**). PCA comparison of female mice, male mice, and male tgXist-expressing mice (male+dox) showed a separation of mice most highly correlated with sex (**Figure 2C**). Two of six tgXist-expressing male mice showed female-like skewing (circled in **Figure 2C**), in keeping with the expected female penetrance percentage and subsequent analysis focused on these two mice to determine the epigenetic and genomic changes contributing to the shift from male-like to female-like.

ATAC-seq revealed distinct clusters of accessibility corresponding to tgXist/Xist-expressing and non-expressing mice (**Figure 2D**). tgXist-induced and pristane-injected males displayed similar chromatin accessibility to pristane-injected female positive control mice, with higher accessibility in cluster 2, and were distinct from male negative control groups (*tgXist* non-induced and wild-type) that displayed higher accessibility in cluster 1 (**Figure 2D**). Interestingly,

the top Reactome categories in Cluster 2, associated with *tgXist*-expressing males and Xist-expressing females, displayed high TLR pathway signatures (**Figure 2E**) not present in Cluster 1 (Supplemental **Figure 2E**). Notably, TLR9, a pathogen sensor in the innate immune pathway that is highly active in SLE, is significantly more accessible in females and tgXist-induced males (**Supplemental Figure 2G**). RNA-seq profiles of gene expression displayed a similar trend and clustering grouped by tgXist/Xist expression with some variability within cohorts (**Figure 2F**), suggesting that ATAC-seq provides a more consistent profile of cells in transition. Comparison of dysregulated genes from ATAC- and RNA-seq revealed an overlap of 384 genes (**Figure 2G**). CIBERSORT[27] deconvolution of gene expression signatures also identified a clear segregation of *Xist*-expressing and non-expressing mice. In particular, *tgXist/Xist*-expressing males and females contained more CD4 memory T-cells while control males expressed greater proportions of naïve T-cells (**Figure 2H**).

tgXist/Xist-expressing animals develop more pathologically severe pristane-induced SLE

We next tested the effect of tgXist expression in the autoimmune-prone SJL/J mouse background, a widely used strain in multiple autoimmune disease models, including pristane-induced SLE[20,28,29]. Pristane-induced SLE in SJLJ/J mice exhibit many characteristics of human SLE (i.e., autoantibody development, TLR7 upregulation, organ involvement) and demonstrate sex-related bias (i.e., XX vs XY)[13,20]. In sex-matched pristane-induced SLE studies, SJL/J female mice display earlier mortality, more severe nephritis, higher levels of autoantibodies, and are 3.4x more likely to die than male mice[20]. This disease model is also more reflective of most human patients than spontaneous SLE models that are restricted to specific genetic mutations[30-33].

A one-time 0.5 mL intraperitoneal injection of pristane into 8-10 week old mice to induce SLE in *tgXist* and wild-type (WT) mice and continuous ingestion of doxycycline in drinking water to activate *tgXist* in selected mice (**Figure 3A**). The experimental cohorts include the wild-type female pristane-treated positive control, negative controls of mock-injected *tgXist*- and WT males, the WT male treatment control, and the test group of *tgXist*-expressing pristane-treated males (**Figure 3B**). Since previous literature determined that WT SJL/J females display mortality as early as 16 weeks post-pristane injection[20], we selected the terminal timepoint of 16 weeks post-injection to avoid premature loss of mice.

Since SLE is a systemic disease, we used H&E staining to assess the pathophysiology of multiple immune-related organs at the terminal collection time (**Figure 3C,D**). In all pristane-treated cohorts, pristane injection caused lipogranulomas in adipose tissues, lymph node hyperplasia and medullary plasmocytosis, as well as varying degrees of extramedullary hematopoiesis and lymphoid expansion in the spleen. However, pristane injection coupled with *tgXist* expression in males or *Xist* in WT female resulted in greater incidence and severity of glomerulonephritis (kidney), hepatic lipogranulomas (liver), and pulmonary hemorrhage and lymphohistiocytic alveolitis (lung), which is reflective of disease damage in the kidney, liver, and lungs observed in severe SLE patients[34] (**Figure 3C**). Manifestation of mild signs of early SLE in some pristane-injected male controls was not unexpected because WT SJL/J males are expected to develop milder disease and display later mortality at 24 weeks post-injection[20].

Based on the lowest female total pathology score, 10 was set as the boundary to categorize diseased and normal limits in Fisher's Exact Test of Independence. We determined that the total systemic damage across all assessed organs significantly increased in pristane-treated *tgXist*-expressing males compared to pristane-treated WT males as well as untreated male controls (statistical significance highlighted in orange). Comparatively, there was no significant difference in total damage scores between pristane-treated WT males and the untreated male controls. Notably, the greatest statistical difference was preserved in the comparison between the female positive controls and the pristane-treated WT males as it was between the positive female and negative male controls (**Figure 3E**, highlighted in orange).

Sera collected at treatment start (day 0, baseline) and terminally (16 weeks post-treatment) were assessed for reactivity to SLE antigens using the Luminex bead-based antigen array[35]. Both the female (WT *Xist*) and *tgXist* males treated with pristane expressed higher reactivity levels to SLE antigens, including the subunits of the ribosomal P antigen, U1-snRNP, and C1q complexes (**Figure 3F**). Based on these studies, similar to female SJL/J mice, *tgXist*-expressing male mice experience more severe disease than their wild-type male counterparts in the pristane-induced SLE model.

Single Cell Gene Expression Analysis reveals elevation of atypical B-cells and suppression of T-cell modulators in diseased *tgXist/Xist*-expressing animals

In order to gain insight to the genes and cellular processes underpinning the heightened autoimmunity in the tgXist/Xist-expressing mice, we created single cell gene expression libraries of CD45+ sorted splenic cells from pristane-injected mice in the SJL/J strain. Using the aforementioned total pathology score boundary of 10 (**Figure 3B,D**), we divided the pristane-induced tgXist-expressing males into disease high (total pathology score ≧ 10) and low (<10) disease groups. Sequencing libraries were made from representative mice selected from the tgXist male high disease (n=4), tgXist male low disease (n=4), and two pristane-injected control groups: wild-type females (n=4) and wild-type males (n=3). Due to high mitochondrial/ribosomal RNA content, 2 of the wild-type females were excluded from the gene expression analysis.

The single cell gene expression UMAP formed 12 distinct clusters and were assigned identities based on key transcription factors and markers (**Figure 4A, Supplemental Figure 4B**). A key advantage of the single cell multiomics approach is the ability to interrogate all the hematopoietic lineages within the spleen in a single assay. While our bulk ATAC- and RNA-sequencing in the C57BL/6J strain was restricted to CD4+ T-cells, the single cell libraries of the whole CD45+ splenic population segregated clearly into B-cell, T-cell, NK cell, myeloid, and erythroid lineages.

Within the B-cell cluster, the tgXist male high disease group overlay with the female wild-type mice, distinct from the wild type male and the relatively unaffected tgXist male low disease animals (**Figure 4B**). Since B cells produce the majority of autoantibodies and proinflammatory cytokines that characterize autoimmune pathogenesis[36,37], the high correlation of B-cells, specifically, of the tgXist- diseased males with those from positive disease control females provides another layer of support for the hypothesis that Xist complexes mediate an environment of higher autoimmunity. The tgXist low disease and wild-type males grouped closely together and most of the differences between the tgXist high disease and wild-type female mice lay in the sex chromosomes or ribosomal protein genes. Inspection of the differentially expressed genes comparing the SLE-affected tgXist male and female mice with the wild-type males (**Figure 4C)** revealed that the majority of both downregulated and upregulated genes were shared between the first two groups (**Figure 4D**). Particularly significant genes included the upregulation of the atypical B-cell marker, Zeb2 tgMhigh vs WT M p-value= 3.08E-69, WT F vs WT M p-value= 1.17E-35and CD22 (tgMhigh vs WT M p-value= 7.63E-46, WT F vs WT M p-value= 2.59E-36), a receptor associated with pathogenic B cells[38] critical for B-cell proliferation and B cell receptor signaling (**Figure 4E**). Concurrently, Siglec-g, a receptor associated with promoting B-cell self-tolerance, deficiency of which is associated with increased B-1 cells and multiple autoimmune diseases[39-41], was significantly downregulated (tgMhigh vs WT M p-value= 1.73E-101, WT F vs WT M p-value= 1.01E-43) (**Figure 4E**). Also downregulated were complement receptor 2 (Cr2, tgMhigh vs WT M p-value= 1.30E-153, WT F vs WT M p-value=3.70E-44) and the paralog to its alternatively spliced form, Cr1l (tgMhigh vs WT M p-value= 4.41E-33, WT F vs WT M p-value= 1.35E-33), both of which are important for suppressing autoimmunity[42,43].

While the UMAP grouping within the T-cell cluster was less distinctly demarcated than the B-cell cluster (**Figure 4B**), the mechanistic patterning was clear in the shared downregulated programs (**Supplemental Figure 4D**). Multiple key T-cell regulation and self-tolerance genes

were downregulated in the tgXist male high disease and female cohorts compared to the wild-type males. Among the significantly downregulated genes were the glucocorticoid receptor Nr3c1 (tgMhigh vs WT M p-value= 1.37E-33, WT F vs WT M p-value= 1.33E-63) involved in Treg modulation of inflammation[44], Cd37 (tgMhigh vs WT M p-value= 4.60E-12, WT F vs WT M p-value= 1.81E-10) which regulates proliferation[45] and complement-mediated apoptosis of autoreactive T-cells[45], the CD4 T-cell immune regulator Cd52[46] (tgMhigh vs WT M p-value= 7.24E-13, WT F vs WT M p-value= 2.22E-72), and invariant chain Cd74 involved in trianing antigen immunity in T-cells[47] (tgMhigh vs WT M p-value= 8.15E-37, WT F vs WT M p-value= 2.27E-96) (representative plots, **Figure 4F**).

At the single cell gene expression level, the severity of disease in tgXist-/Xist-expressing animals may be driven by increased atypical B-cell activity and decreased immune modulatory programs in both B-and T-cells. Combined with the heightened pathophysiology scores and increased autoantigen levels, the gene expression data corroborates a role for Xist RNPs in the development of increased and/or more severe autoimmunity in the pristane-induced SLE model.

Autoimmune Patients Display Serum Reactivity to Multiple Antigens in the XIST RNP

Although mouse models provide critical insights into biological mechanisms, there remain limitations in translating discoveries in mice to humans. In order to test whether the XIST RNP is immunogenic in humans, we obtained de-identified patient serum from dermatomyocitis (DMM), scleroderma (SSc) and SLE to test for reactivity to XIST complex proteins[14,48]. SciLifeLab created protein fragments called Protein Signature Tags (PrESTs) for 116 of our proteins of interest and 51 for proteins used as clinical markers for DMM, SSc and SLE. 6 of the clinical protein antigens were also detected in the XIST ChIRP-MS lists (SSB, HARS1, SNRPC, SNRPD1, SNRPD2, SNRPF). When possible, multiple PrESTs spanning different regions of each protein were used (**Figure 5A**).

As a general population control, we obtained serum from research byproduct donations to the Stanford Blood Center. However, since the mean and median age of these donors were 58 and 61 years at the time of donation, these donors may express some autoreactivity due to advanced age. Nevertheless, autoimmune patients were significantly more reactive to 61 proteins, 17 of which were disease markers, 3 of which overlapped with XIST proteins, and the remaining antigens from the XIST complex list (Figure 5B). Distinct reactive antigens for each autoimmune disease arose when grouped by disease (**Figure 5C**) and 9 antigens were shared among all 3 diseases (**Figure 5D**). Of these, TRIM33, or TIF1-γ, is a clinical marker for autoimmune disease (DMM)[49]. The other 8 were XIST-complex proteins, several of which have recently discovered roles in autoimmunity. HMGB1 was recently identified as an autoantibody in SLE/Sjogren that may be another SS-protein in the SS-A/SS-B family[50,51], HNRNPK is an autoantibody in a subset of Raynaud's disease[52,53] and aplastic anemia, SAFB is a novel autoantibody in connective tissues detected in interstitial lung disease, and XPO5 is in the SSB/La processing pathway[54].

Several antigens from the XIST RNPs elicited significant reactivity in sera from DMM, SSc, and SLE patients above the general population blood donor baseline. It is possible that a younger general population baseline would reveal more significant autoreactivity that may be masked by aging in our donor group. Increasing the number of patient sera, particularly from the DMM cohort, would also increase the power of our comparative study. Nevertheless, our study was able to confirm serum reactivity towards multiple components of the XIST RNP complex above baseline level in three different types of autoimmune patients.

DISCUSSION

Our study is the first to link *Xist* ribonucleoprotein particles as triggers underlying the greater prevalence of autoimmune diseases in females. Although it is a well-documented fact that females are more prone to autoimmune diseases than males, previous studies primarily examined

differences in gene contribution and hormonal background. And, while more recent *Xist*-involvement studies address altered X-inactivation and the subsequent impact of *Xist* silencing mechanisms on related X-linked genes[48,55,56], this study is the first to investigate the immunogenicity of *Xist* itself as part of an immune complex. We have shown that expression of Xist RNPs in male mice is sufficient to increase disease severity and change the expression and epigenomic profiles of both the B-cell and T-cell effectors of SLE pathogenesis.

There are more than 100 known autoimmune diseases that, in aggregate, combine to afflict as many 50 million afflicted Americans and comprise one of the top ten leading causes of death for women under the age of 65[57]. Worryingly, cases are increasingly yearly on a global scale and a recent seropositivity studies revealed a steep rise of increasing ANAb reactivity[58,59]. Understanding the risk factors and drivers of autoimmunity has become even more critical in the race to develop effective therapies and sensitive diagnostics specific to each autoimmune disease. However, the high heterogeneity within autoimmune diseases and overlapping traits across diseases have limited our ability to tailor effective therapies and sensitive diagnostics specific to each autoimmune disease[60]. Our discovery of seropositivity towards multiple XIST-associating proteins in autoimmune patients introduces a novel antigen set with great clinical potential for enhancing diagnostics and disease monitoring, as autoantigens are often detected prior to or early in disease onset[61,62]. In addition to serving as effective biomarkers for disease, studies in SSc have also demonstrated the effectiveness of autoantigen analysis in patient stratification and identifying pathogenic pathways[63,64].

Currently, there are few targeted therapies for autoimmune diseases available. The most common therapies involve B-cell depletion (BCDT), but are not always effective. Interestingly, CD22, a significantly upregulated gene in the diseased male and female animals expressing Xist, is a more recently developed BCDT target, epratuzumab[65-67]. There remains a need for more specific pathogenic leukocyte targets. Identifying novel antigens and their associated pathogenic pathways hold great therapeutic potential. Future studies will involve both an increased cohort of autoimmune patient serum and more representative general population group to serve as a comparative baseline. Profiling the XIST RNP in primary cells, both healthy and diseased, will also be essential in advancing our understanding of the aberrant autoreactivity towards proteins within the complex and identify more potential autoantigens. And, examining lncRNAs and their associated protein complexes, not limited to XIST, and their interactions with immune regulatory pathways, will be an important consideration outside their current roles as disease-associated biomarkers.

ACKNOWLEDGEMENTS

We thank members of the Chang and Utz labs for discussion and Adrianne Woods and Gwendolyn Leatherman for assistance with serum compilation from Johns Hopkins. We thank the Stanford Veterinary Service Center and the Stanford Breeding Colony Management Services for assistance with animal husbandry and welfare, the Stanford Comparative Medicine Services for histology preparation, the Stanford Transgenic, Knockout and Tumor Model Research Center for tissue sample preparation, the Stanford Functional Genomics Facility for bulk sequencing assistance, and the Stanford Shared FACS Facility for assistance in cell sorting. This work was supported by Scleroderma Research Foundation (H.Y.C.), NIAMS T32 AR007422 and NIAMS K99/R00 (D.R.D.), and NIAMS T32 AR050942 (B.T.A.). H.Y.C. is an Investigator of the Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

Conceptualization: D.R.D., P.J.U., and H.Y.C.; Methodology and Investigation: D.R.D, K.M.C., R.L., D.C., K.K., C.H., R.S., S.C., A.F., D.W.G., A.A.S., M.P, F.M.W., L.S.C., D.F.F., and A.W.; Data Analysis: D.R.D., Y.Z., J.A.B., Y.Z., K.M.C., B.Y., B.A., C.H., R.S., A.F., B.A., and E.K.L.;

Writing: D.R.D., K.M.C., L.S.C., D.F.F., P.J.U., and H.Y.C.; Funding Acquisition: H.Y.C.; Resources: H.Y. Chang; Supervision: H.Y.C.

DISCLOSURE

H.Y.C. is a co-founder of Accent Therapeutics, Boundless Bio, and an advisor to 10x Genomics, Arsenal Biosciences, and Spring Discovery. A.A.S. receives research grant funding from the following companies to support clinical trials in SSc: Arena Pharmaceuticals, Eicos Sciences, Kadmon Corporation, Medpace.

MATERIALS AND METHODS

Mice

Strains

Wild-type C57BL/6J (000664) and SJL/J (000686) mice were purchased from Jackson Laboratories. To create the transgenic lines, mice carrying the tet-O-ΔRepA-Xist construct (as described in Figure 1A) from Anton Wutz's lab were backcrossed into the C57BL/6J or SJL/J mice for multiple generations. Tail tips were collected from 5-7 female mice from each generation for speed congenics selection using Charles Rivers' MAX-BAX 384 SNP panel. The two females with the highest match in the desired background from each generation were selected as breeders. The tgXist-C57BL6/J mice used in this study are > 80% in the C57BL/6J background. The tgXist-SJL/J mice are > 99.99% in the SJL/J strain background. Mice were genotyped for both the Neomycin resistant cassette and from the CMV promoter into the truncated Xist transgene. Studies in the SJL/J background used both mice heterozygous for tgXist and wild-type littermates while studies involving SJL/J mice while wild-type control C57BL/6J mice were obtained from Jackson Laboratories. Genotyping primers are listed in **Supplemental Table 1**.

Pristane-induction of SLE

Wild-type and tg-Xist mice were injected with a one-time injection of 0.5 mL of pristane (Sigma-Aldrich, P9622-10X1ML) at 8-10 weeks old (SJL/J strain) or 12-14 weeks old (C57BL/6J studies). Control animals were injected with PBS 1x (Thermo Fisher Scientific, 10010049).

Induction of transgene (tgXist)

Simultaneous to the injection date, doxycycline (Sigma-Aldrich, S9891-5G) was continuously administered at 0.2 g/mL in drinking water to select mice until the terminal timepoint.

For the *tgXist* expression validation studies, *tgXist*$^{+/-}$ and wild-type mice in the C57BL/6J strain were administered doxycycline for 2 weeks and tissues were harvested for qRT-PCR and FISH analysis.

Xist qRT-PCR

Mechanically dissociated thymus, spleen, kidney, and liver cells harvested from *tgXist*$^{+/-}$ and wild-type mice in the C57BL/6J were strained through a cell filter, pelleted, and frozen in RLT Buffer with 0.1% BME. RNA was extracted using the RNeasy Mini kit (Qiagen, 74106), genomic DNA was removed using amplification grade DNAse I (Thermo Fisher Scientific, 18068015). RNA concentration was quantified on a Nanodrop. cDNA was made from 1 ug of RNA/sample using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, 4368814). Each Ct value was measured using Lightcycler 480 (Roche) and each mean dCt was averaged from triplicate qRT-PCR reactions. Relative Xist RNA levels was calculated by ddCt method compared to GAPDH controls. Statistical significance was calculated using the Student's t-test. Primer sequences are listed in the **Supplemental Table 1**.

Xist FISH

Mice organs frozen in O.C.T. compound were sectioned on a crysostat to 15 mm onto microscope slides (Fisher Scientific, 12-550-15). Sections were washed once with PBS 1x, fixed in 4% paraformaldehyde for 10 min at room temperature, washed twice with PBS 1x for 2-5 min each, then permeabilized on ice with ice cold 0.5% TritonX-100 in PBS for 10 min. Slides were then rinsed once with PBS and dehydrated sequentially in 70%, 90% and 100% Ethanol (Gold Shield, 412804) for 5 minutes each. Sectioned slides were then allowed to air dry before immersion in freshly made Stellaris® RNA FISH Wash Buffer A (SMF-WA1-60) for 2-5 minutes at room temperature. Sections were then hybridized overnight for 16-18 hours in the dark at 42C in 250 nM of Mouse Xist Stellaris® FISH Probes with Quasar® 570 Dye (SMF-3011-1) in Stellaris® RNA FISH Hybridization Buffer (SMF-HB1-10). The next day, slides were incubated twice in Wash Buffer A for 30 minutes at 37C followed by immersion in Stellaris® RNA FISH Wash Buffer B for 2-5 minutes at 27C, mounted in VECTASHIELD Mounting Medium with DAPI (Vector Laboratories, H-1200) and sealed with cover glass (VWR, 48366-227) and nail polish clear coat. Sectioned organs were imaged on the Zeiss Observer Z.1 using the 63x oil objective, X-Cite Series120 laser, and AxioVision Rel. 4.8 software.

Mouse sera collection

Mice were sedated with isofluorane and retro-orbitally bled immediately prior to the injection and viably bled every 4 weeks until the terminal date. On the terminal date, blood was collected through cardiac puncture. Blood was allowed to clot for 2 hours at room temperature, then spun for 15 minutes at 1500xg at room temperature. Serum was immediately collected, flash-frozen on dry ice and stored at -80C.

ANAb ELISA

Serum ANAb levels were measured using the Mouse Anti-Nuclear Antibody Kit (MyBioSource.com, MBS731183) and assessed on an ELISA plate reader according to manufacturer's protocol. Titration curves and values were calculated using a Four Parameter Logistic Curve.

Tissue collection and preparation of pristane-induced SLE studies

Mice were euthanized by CO_2 asphyxiation and cardiac exsanguination at the terminal timepoints of 16- or 52-weeks post-injection. Terminal cardiac blood was aliquoted into EDTA tubes (400 ml per mouse) and Eppendorf tubes (for serum collection). Total body weights were obtained, and the following organs were weighed individually: thymus, heart, liver, spleen, testes. Organ weights as a percentage of total body weight were calculated. Gross necropsies were performed on all mice. The caudate and papillary liver lobes, left kidney, one half of the thymus and one half of the spleen were dissected and placed on ice in HypoThermosol FRS (BioLifeSolutions, 101102) until all dissections concluded and were either mechanically dissociated using clean scalpels and syringes or frozen in O.C.T. Compound (Tissue-Tek, 4583) at -80C on the same day. The remaining tissues were immersion-fixed in 10% neutral buffered formalin for 72 hours for downstream histology analysis. Dissociated cells were passed through cell filters (Sysmex, 04-004-2327), pelleted at 300xg, incubated for 1 minute at room temperature in eBioscience™ 1X RBC Lysis Buffer (00-4333-57) to remove red blood cells, and quenched with 10x the volume of PBS 1X. For viable stocks, dissociated cells were frozen in BAMBANKER (Wako, #203-14681) and stored in liquid nitrogen. For RNA extraction, cells were frozen in either TRI Reagent® (Molecular Research Center, TR118) or Buffer RLT Plus (Qiagen, 1053393) and stored at -80C.

CD4+ isolation

CD4+ cells were isolated from freshly dissociated mouse spleen using the Easysep Mouse CD4+ T cell Isolation Kit (STEMCELL Technologies, 19852). CD4+ cells were then immediately used for bulk ATAC-seq, frozen in RLT Buffer Plus, or viably frozen in BAMBANKER. A small aliquot of cells were stained in PBS 1x with the viability marker 7-AAD (BD Biosciences, 559925) and T-cell marker CD4 FITC (clone GK1.5, BioLegend, 100406) to assess purity and viability on a FACS analyzer.

CD45+ isolation

Viable cells frozen in BAMBANKER were thawed at 37C in a bead bath, resuspended in RPMI, and spun at 300xg to pellet and remove the buffer. Cells were stained with the pan-hematopoietic marker CD45 APC (clone 30-F11, Biolegend, 103112) and 7-AAD for viability. Viable CD45+ cells were sorted on the BD FACSAria II sorter using a 70 uM nozzle into chilled FBS in the Stanford Shared FACS Facility. Sorted cells were then immediately used to prepare sequencing libraries.

Histopathology

Formalin-fixed tissues were processed routinely, embedded in paraffin, sectioned at 5 mm, and stained with hematoxylin and eosin. All tissues were evaluated blindly by a board-certified veterinary pathologist (KMC). An ordinal histopathologic grading scale (score 0-4) was designed to evaluate glomerulonephritis, hepatic lipogranulomas, pulmonary lymphohistiocytic alveolitis, and splenic extramedullary hematopoiesis. A binary histopathologic grading scale (score 0 or 1) was used to evaluate for intraabdominal lipogranuloma formation and ectopic lymphoid tissue, pulmonary hemorrhage, hemosiderosis, and/or vascular thrombosis, splenic lymphoid hyperplasia and plasmacytosis, and lymph node hyperplasia and medullary plasmacytosis. A total composite score was derived for each mouse. The Fisher's Exact Test was used to calculate significance between treatment groups.

Sequencing Library Preparation

Bulk sequencing: ATAC-seq libraries of freshly isolated CD4+ cells were prepared using the Omni-ATAC protocol[68]. RNA was extracted using the RNeasy Plus Mini Kit (Qiagen, 74136). TruSeq® Stranded mRNA Library Prep (Illumina, 20020594) was used to generate polyA-selected RNA-sequencing libraries, cleanup performed on magnets with AMPure XP beads (Beckman Coulter, A63880). The bulk ATAC-seq and RNA-seq libraries were sequenced with paired-end 75 bp read lengths on in the Stanford Functional Genomics Facility an Illumina HiSeq 4000 that was purchased with funds from NIH under award number S10OD018220.

Single Cell Sequencing: The Chromium Single Cell Multiome ATAC + Gene Expression (10x Genomics) was used to prepare libraries for CD45+ sorted cells with a target of 10,000 cells/sample. Libraries were sent to Novogene for Bioanalyzer trace quality control check and sequencing. Libraries were sequenced on the NovaSeq 6000 at a depth of 20,000 paired reads per cell for gene expression libraries and 25,000 paired reads per cell for ATAC libraries.

Sequencing Library Analysis

Bulk sequencing: The bulk RNA-seq data was aligned to mm10 using STAR. The gene expression read counts was generated using RSEM. The adaptor of paired-end ATAC-seq data were first trimmed by an in-house software, and then aligned to hg38 genome using bowtie2. The

mitochondrial reads and reads with low alignment score (<10) were removed. The aligned sam files were converted to bam files and sorted by Samtools. Picard was used to remove duplicate reads and Macs2 was used to call peaks. Each ATAC-seq peak was annotated by its nearby genes using GREAT under the basal plus extension default setting.

The raw bulk RNA-seq and ATAC-seq read counts were then normalized and analyzed using DESeq2. The differential expressed genes and ATAC-seq peaks were identified using the negative binomial models. Benjamini hochberg procedure was used to adjust for multiple hypothesis testing. Peaks with FDR < 0.2 and absolute fold change larger than 1.5 were selected as significant. Bioinformatics tool Gprofiler was used for pathway enrichment analysis. CIBERSORT was used to estimate the abundance of immune cells based on normalized RNA-seq data.

Single cell multiomics sequencing (gene expression): The single-cell paired RNA and ATAC-seq reads were aligned to the mm10 reference genome using cellranger-arc count (10x Genomics, version 2.0.1). Gene expression data was filtered to include only barcodes that had nFeature > 500 and pct.mt < 12. Two female samples (F100 and F113) were excluded because, even after stringent filtering, cells from these samples had lower quality metrics than other samples. R version 4.1.3 Seurat_4.1.1 and ggplot2_3.3.6 were used for downstream analysis and visualization. After filtering, clustering and dimensionality reduction was performed using the top 10 principle components (dims=1:10) and a clustering resolution of 0.25. Differential gene analysis was performed using Seurat function "FindMarkers" with default settings.

Clinical Cohorts

All human patient samples were de-identified in this study.

General Population: Donated whole blood from a total of 9 male and 8 female donors were obtained from the Stanford Blood Bank following IRB-approved protocols with confirmed assent for use in research. Serum was collected as described above in the mouse serum section and frozen in -80C for long-term storage. During data analysis, 1 of the 9 male samples was filtered out for high IgG background.

Dermatomyocitis (Stanford): Serum from 1 male and 3 female dermatomyocitis patients were used in this study. All samples were collected from patients seen at the Stanford outpatient clinics under an IRB-approved protocol, and all patients provided informed consent to participate. The dermatomyositis cohort has been described previously[69]. All patients met probable or definite DM by 2017 ACR/EULAR IIM Classification Criteria[70]. All sera used in the study were also known to contain antibodies against TIF1-γ which was assayed as previously described[71].

Systemic Sclerosis (Stanford): Serum from 24 patients (2 male and 22 female) with systemic sclerosis (SSc) were included in this study from the Stanford cohort. All samples were collected from patients seen at the Stanford outpatient clinics under an IRB-approved protocol, and all patients provided informed consent to participate. All patients fulfilled 2013 ACR/EULAR classification criteria for SSc. Twelve patients had diffuse and 12 had limited cutaneous SSc. Seven patients had the Scl-70 antibody, 5 had the anti-centromere antibody, 2 patients had RNA polymerase III, 2 had nucleolar ANA, 2 had PM/Scl, and 1 had U1RNP, with the remainder having positive ANA but no known SSc-specific autoantibody.

Systemic Sclerosis (Johns Hopkins): Serum from 9 males and 31 females were shipped on dry ice to Stanford. Patients who are enrolled in the Johns Hopkins Scleroderma Center Research registry agree to have up to 7 tablespoons of blood drawn at the time of a scheduled clinic visit

collected no more than 8 times per year. This blood is drawn at the time of a clinical blood draw when possible. Scleroderma participants in the registry meet at least one of the following criteria for systemic sclerosis: 1) 2013 ACR/EULAR classification criteria for scleroderma, 2) 1980 ACR classification criteria, 3) having at least 3 of 5 features of the CREST syndrome (calcinosis, Raynaud's phenomenon, esophageal dysmotility, sclerodactyly, telangiectasia), or 4) having definite Raynaud's phenomenon, abnormal nailfold capillaries and a scleroderma-specific autoantibody. Serum samples are kept in a -80 freezer for long term storage.

Systemic Lupus Erythematosus (Johns Hopkins): Patients provided written informed consent to participate in the Hopkins Lupus Cohort (IRB study number NA_00039294). Blood was drawn at the time of a clinical blood draw, serum collected and kept in a -80 freezer for long term storage. For this study, 40 serum samples (3 males and 37 females) were randomly selected from patients who were ever: 1) ANA positive with a titer of at least 1:320; and 2) positive to the dsDNA autoantigen. Patients enrolled in the Hopkins Lupus Cohort met either the revised American College of Rheumatology (ACR)[72] or Systemic Lupus International Collaborating Clinics (SLICC)[73] criteria for SLE. Samples were shipped on dry ice to Stanford and stored at –80C.

SLE autoantigen Panel Array

Assay

Mouse serum reactivity autoantigen was assessed using the SLE autoantigen panel and array as previously described[35]. Serum from the baseline timepoint of 0 weeks (injection start) and terminal timepoints (16 weeks for SJL/J and 52 weeks for C57BL/6J strains) were run in the same assay (sample-matched). Serum was stored long-term at –80C and prepared as previously described[35].

Analysis

Raw MFI scores were first normalized by subtracting the baseline value of bare bead. The difference between terminal and baseline timepoints was then calculated using the normalized values. Values were plotted in Graphpad PRISM and the student's t-test was used to test for statistical significance.

XIST Autoantigen Array

Autoantigen list

XIST-associating proteins were chosen from published XIST ChIRP-MS datasets[14,48]. Autoantigens clinically used to screen for DMM, SLE or SSc were included as positive controls. SciLifeLab also included in-house controls, such as EBNA1 (Eppstein Barr virus), bare bead (unconjugated), and human and mouse IgG. The full list of autoantigens and their categorization can be found in **Supplemental Table 2**.

The suspension bead array (SBA) includes 384 bead IDs, out of which 6 are technical controls. All antigens (except for the controls) are protein fragments called Protein Signature Tags (PrESTs) expressed in Escherichia coli with a six histidine and albumin binding protein tag (His6ABP), and have been produced within the Human Protein Atlas project (https://www.proteinatlas.org/). The PrESTs have been designed to not contain transmembrane regions and to have less than 60% sequence similarity (on gene-level) towards any other human gene. If the sequence similarity criteria have failed to be fulfilled fragments mapping to a group of similar proteins have been produced instead. In the latter case one of the genes the fragment maps to is included with an asterisk (*) afterwards in the short hand identifiers, and the full list of

genes the fragment maps to is found in the antigen meta data. In this project, the PrESTs are 20-151 amino acids long.

Sample Preparation

25 uL of serum per sample were aliquoted in a pre-determined randomized order onto 96-well plates (Thermo Fisher Scientific, AB800150) with mouse and human samples on separate plates. Thermal sealed plates were shipped on dry ice to SciLifeLab in Sweden and stored at –80C upon arrival.

Suspension Bead Array Assay (SciLifeLab)

The antigens and six technical controls were covalently coupled to color coded magnetic beads (MagPlex, Luminex Corp., Austin, TX) using NHS and EDC based chemistry. The controls consisted of His6ABP (negative, tag binding control for the protein fragments; Human Protein Atlas), buffer (negative, bead binding control), rabbit anti human IgG (positive, sample loading control; 309-005-082, Jackson immunoresearch), donkey anti mouse IgG (positive, sample loading control; 715-005-150, Jackson immunoresearch), purified mouse IgG (positive, secondary control; PMP01X, BioRad), and EBNA1 (semi positive control for humans; ab138345, abcam).

In total, one assay was performed in one day as follows: The human samples were diluted 1:250 and the mouse samples were diluted 1:100 in assay buffer (5% BSA in PBS supplemented with 0.1% Tween-20 and 0.16 mg/ml His6ABP tag). Commercial plasma (mixed 50% male & female pool; HUMANPLK2PNN, BioIVT) and non-commercial in-house available mouse serum (AbClone) were diluted like the test samples and added to the dilution plates as technical controls. Before incubation with the antigen bead array, all diluted samples incubated for 1 hour at room temperature in order for the His6ABP in the buffer to pre-block any potential antibodies toward the tag (derived from Streptococcal protein G). The diluted samples were subsequently incubated with the antigen bead array for 2 hours. Any antibody binding to the beads were fixated by incubating the beads in 0.2% paraformaldehyde for 10 min. R-Phycoerythrin labelled secondary reagents targeting human IgG (Goat anti-Human IgG Fc Secondary Antibody, PE, eBioscienceTM; 12-4998-82, Invitrogen) and mouse IgG (R-Phycoerythrin AffiniPure F(ab')2 Fragment Goat Anti-Mouse IgG (H+L); 115-116-146, Jackson immunoresearch), respectively, were then applied for 30 min to enable a read out using a FlexMap 3D (Luminex Corp., Austin, TX). The read out consists of the median fluorescent intensity (MFI) and number of beads for each antigen (bead ID) in each sample that the MFI is based on (median of signals from individual beads with the same bead ID).

Data handling and analysis

All data handling and analysis was performed using R version 4.1.1. The antigens have antigen specific background levels, causing potential bias during the sample specific transformation and making comparisons between antigens difficult. In order to account for this, the data is centered so that the 10th percentile of each antigen (excluding filtered samples/antigens and control wells) is equal to the 10th percentile of the entire data set (excluding filtered samples/ antigens, control wells and positive control beads). The raw MFI is denoted MFIorg and the percentile adjusted MFI is denoted MFIadj in the data files. Since the samples also have sample specific background levels, direct comparisons between the MFI values can be skewed. Using the assumption that in any given sample, a portion of the antigens will not be reactive, an estimate background level can be found per sample to center the data. The antigen percentile adjusted MFI values were transformed per sample into number of "median absolute deviations" (MADs) around the sample median (assuming that at least half of the antigens are definitely in the background in the sample), using the algorithm:

$$MADs_{ag,sample} = (MFIadj_{ag,sample} - median_{sample}) / MAD_{sample}$$

FIGURES

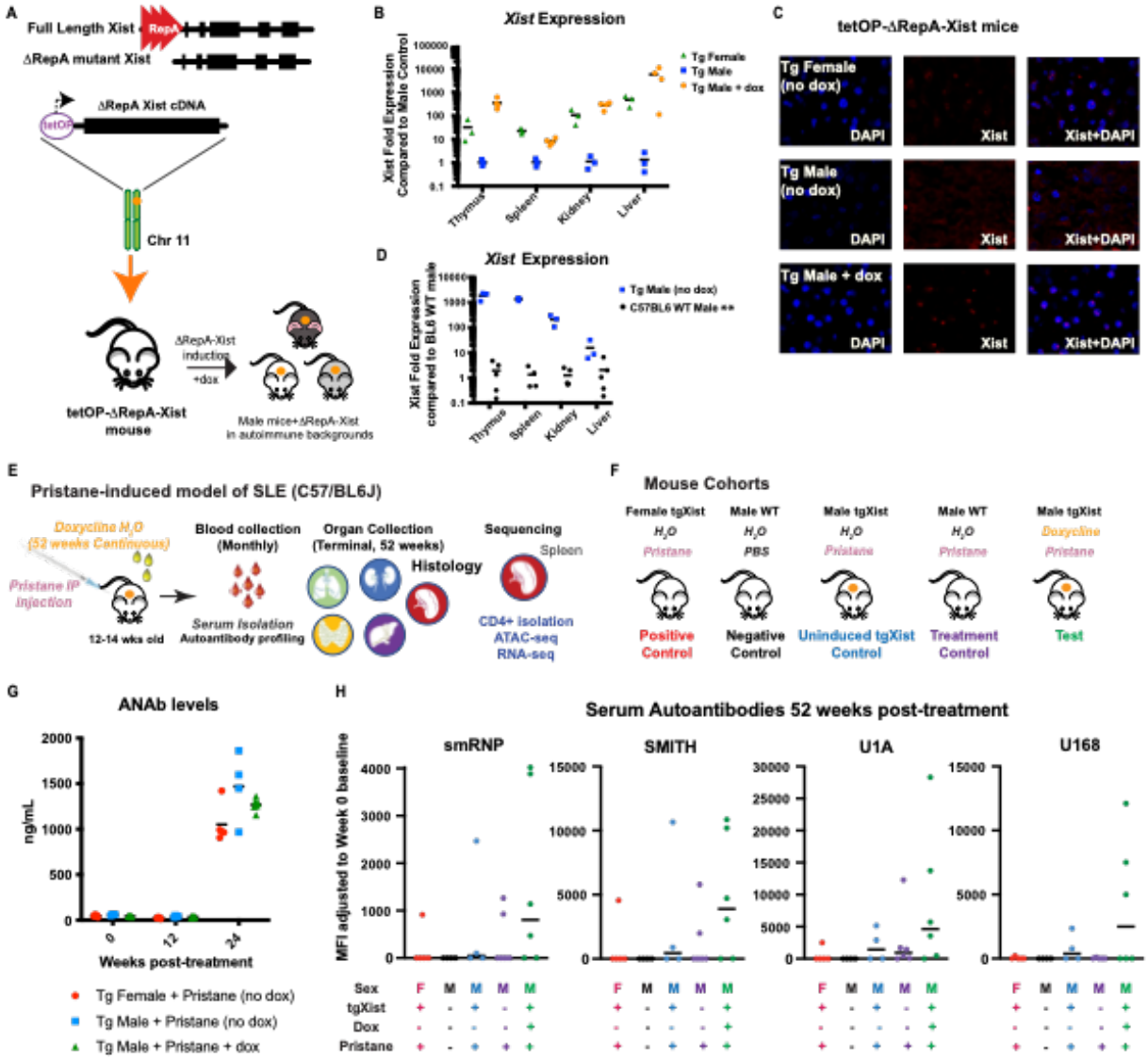


Figure 1: tgXist-mice inducibly express ΔrepA-Xist under doxycycline exposure. (A) Transgene construct of truncated ΔrepA Xist under tet-o-promoter in mice. (B) qRT-PCR and (C) FISH of Xist expression in tgXist-mice after 2 weeks of doxycycline administered in drinking water. (D) qRT-PCR comparison of basal tgXist levels in tgXist male mice without doxycycline exposure and WT male mice. (E) Schematic of tissue collection in pristane-induced SLE model in the C57/BL6J strain. (F) Cohorts in the induced SLE model. (G) ELISA of Serum ANAb levels. H) Median Fluorescence Intensity (MFI) of serum reactivity to autoantigens from lupus antigen array after 52 weeks of treatment.

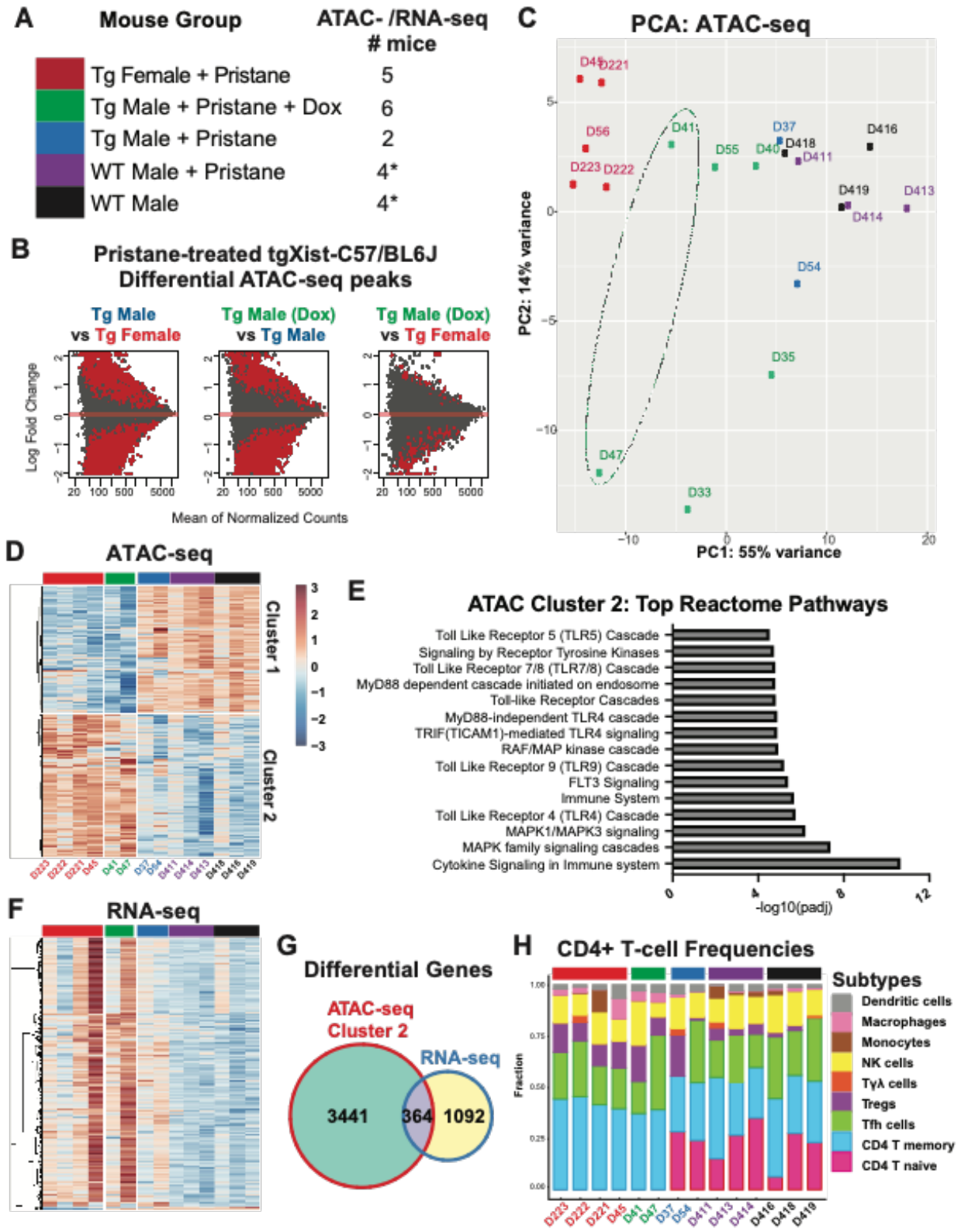
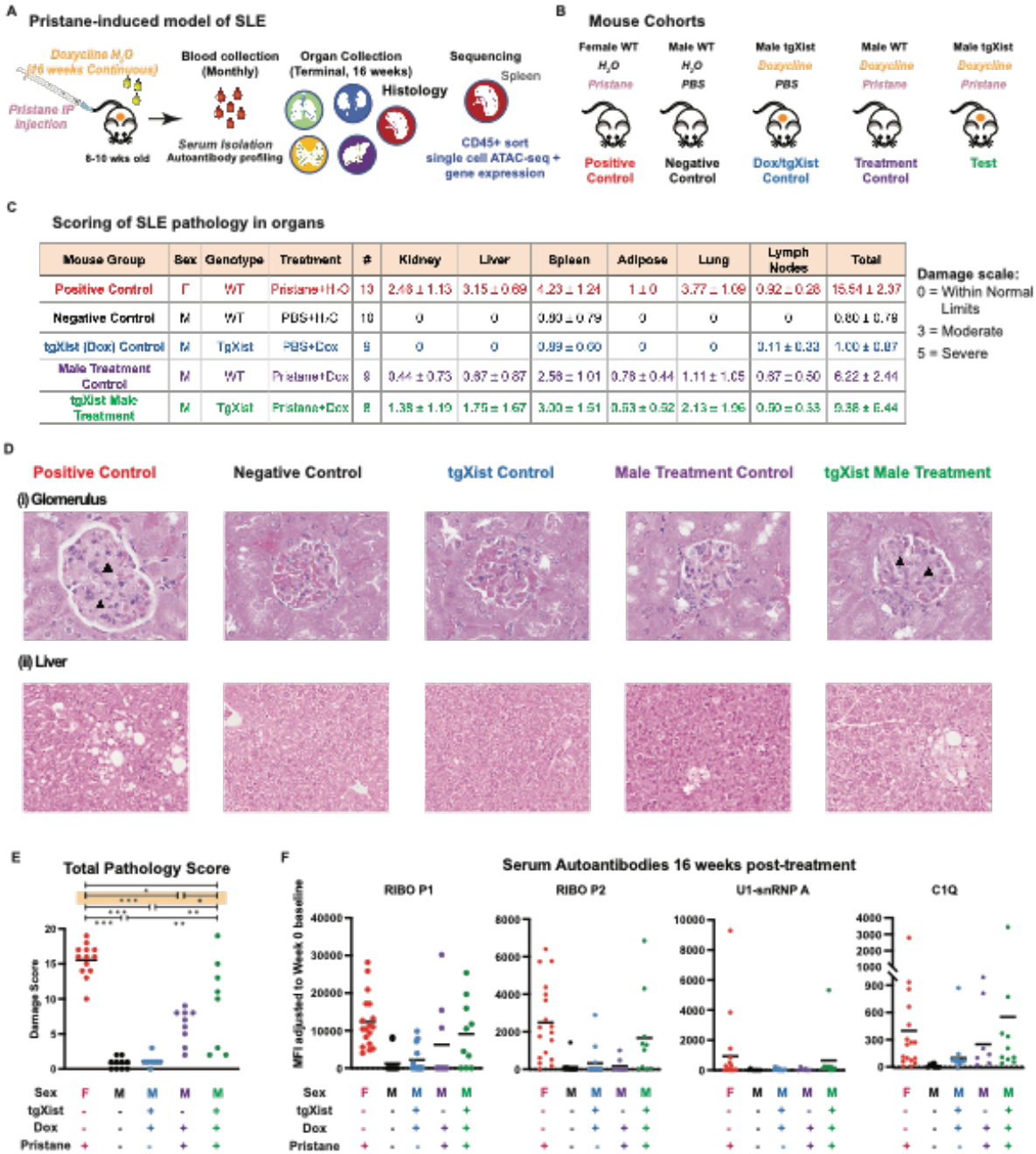


Figure 2: Bulk ATAC- and RNA-sequencing of splenic CD4+ T-cells from C57/BL6J mice reveal closer correlation between tgXist- and Xist-expressing mice in the pristane-induced

SLE model. (A) Table of treatment cohorts: colors correspond to each treatment cohort and numbers indicate the number of ATAC- and RNA-seq samples. *Due to library quality, only 3 RNA-seq libraries were used. (B) MA plots comparing differential regions of genomic accessibility in transgenic cohorts. (C) Poly Component Analysis (PCA) plots of ATAC-seq libraries from all mice in the study. Heatmap of (D) ATAC-seq Z-scores and (F) RNA-seq differential gene expression from representative mice from each treatment cohort. (E) Top 15 differential reactomes associated with cluster 2 genomic regions. (G) Metrics of differential genes from cluster 2 ATAC-seq and RNA-seq datasets. (H) CIBERSORT prediction of T-cell subset composition from CD4+ T-cells RNA-seq gene expression libraries.



Mouse Group	Sex	Genotype	Treatment	#	Kidney	Liver	Spleen	Adipose	Lung	Lymph Nodes	Total
Positive Control	F	WT	Pristane+H_2O	13	2.46 ± 1.13	3.15 ± 0.69	4.23 ± 1.24	1 ± 0	3.77 ± 1.09	0.92 ± 0.28	15.54 ± 2.37
Negative Control	M	WT	PBS+H_2O	10	0	0	0.90 ± 0.79	0	0	0	0.90 ± 0.79
tgXist (Dox) Control	M	TgXist	PBS+Dox	9	0	0	0.89 ± 0.60	0	0	0.11 ± 0.33	1.00 ± 0.87
Male Treatment Control	M	WT	Pristane+Dox	9	0.44 ± 0.73	0.67 ± 0.87	2.56 ± 1.01	0.78 ± 0.44	1.11 ± 1.05	0.67 ± 0.50	6.22 ± 2.44
tgXist Male Treatment	M	TgXist	Pristane+Dox	8	1.38 ± 1.19	1.75 ± 1.67	3.00 ± 1.51	0.63 ± 0.52	2.13 ± 1.96	0.50 ± 0.53	9.38 ± 6.44

Figure 3: Increased phenotypic changes in pathophysiology and autoantigen levels in tgXist- and Xist-expressing mice in the SJL/J strain in the pristane-induced SLE model. (A) Schematic of pristane-induced SLE and strategy for histopathology, serum and sequencing analysis in the SJL/J strain. (B) Treatment cohorts in the SJL/J strain. (C) Table of the mean severity of damage across multiple tissue sites and mice. (D) Representative H&E images of (i) glomeruli (kidney) and (ii) liver sections from each treatment cohort. Arrows demarcate mesangial thickening. (E) Graph of total sum of pathophysiology damage scores. (F) Median Fluorescence Intensity (MFI) of serum reactivity to autoantigens from lupus antigen array after 16 weeks of treatment.

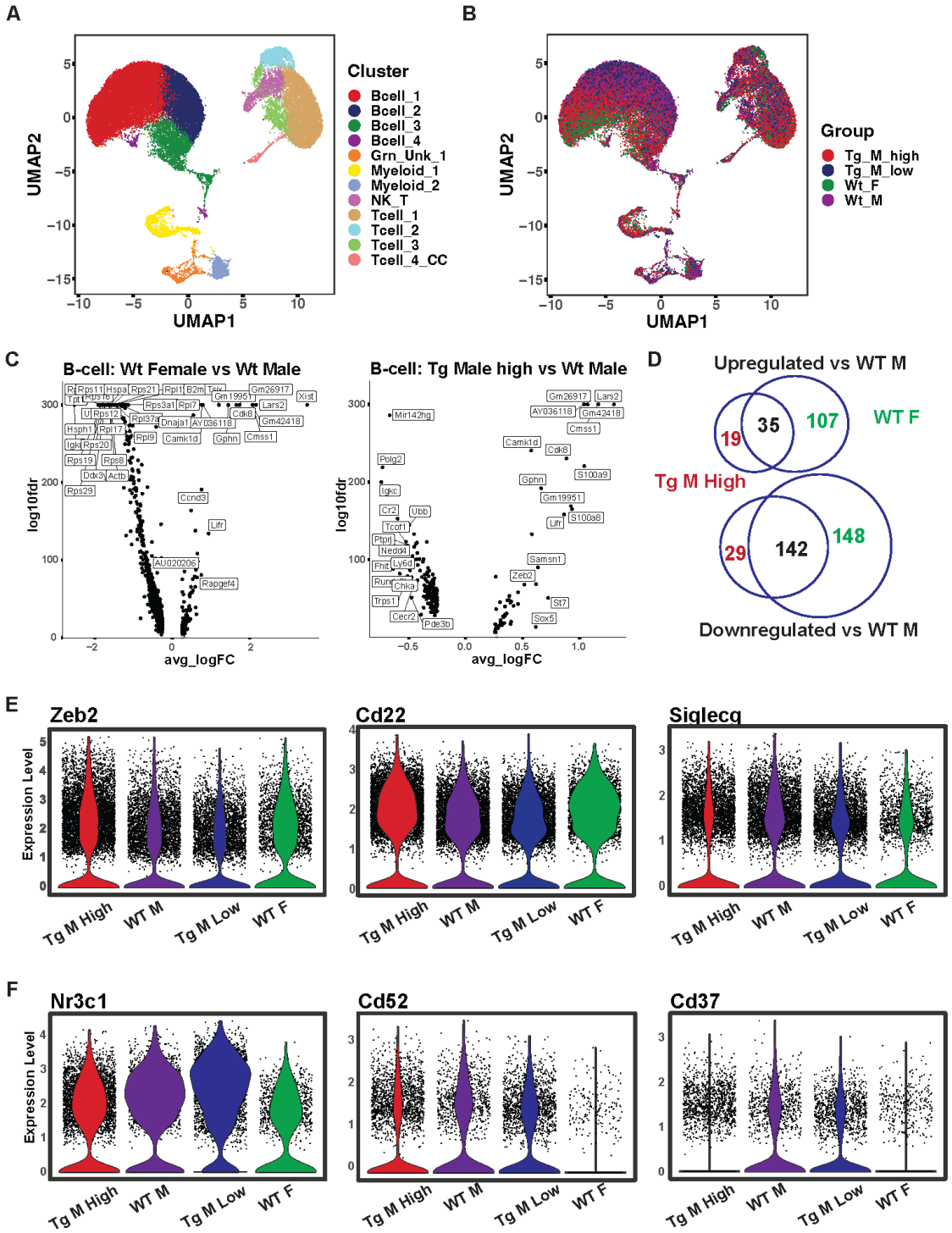


Figure 4: Single cell gene expression comparisons of splenic CD45+ hematopoietic cells from pristane-treated mice in the SJL/J strain.

(A) Cell type cluster identities and (B) pristane-treated mouse groups displayed on the single cell gene expression clustering UMAP. (C) Volcano plots and (D) metrics of differentially expressed B-cell cluster genes of tgXist Male high disease and wild-type female compared to wild-type animals. (E) Representative gene expression plots from the B-cell and (F) T-cell cluster highlighting significant differentially expressed genes. Pristane-treated mouse groups

shown: tgXist male disease high (n=4), tgXist male disease low (n=4), wild-type male (n=3), and wild-type female (n=2\). Significance was calculated using the Wilcoxon Rank Sum test.

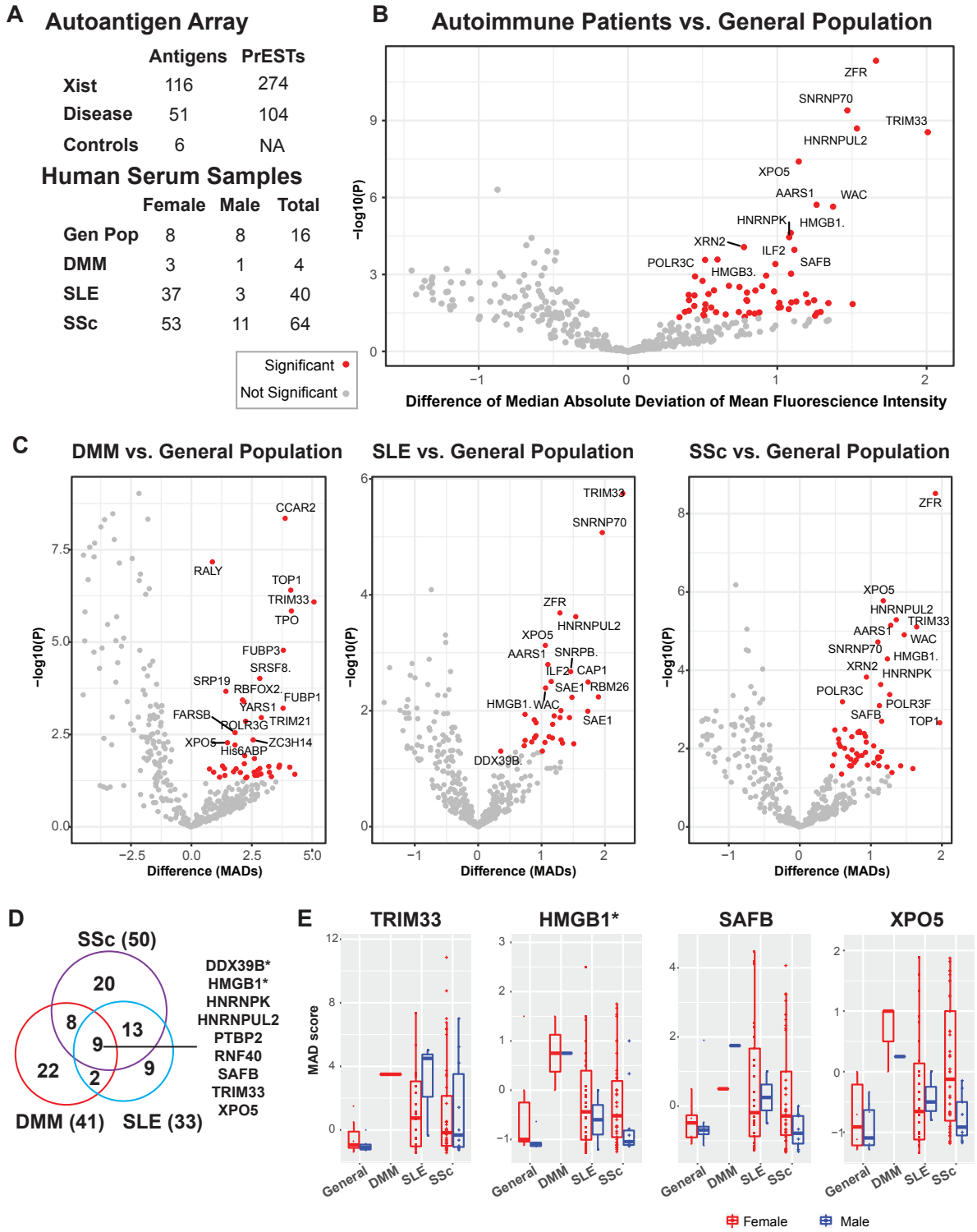


Figure 5: Autoimmune disease individuals experience increased serum reactivity towards antigens of the XIST RNP. (A) Table of the XIST complex-associated antigen array design showing the number of total unique proteins and corresponding PrESTs counts drawn

from XIST ChIRP-MS, clinical disease panel, and control protein lists; and sample numbers of autoimmune patient and general population serum. (B) Volcano plot of serum reactivity of all autoimmune patients and (C) sera grouped by DMM, SSc, and SLE patients compared to general population baseline of serum from blood donors. The log(adjusted p-value) significance calculated by Student's T-test and reactivity difference calculated from Median Average Deviation (MAD) difference between disease and general population groups. Significant differentially reactive antigens, defined as $p_{adj} < 0.05$ and MAD difference ³ 0, labeled in red. (D) Metrics of unique antigens from the array with significant elevated serum activity in autoimmune patients compared to the general population. (E) Serum reactivity (MAD) plots of representative antigens significantly reactive in all three autoimmune patient cohorts grouped by disease type and colored by sex (red=female, blue=male).

REFERENCES

1. Fairweather, D., and Rose, N.R. (2004). Women and autoimmune diseases. Emerging infectious diseases *10*, 2005-2011. 10.3201/eid1011.040367.
2. Libert, C., Dejager, L., and Pinheiro, I. (2010). The X chromosome in immune functions: when a chromosome makes the difference. Nat Rev Immunol *10*, 594-604. 10.1038/nri2815.
3. Fairweather, D., Frisancho-Kiss, S., and Rose, N.R. (2008). Sex differences in autoimmune disease from a pathological perspective. The American journal of pathology *173*, 600-609. 10.2353/ajpath.2008.071008.
4. Moulton, V.R. (2018). Sex Hormones in Acquired Immunity and Autoimmune Disease. Front Immunol *9*, 2279. 10.3389/fimmu.2018.02279.
5. Syrett, C.M., Sierra, I., Beethem, Z.T., Dubin, A.H., and Anguera, M.C. (2020). Loss of epigenetic modifications on the inactive X chromosome and sex-biased gene expression profiles in B cells from NZB/W F1 mice with lupus-like disease. J Autoimmun *107*, 102357. 10.1016/j.jaut.2019.102357.
6. Syrett, C.M., Sindhava, V., Sierra, I., Dubin, A.H., Atchison, M., and Anguera, M.C. (2018). Diversity of Epigenetic Features of the Inactive X-Chromosome in NK Cells, Dendritic Cells, and Macrophages. Front Immunol *9*, 3087. 10.3389/fimmu.2018.03087.
7. Subramanian, S., Tus, K., Li, Q.-Z., Wang, A., Tian, X.-H., Zhou, J., Liang, C., Bartov, G., McDaniel, L.D., Zhou, X.J., et al. (2006). A Tlr7 translocation accelerates systemic autoimmunity in murine lupus. Proceedings of the National Academy of Sciences *103*, 9970-9975. 10.1073/pnas.0603912103.
8. Christensen, S.R., Shupe, J., Nickerson, K., Kashgarian, M., Flavell, R.A., and Shlomchik, M.J. (2006). Toll-like receptor 7 and TLR9 dictate autoantibody specificity and have opposing inflammatory and regulatory roles in a murine model of lupus. Immunity *25*, 417-428. 10.1016/j.immuni.2006.07.013.
9. Holroyd, C.R., and Edwards, C.J. (2009). The effects of hormone replacement therapy on autoimmune disease: rheumatoid arthritis and systemic lupus erythematosus. Climacteric *12*, 378-386. 10.1080/13697130903025449.
10. Bogdanos, D.P., Smyk, D.S., Rigopoulou, E.I., Mytilinaiou, M.G., Heneghan, M.A., Selmi, C., and Gershwin, M.E. (2012). Twin studies in autoimmune disease: genetics, gender and environment. J Autoimmun *38*, J156-169. 10.1016/j.jaut.2011.11.003.
11. Awate, S., Babiuk, L.A., and Mutwiri, G. (2013). Mechanisms of action of adjuvants. Front Immunol *4*, 114. 10.3389/fimmu.2013.00114.
12. Scofield, R.H., Bruner, G.R., Namjou, B., Kimberly, R.P., Ramsey-Goldman, R., Petri, M., Reveille, J.D., Alarcón, G.S., Vilá, L.M., Reid, J., et al. (2008). Klinefelter's syndrome (47,XXY) in male systemic lupus erythematosus patients: support for the notion of a gene-dose effect from the X chromosome. Arthritis Rheum *58*, 2511-2517. 10.1002/art.23701.
13. Smith-Bouvier, D.L., Divekar, A.A., Sasidhar, M., Du, S., Tiwari-Woodruff, S.K., King, J.K., Arnold, A.P., Singh, R.R., and Voskuhl, R.R. (2008). A role for sex chromosome complement in the female bias in autoimmune disease. J Exp Med *205*, 1099-1108. 10.1084/jem.20070850.
14. Chu, C., Zhang, Q.C., da Rocha, S.T., Flynn, R.A., Bharadwaj, M., Calabrese, J.M., Magnuson, T., Heard, E., and Chang, H.Y. (2015). Systematic discovery of Xist RNA binding proteins. Cell *161*, 404-416. 10.1016/j.cell.2015.03.025.
15. Lau, C.M., Broughton, C., Tabor, A.S., Akira, S., Flavell, R.A., Mamula, M.J., Christensen, S.R., Shlomchik, M.J., Viglianti, G.A., Rifkin, I.R., and Marshak-Rothstein, A. (2005). RNA-associated autoantigens activate B cells by combined B cell antigen receptor/Toll-like receptor 7 engagement. J Exp Med *202*, 1171-1177. 10.1084/jem.20050630.

16. Busconi, L., Lau, C.M., Tabor, A.S., Uccellini, M.B., Ruhe, Z., Akira, S., Viglianti, G.A., Rifkin, I.R., and Marshak-Rothstein, A. (2006). DNA and RNA autoantigens as autoadjuvants. J Endotoxin Res *12*, 379-384. 10.1179/096805106X118816.
17. Means, T.K., Latz, E., Hayashi, F., Murali, M.R., Golenbock, D.T., and Luster, A.D. (2005). Human lupus autoantibody-DNA complexes activate DCs through cooperation of CD32 and TLR9. J Clin Invest *115*, 407-417. 10.1172/JCI23025.
18. Wutz, A., Rasmussen, T.P., and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nat Genet *30*, 167-174. 10.1038/ng820.
19. Graham, K.L., Thibault, D.L., Steinman, J.B., Okeke, L., Kao, P.N., and Utz, P.J. (2005). Granzyme B is dispensable for immunologic tolerance to self in a murine model of systemic lupus erythematosus. Arthritis Rheum *52*, 1684-1693. 10.1002/art.21092.
20. Smith, D.L., Dong, X., Du, S., Oh, M., Singh, R.R., and Voskuhl, R.R. (2007). A female preponderance for chemically induced lupus in SJL/J mice. Clin Immunol *122*, 101-107. 10.1016/j.clim.2006.09.009.
21. Satoh, M., Richards, H.B., Shaheen, V.M., Yoshida, H., Shaw, M., Naim, J.O., Wooley, P.H., and Reeves, W.H. (2000). Widespread susceptibility among inbred mouse strains to the induction of lupus autoantibodies by pristane. Clin Exp Immunol *121*, 399-405. 10.1046/j.1365-2249.2000.01276.x.
22. Reeves, W.H., Lee, P.Y., Weinstein, J.S., Satoh, M., and Lu, L. (2009). Induction of autoimmunity by pristane and other naturally occurring hydrocarbons. Trends in immunology *30*, 455-464. 10.1016/j.it.2009.06.003.
23. Izui, S., Higaki, M., Morrow, D., and Merino, R. (1988). The Y chromosome from autoimmune BXSB/MpJ mice induces a lupus-like syndrome in (NZW x C57BL/6)F1 male mice, but not in C57BL/6 male mice. Eur J Immunol *18*, 911-915. 10.1002/eji.1830180612.
24. Luckheeram, R.V., Zhou, R., Verma, A.D., and Xia, B. (2012). $CD4^+$T cells: differentiation and functions. Clin Dev Immunol *2012*, 925135. 10.1155/2012/925135.
25. Barron, L., and Wynn, T.A. (2011). Fibrosis is regulated by Th2 and Th17 responses and by dynamic interactions between fibroblasts and macrophages. Am J Physiol Gastrointest Liver Physiol *300*, G723-728. 10.1152/ajpgi.00414.2010.
26. Palmer, M.T., and Weaver, C.T. (2010). Autoimmunity: increasing suspects in the CD4+ T cell lineup. Nature Immunology *11*, 36-40. 10.1038/ni.1802.
27. Newman, A.M., Steen, C.B., Liu, C.L., Gentles, A.J., Chaudhuri, A.A., Scherer, F., Khodadoust, M.S., Esfahani, M.S., Luca, B.A., Steiner, D., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. Nature Biotechnology *37*, 773-782. 10.1038/s41587-019-0114-2.
28. Storer, J.B. (1966). Longevity and gross pathology at death in 22 inbred mouse strains. J Gerontol *21*, 404-409.
29. Constantinescu, C.S., Farooqi, N., O'Brien, K., and Gran, B. (2011). Experimental autoimmune encephalomyelitis (EAE) as a model for multiple sclerosis (MS). Br J Pharmacol *164*, 1079-1106. 10.1111/j.1476-5381.2011.01302.x.
30. Celhar, T., and Fairhurst, A.-M. (2016). Modelling clinical systemic lupus erythematosus: similarities, differences and success stories. Rheumatology *56*, i88-i99. 10.1093/rheumatology/kew400.
31. Morel, L. (2004). Mouse models of human autoimmune diseases: essential tools that require the proper controls. PLoS biology *2*, E241-E241. 10.1371/journal.pbio.0020241.
32. Bender, A.T., Wu, Y., Cao, Q., Ding, Y., Oestreicher, J., Genest, M., Akare, S., Ishizaka, S.T., and Mackey, M.F. (2014). Assessment of the translational value of mouse lupus models using clinically relevant biomarkers. Translational Research *163*, 515-532. https://doi.org/10.1016/j.trsl.2014.01.003.

33. Perry, D., Sang, A., Yin, Y., Zheng, Y.-Y., and Morel, L. (2011). Murine Models of Systemic Lupus Erythematosus. Journal of Biomedicine and Biotechnology *2011*, 271694. 10.1155/2011/271694.
34. Cojocaru, M., Cojocaru, I.M., Silosi, I., and Vrabie, C.D. (2011). Manifestations of systemic lupus erythematosus. Maedica (Bucur) *6*, 330-336.
35. Degn, S.E., van der Poel, C.E., Firl, D.J., Ayoglu, B., Al Qureshah, F.A., Bajic, G., Mesin, L., Reynaud, C.A., Weill, J.C., Utz, P.J., et al. (2017). Clonal Evolution of Autoreactive Germinal Centers. Cell *170*, 913-926.e919. 10.1016/j.cell.2017.07.026.
36. Musette, P., and Bouaziz, J.D. (2018). B Cell Modulation Strategies in Autoimmune Diseases: New Concepts. Front Immunol *9*, 622. 10.3389/fimmu.2018.00622.
37. Lee, D.S.W., Rojas, O.L., and Gommerman, J.L. (2021). B cell depletion therapies in autoimmune disease: advances and mechanistic insights. Nat Rev Drug Discov *20*, 179-199. 10.1038/s41573-020-00092-2.
38. Clark, E.A., and Giltiay, N.V. (2018). CD22: A Regulator of Innate and Adaptive B Cell Responses and Autoimmunity. Front Immunol *9*, 2235. 10.3389/fimmu.2018.02235.
39. Müller, J., Lunz, B., Schwab, I., Acs, A., Nimmerjahn, F., Daniel, C., and Nitschke, L. (2015). Siglec-G Deficiency Leads to Autoimmunity in Aging C57BL/6 Mice. J Immunol *195*, 51-60. 10.4049/jimmunol.1403139.
40. Brzezicka, K.A., and Paulson, J.C. (2022). Impact of Siglecs on autoimmune diseases. Mol Aspects Med, 101140. 10.1016/j.mam.2022.101140.
41. Poe, J.C., and Tedder, T.F. (2012). CD22 and Siglec-G in B cell function and tolerance. Trends Immunol *33*, 413-420. 10.1016/j.it.2012.04.010.
42. Holers, V.M., and Boackle, S.A. (2004). Complement receptor 2 and autoimmunity. Curr Dir Autoimmun *7*, 33-48. 10.1159/000075686.
43. Song, W.C. (2006). Complement regulatory proteins and autoimmunity. Autoimmunity *39*, 403-410. 10.1080/08916930600739647.
44. Engler, J.B., Kursawe, N., Solano, M.E., Patas, K., Wehrmann, S., Heckmann, N., Lühder, F., Reichardt, H.M., Arck, P.C., Gold, S.M., and Friese, M.A. (2017). Glucocorticoid receptor in T cells mediates protection from autoimmunity in pregnancy. Proc Natl Acad Sci U S A *114*, E181-E190. 10.1073/pnas.1617115114.
45. Oostindie, S.C., van der Horst, H.J., Lindorfer, M.A., Cook, E.M., Tupitza, J.C., Zent, C.S., Burack, R., VanDerMeid, K.R., Strumane, K., Chamuleau, M.E.D., et al. (2019). CD20 and CD37 antibodies synergize to activate complement by Fc-mediated clustering. Haematologica *104*, 1841-1852. 10.3324/haematol.2018.207266.
46. Toh, B.H., Kyaw, T., Tipping, P., and Bobik, A. (2013). Immune regulation by CD52-expressing CD4 T cells. Cell Mol Immunol *10*, 379-382. 10.1038/cmi.2013.35.
47. Mensali, N., Grenov, A., Pati, N.B., Dillard, P., Myhre, M.R., Gaudernack, G., Kvalheim, G., Inderberg, E.M., Bakke, O., and Wälchli, S. (2019). Antigen-delivery through invariant chain (CD74) boosts CD8 and CD4 T cell immunity. Oncoimmunology *8*, 1558663. 10.1080/2162402X.2018.1558663.
48. Yu, B., Qi, Y., Li, R., Shi, Q., Satpathy, A.T., and Chang, H.Y. (2021). B cell-specific XIST complex enforces X-inactivation and restrains atypical B cells. Cell *184*, 1790-1803.e1717. 10.1016/j.cell.2021.02.015.
49. Satoh, M., Tanaka, S., Ceribelli, A., Calise, S.J., and Chan, E.K. (2017). A Comprehensive Overview on Myositis-Specific Antibodies: New and Old Biomarkers in Idiopathic Inflammatory Myopathy. Clin Rev Allergy Immunol *52*, 1-19. 10.1007/s12016-015-8510-y.
50. Chen, M., Zhou, Y., Xue, M., Zhu, R., Jing, L., Lin, L., He, C., and Qin, Y. (2022). Anti-HMGB1 antibody is a potential characteristic autoantibody for Sjögren's syndrome. Sci Rep *12*, 6020. 10.1038/s41598-022-10007-3.
51. Takeuchi, A., Matsushita, T., Kaji, K., Okamoto, Y., Yasui, M., Hirata, M., Oishi, N., Higashi, A., Seishima, M., Asano, T., et al. (2017). Autoantibody to scaffold attachment

factor B (SAFB): A novel connective tissue disease-related autoantibody associated with interstitial lung disease. J Autoimmun *76*, 101-107. 10.1016/j.jaut.2016.09.006.

52. Yang, L., Fujimoto, M., Murota, H., Serada, S., Honda, H., Yamada, K., Suzuki, K., Nishikawa, A., Hosono, Y., Yoneda, Y., et al. (2015). Proteomic identification of heterogeneous nuclear ribonucleoprotein K as a novel cold-associated autoantigen in patients with secondary Raynaud's phenomenon. Rheumatology (Oxford) *54*, 349-358. 10.1093/rheumatology/keu325.
53. Qi, Z., Takamatsu, H., Espinoza, J.L., Lu, X., Sugimori, N., Yamazaki, H., Okawa, K., and Nakao, S. (2010). Autoantibodies specific to hnRNP K: a new diagnostic marker for immune pathophysiology in aplastic anemia. Ann Hematol *89*, 1255-1263. 10.1007/s00277-010-1020-3.
54. Hasler, D., Lehmann, G., Murakawa, Y., Klironomos, F., Jakob, L., Grässer, F.A., Rajewsky, N., Landthaler, M., and Meister, G. (2016). The Lupus Autoantigen La Prevents Mis-channeling of tRNA Fragments into the Human MicroRNA Pathway. Mol Cell *63*, 110-124. 10.1016/j.molcel.2016.05.026.
55. Mousavi, M.J., Mahmoudi, M., and Ghotloo, S. (2020). Escape from X chromosome inactivation and female bias of autoimmune diseases. Molecular Medicine *26*, 127. 10.1186/s10020-020-00256-1.
56. Syrett, C.M., Paneru, B., Sandoval-Heglund, D., Wang, J., Banerjee, S., Sindhava, V., Behrens, E.M., Atchison, M., and Anguera, M.C. (2019). Altered X-chromosome inactivation in T cells may promote sex-biased autoimmune diseases. JCI Insight *4*. 10.1172/jci.insight.126751.
57. Walsh, S.J., and Rau, L.M. (2000). Autoimmune diseases: a leading cause of death among young and middle-aged women in the United States. Am J Public Health *90*, 1463-1466. 10.2105/ajph.90.9.1463.
58. Lerner, A., Patricia, W., and Matthias, T. (2015). The World Incidence and Prevalence of Autoimmune Diseases is Increasing. International Journal of Celiac Disease *3*, 151-155. 10.12691/ijcd-3-4-8.
59. Dinse, G.E., Parks, C.G., Weinberg, C.R., Co, C.A., Wilkerson, J., Zeldin, D.C., Chan, E.K.L., and Miller, F.W. (2020). Increasing Prevalence of Antinuclear Antibodies in the United States. Arthritis Rheumatol *72*, 1026-1035. 10.1002/art.41214.
60. Jury, E.C., D'Cruz, D., and Morrow, W.J.W. (2001). Autoantibodies and overlap syndromes in autoimmune rheumatic disease. Journal of Clinical Pathology *54*, 340-347. 10.1136/jcp.54.5.340.
61. Rosen, A., and Casciola-Rosen, L. (2016). Autoantigens as Partners in Initiation and Propagation of Autoimmune Rheumatic Diseases. Annu Rev Immunol *34*, 395-420. 10.1146/annurev-immunol-032414-112205.
62. Arbuckle, M.R., McClain, M.T., Rubertone, M.V., Scofield, R.H., Dennis, G.J., James, J.A., and Harley, J.B. (2003). Development of autoantibodies before the clinical onset of systemic lupus erythematosus. N Engl J Med *349*, 1526-1533. 10.1056/NEJMoa021933.
63. Dou, D.R., Zhao, Y., Abe, B., Li, R., Zaba, L.C., Aren, K., Carns, M., Chung, L.S., Hinchcliff, M., and Chang, H.Y. (2021). Distinct T cell chromatin landscapes in scleroderma subtypes. bioRxiv, 2021.2001.2010.426131. 10.1101/2021.01.10.426131.
64. Xu, G.J., Shah, A.A., Li, M.Z., Xu, Q., Rosen, A., Casciola-Rosen, L., and Elledge, S.J. (2016). Systematic autoantigen analysis identifies a distinct subtype of scleroderma with coincident cancer. Proceedings of the National Academy of Sciences *113*, E7526-E7534. 10.1073/pnas.1615990113.
65. Giltiay, N.V., Shu, G.L., Shock, A., and Clark, E.A. (2017). Targeting CD22 with the monoclonal antibody epratuzumab modulates human B-cell maturation and cytokine production in response to Toll-like receptor 7 (TLR7) and B-cell receptor (BCR) signaling. Arthritis Research & Therapy *19*, 91. 10.1186/s13075-017-1284-2.

66. Dörner, T., Shock, A., Goldenberg, D.M., and Lipsky, P.E. (2015). The mechanistic impact of CD22 engagement with epratuzumab on B cell function: Implications for the treatment of systemic lupus erythematosus. Autoimmun Rev *14*, 1079-1086. 10.1016/j.autrev.2015.07.013.
67. Li, J., Wei, M.M., Song, Q., Guo, X.H., Shao, L., and Liu, Y. (2019). Anti-CD22 epratuzumab for systemic lupus erythematosus: A systematic review and meta-analysis of randomized controlled trials. Exp Ther Med *18*, 1500-1506. 10.3892/etm.2019.7630.
68. Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods. 10.1038/nmeth.4396.
69. Fiorentino, D.F., Mecoli, C.A., Rosen, M.C., Chung, L.S., Christopher-Stine, L., Rosen, A., and Casciola-Rosen, L. (2022). Immune responses to CCAR1 and other dermatomyositis autoantigens are associated with attenuated cancer emergence. J Clin Invest *132*. 10.1172/JCI150201.
70. Lundberg, I.E., Tjärnlund, A., Bottai, M., Werth, V.P., Pilkington, C., de Visser, M., Alfredsson, L., Amato, A.A., Barohn, R.J., Liang, M.H., et al. (2017). 2017 European League Against Rheumatism/American College of Rheumatology Classification Criteria for Adult and Juvenile Idiopathic Inflammatory Myopathies and Their Major Subgroups. Arthritis Rheumatol *69*, 2271-2282. 10.1002/art.40320.
71. Fiorentino, D.F., Gutierrez-Alamillo, L., Hines, D., Yang, Q., and Casciola-Rosen, L. (2019). Distinct dermatomyositis populations are detected with different autoantibody assay platforms. Clin Exp Rheumatol *37*, 1048-1051.
72. Hochberg, M.C. (1997). Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum *40*, 1725. 10.1002/art.1780400928.
73. Petri, M., Orbai, A.M., Alarcón, G.S., Gordon, C., Merrill, J.T., Fortin, P.R., Bruce, I.N., Isenberg, D., Wallace, D.J., Nived, O., et al. (2012). Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. Arthritis Rheum *64*, 2677-2686. 10.1002/art.34473.