SUPPLEMENTAL METHODS

Whole peptidome array and statistical analysis

In order to account for signal and noise characteristics of the peptide array, we used a large-scale testing tool, MixTwice [1], and r-value [2] to prioritize peptides for differential signal intensity between two groups. These empirical Bayesian tools enable ranking and calibration via local false discovery rate (locFDR) in low signal/noise settings by accounting for shared attributes of peptide-specific sampling distributions.

We assigned peptides a locFDR for sensitive filtering, combined with data on binding affinity, protein context, and peptide sequence. We defined a nearest neighbor peptide as on the same protein and at the immediate neighboring position. The nearest neighbor locFDR of a peptide is the averaged locFDR of its nearest neighbor peptide(s). We used a combination of r-value <0.01, locFDR <0.01, and nearest neighbor locFDR <0.05 on peptides transformed by empirical cumulative distribution function.

Selection of peptides for ELISA validation of whole peptidome array

We narrowed our peptides by requiring a fold increase in SSA- SjD versus controls of at least 10, at least two significant peptides to be bound in the same protein, and at least half of SjD participants bound more or less than a predefined threshold (mean plus one standard deviation of all peptide signals on the array). The resulting 24 peptides (Figure 1) were synthesized and used in ELISA to validate array findings with sera from the same participants.

Selection of peptides from the internal ELISA validation of the array for external validation

Among peptides bound more by SSA- SjD than control serum IgG on the array validation ELISA, we selected peptides for external validation according to the following criteria: (1) average IgG binding at an optical density (OD) of ≥ 0.05 for SSA- SjD participants and (2) IgG binding of at least half of SSA- SjD participants with an OD greater than the standard error of

the mean (SEM) of the control participants (Figure 1). None of the peptides that were bound less by SSA- SjD than control IgG met criteria to proceed to external validation.

Statistical analysis of ELISA results

Functional annotations and protein domains were generated using NIH DAVID 6.8 online tool to determine function (biological, cellular, and molecular processes) and protein domains (InterPro, SMART, UP_KW_Domain) of top peptides (website: https://david-d.ncifcrf.gov/) [3, 4]. Motif analysis was performed with Meme Suite [5] and PROSITE [6-8] was used to identify proteins containing the identified motifs.

Mann-Whitney, Kruskal-Wallis and Wilcoxon rank-sum tests (2-sided) were used to compare adjusted optical density of each peptide among groups of interest (Graphpad Prism software [Graphpad software, La Jolla, CA; R v4.2.2]). Two-group comparisons used one-sided p-values, the direction being confirmed from the initial array findings and internal validation. The Benjamini-Hochberg method was used to form adjusted p-values (q-values) that are adjusted for false discovery rate (FDR) within each block of tests on the (15) peptides entering external validation. The Nemenyi test was used to adjust for multiple comparison of SSA- SjD, SSA+ SjD, and the Augmented Control groups (combined control with SLE and RA samples).

SICCA registry/repository immunoassays

Serology was performed in a standardized fashion for all SICCA registry participants. As described previously in the literature, the first 876 participants were tested for anti-SSA and -SSB with Bio-Rad Autoimmune EIA (Bio-Rad, Hercules, California, USA) and reported as a semiquantitative enzyme immunoassay using purified native antigens. Results were reported as either enzyme units or index values. The next 2,421 participants were tested for anti-SSA and -SSB with a newly introduced Bio-Rad Bioplex 2200 multiplex flow immunoassay. This assay used Ro52 as a recombinant antigen while Ro60 and SSB were native antigens. Results were

reported in international units [9]. Participants who had either one or both Ro52 or Ro60 positive, were classified as anti-SSA antibody positive.

ELISA

We first optimized serum dilutions and peptide concentrations to optimize binding. Biotinylated peptides (Kitchener, Ontario, Canada) were prepared per the manufacturer's directions to a concentration of 1 mg/mL. High-binding ELISA plates (Costar 3590, Corning, Corning, NY) were coated with 0.1 mL of 5 µg/mL streptavidin (Thermofisher, Waltham, MA) diluted in PBS overnight at 4° Celsius. After washing twice with 200 µL of 1x PBS, biotinylated peptide was diluted 1:500 in 1x PBS, 0.1 mL was added to the wells and the plate incubated at room temperature for 1 hour. After incubation, the plate was washed three times with 0.2% Tween-20 in 1x PBS. We blocked the wells with 0.1 mL of 5% non-fat dried milk in 0.2% Tween-20 in PBS for 2.5 hours at room temperature. Next, the serum samples and controls were diluted 1:100 in 5% non-fat dried milk in 0.2% Tween-20 in PBS and 0.1 mL was added in duplicate to the plate overnight at 4° Celsius. The plate was then washed four times with 0.2% Tween-20 in 1x PBS. Next, an HRP-conjugated mouse anti-human IgG clone JDC-10 (Southern Biotech, Oxmoore Blvd, Birmingham, AL) was diluted 1:5000 in 5% non-fat dried milk in 0.2% Tween-20 in PBS for 1 hour at room temperature in the dark. The plate was washed four times with 0.2% Tween-20 in 1x PBS and 0.1 ml of TMB-Slow ELISA formulation (ThermoFisher Scientific, Coraopolis, Pennsylvania, USA) was added to each well and developed in the dark at room temperature for 15 minutes. 0.2 M H2SO4 stop solution was added and the plate was read at 450 and 540 nm. Coated plates at either peptide concentrations of 5 ug/mL or 1 ug/mL with pbs alone wells. Overnight at 4 degree. Washed twice with PBS. Blocked with 200 uL 5% milk and 0.2% tween. Blocked 2.5 hrs. Diluted serum 1:25 and 1:100 for each plate. Overnight at 4 degrees. Detection with 1:5000. Standards were run with polyclonal antibodies from 500 ng/mL up to 50 ng/mL and 5 ng/mL in milk. Adjusted ODs were reported after adjusting for the

following controls: subtracting 450-540 (correcting for optical imperfections), subtract OD for wells with no peptide or serum (background), subtract OD for wells with just serum and no peptide (nonspecific antibody binding), and dividing by a pooled positive control (to account for plate to plate variation).

Inhibition Assays

We incubated anti-DTD2 antibodies with various concentrations of DTD2 peptide or a peptide that showed low reactivity (a peptide from GASP1). The optimum serum dilution was determined in preliminary experiments to be 1:100 (50% max binding) using two sera samples. The final added peptide concentrations varied from 2-100 µg/mL. The mixture of serum and peptides as incubated for 1 hour at 37 degrees C and then overnight at 4 degrees C. ELISA was performed for reactivity against the peptides at all the dilutions. Inhibition percentages were performed by the following equations [(OD 0 mg/mL)-(OD XYZ mg/mL)]/(OD 1 mg/mL) x 100% for each serum as previously reported [10-12]. For homologous inhibition, ELISA was performed for the same antigen as the incubation for heterologous inhibition, the peptide incubation and ELISA targets differed.

Anti-DTD2 Protein ELISA

High-binding ELISA plates (Costar 3590, Corning, Corning, NY) were coated with 1-5 ug/mL recombinant DTD2 protein (Novus Biologicals, Centennial, CO, USA) in 1x PBS or 1x PBS, overnight at 4 Celsius. Then, the wells were washed twice with 200 µL of 1x PBS and blocked with 100 µL of 5% non-fat dried milk in 0.2% Tween-20 in PBS for 2.5 hours at room temperature. Serum samples were diluted from 1:25-1:100 in 5% non-fat dried milk in 0.2% Tween-20 in PBS, overnight at 4 Celsius. The plate was then washed four times with 0.2% Tween-20 in 1x PBS. Next, an HRP-conjugated mouse anti-human IgG clone JDC-10 (Southern Biotech, Oxmoore Blvd, Birmingham, AL) was diluted 1:5000 in 5% non-fat dried milk in 0.2%

Tween-20 in PBS for 1 hour at room temperature in the dark. The plate was washed four times with 0.2% Tween-20 in 1x PBS and 0.1 ml of TMB-Slow ELISA formulation (ThermoFisher Scientific, Coraopolis, Pennsylvania, USA) was added to each well and developed in the dark at room temperature for 15 minutes. 0.2 M H2SO4 stop solution was added and the plate was read at 450 and 540 nm.

Immunohistochemistry (IHC)

Human minor salivary gland biopsy tissues were fixed in paraformaldehyde, embedded in paraffin and cut into 5 μm sections. The paraffin-embedded tissue sections were baked at 60 degrees C for 20 minutes. Deparaffinization was completed by washing the slides washed three times in xylene, followed by rehydration through a series of ethanol solutions (100%, 2 washes; 95%, 2 washes; 70%, 1 wash) and ddH_2O. Antigen retrieval was completed using a Biocare Medical (Pacheco, CA) decloaking chamber, according to manufacturer's protocol. Briefly, the slides were placed in polyethylene Coplin staining jars containing 10 mM citrate buffer, pH 6.0. The tissue sections were heated to 125 degrees C for 2 minutes at 22 psi. The slides were subsequently washed 5 times with ddH_2O and incubated for 5 minutes in 1X Tris-buffered saline with 0.1% Tween (TBST). Primary DTD2 rabbit polyclonal antibody (Novusbio, Centennial, CO; catalog number NBP2-14375), diluted 1:20 in 1X phosphate buffered saline (PBS), was applied to the tissue sections and incubated overnight at 4 degrees C. The following morning, the slides were washed three times, 6 minutes per wash, in 1X TBST. Endogenous peroxidase activity was blocked by incubating the slides in 3% H_2O_2 (Aqua Solutions, Inc., Deer Park, TX; catalog number UC208) for 10 minutes. Slides were washed 2 times in 1X TBST, for 5 minutes per wash. HRP-conjugated horse anti-rabbit secondary antibody (ImmPRESS, Vector Laboratories, Newark, CA; catalog number MP-7401-50) was applied to the slides and incubated for 45 minutes at room temperature. The slides were washed three times with 1X TBST, 6 minutes per wash, and incubated for 4 minutes at room temperature with IMPACT DAB chromogen

peroxidase substrate (Vector Laboratories, Newark, CA; catalog number SK-4105), diluted according to manufacturer's instructions. The tissue sections (slides) were washed with ddH_2O for 5 minutes. The sections were counterstained with hematoxylin for 1.5 minutes, which was followed by washing and dehydration steps (ddH2O; 1% HCl/70% ethanol; ddH2O; 1% ammonia in ddH2O; ddH2O; 95% ethanol, 2 washes; 100% ethanol, 2 washes; xylene, 2 washes). Following the final xylene wash, the tissues were sealed and coverslipped using Cytoseal XYL (Epredia, Kalamazoo, MI; catalog number 8312-4) (7 drops per slide).

Image Acquisition and Analysis

Light microscopy images were obtained on an Olympus BX51 microscope with an Olympus KP70 camera (Olympus America, Inc., Waltham, MA). All Images were captured using a 10X objective. Due to the variable size of the respective salivary glands, the images were unbiasedly cropped to a uniform area and saved as individual TIFF files for DAB chromogenic analysis. Semi-quantitative H DAB chromogenic staining was measured using Fiji Image J2 [13], with minimum/maximum threshold values of 123/172, respectively. Data, percent area (± SD), was plotted using GraphPad Prism (V10.102). Statistical significance was calculated using an ordinary one-way Anova with Tukey's multiple comparisons test.

REFERENCES

1 Zheng Z, Mergaert AM, Ong IM, Shelef MA, Newton MA. MixTwice: large-scale hypothesis testing for peptide arrays by variance mixing. Bioinformatics 2021;37(17):2637-43.
2 Henderson NC, Newton MA. Making the cut: improved ranking and selection for large-scale inference. J R Stat Soc Series B Stat Methodol 2016;78(4):781-804.
3 Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 2009;4(1):44-57.
4 Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1-13.
5 Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. Nucleic Acids Res 2015;43(W1):W39-49.
6 Sigrist CJ, de Castro E, Cerutti L, et al. New and continuing developments at PROSITE. Nucleic Acids Res. 2013;41(Database issue):D344-7.
7 Sigrist CJ, Cerutti L, Hulo N, et al. PROSITE: a documented database using patterns and profiles as motif descriptors. Brief Bioinform 2002;3(3):265-74.

8 Hulo N, Bairoch A, Bulliard V, et al. The PROSITE database. Nucleic Acids Res. 2006;34(Database issue):D227-30.
9 Infantino M, Bentow C, Seaman A, et al. Highlights on novel technologies for the detection of antibodies to Ro60, Ro52, and SS-B. Clin. Dev. Immunol. 2013;2013:978202.
10 Goules JD, Goules AV, Tzioufas AG. Fine specificity of anti-citrullinated peptide antibodies discloses a heterogeneous antibody population in rheumatoid arthritis. Clin. Exp. Immunol. 2013;174(1):10-7.
11 Priest JW, Plucinski MM, Huber CS, et al. Specificity of the IgG antibody response to Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae, and Plasmodium ovale MSP1(19) subunit proteins in multiplexed serologic assays. Malar. J. 2018;17(1):417.
12 Tsiakalos A, Routsias JG, Kordossis T, Moutsopoulos HM, Tzioufas AG, Sipsas NV. Fine epitope specificity of anti-erythropoietin antibodies reveals molecular mimicry with HIV-1 p17 protein: a pathogenetic mechanism for HIV-1-related anemia. J. Infect. Dis. 2011;204(6):902-11.
13 Crowe AR, Yue W. Updated: Semi-quantitative Determination of Protein Expression Using Immunohistochemistry Staining and Analysis. Bio Protoc 2023;2(13).