



OPEN

Temporal reproducibility of IgG and IgM autoantibodies in serum from healthy women

T. V. Clendenen^{1,3}, S. Hu^{2,3}, Y. Afanasyeva¹, M. Askenazi^{1,2}, K. L. Koenig¹, T. Hulett², M. Liu¹, S. Liu², F. Wu¹, A. Zeleniuch-Jacquotte¹ & Y. Chen¹✉

Autoantibodies are present in healthy individuals and altered in chronic diseases. We used repeated samples collected from participants in the NYU Women's Health Study to assess autoantibody reproducibility and repertoire stability over a one-year period using the HuProt array. We included two samples collected one year apart from each of 46 healthy women (92 samples). We also included eight blinded replicate samples to assess laboratory reproducibility. A total of 21,211 IgG and IgM autoantibodies were interrogated. Of those, 86% of IgG ($n = 18,303$) and 34% of IgM ($n = 7,242$) autoantibodies showed adequate lab reproducibility (coefficient of variation [CV] < 20%). Intraclass correlation coefficients (ICCs) were estimated to assess temporal reproducibility. A high proportion of both IgG and IgM autoantibodies with CV < 20% (76% and 98%, respectively) showed excellent temporal reproducibility (ICC > 0.8). Temporal reproducibility was lower after using quantile normalization suggesting that batch variability was not an important source of error, and that normalization removed some informative biological information. To our knowledge this study is the largest in terms of sample size and autoantibody numbers to assess autoantibody reproducibility in healthy women. The results suggest that for many autoantibodies a single measurement may be used to rank individuals in studies of autoantibodies as etiologic markers of disease.

Novel technologies have led to the recognition that numerous antibodies against self molecules, autoantibodies (AABs), are present in healthy, free of auto-immune diseases, individuals^{1,2}. AABs are highly individualized, accumulate throughout life and have been observed to increase with age, including among individuals without autoimmune disease³. Some AABs, particularly IgMs, are characterized by poly-affinity for various antigens, though many highly specific and high binding affinity IgG and IgM AABs have been identified⁴. AABs are thought to play a role in apoptosis and apoptotic waste cleanup and may have a role in the regulation of some anti-inflammatory responses⁴⁻⁶.

It has been hypothesized that AAB profiles may be altered in various diseases, including cardiovascular disease, dementia, and cancer⁷⁻⁹. AABs are thought to play a role in the recognition and removal of malignant and pre-cancerous cells⁹. To be able to interpret the presence and patterns of AABs in patients with various diseases, it is important to identify AABs that are commonly found in healthy individuals. Estimates of the stability of AABs over time in healthy individuals are also needed to conduct prospective studies focusing on incidence of chronic diseases. AABs that are stable in healthy individuals but increase or decrease when individuals develop a specific disease may be good candidate biomarkers for early detection or progression of disease.

Only with the recent development of auto-antigen arrays, has it been possible to examine large numbers of AABs in healthy individuals. Nagele et al.³ observed an abundance of AABs (IgG) in 57 healthy individuals on a 10,000 human protein microarray. Two studies from the same group examined IgG autoreactivity to 335 protein fragments and found that AAB profiles were relatively stable over one year in samples collected from four individuals^{3,10}. To date, there are no studies that include large numbers of IgG AABs and longitudinal samples, and no characterization of IgM AAB profiles in healthy women.

To address this knowledge gap, we conducted a study to identify both IgM and IgG AABs commonly found in healthy women, using arrays of over 21,000 proteins and protein fragments (herein referred to as autoantigens), and examined the temporal reproducibility of their measurement using longitudinally collected samples.

¹Department of Population Health, Division of Epidemiology, NYU Langone Health School of Medicine, 180 Madison Ave. 5th Floor, New York, NY 10016, USA. ²CDI Laboratories, Baltimore, MD, USA. ³These authors contributed equally: T. V. Clendenen and S. Hu. ✉email: yu.chen@nyulangone.org

Methods

Study subjects. We used blood samples from participants in the NYU Women's Health Study. The NYU Women's Health Study is a cohort of 14,274 women enrolled between 1985 and 1991 at a breast cancer screening center in New York City. Participants are followed up through self-completed questionnaires every few years and linkage with the National Death Index and tumor registries of New York, New Jersey, and Florida (where over 85% of the participants reside). We collected blood samples at enrollment and at annual repeat screening visits until 1991. Serum samples were stored at -80°C . Two or more yearly serum samples are available for 52% of the participants. We selected 46 participants in the NYUWHS who: (1) donated two or more blood samples at 1-year intervals; (2) had not been diagnosed with cancer or CVD up to last complete follow-up (2016). This study was approved by the NYU Langone School of Medicine Institutional Review Board in accordance with the policies and regulations governing research with human subjects. Informed consent was obtained from all participants in the NYUWHS.

Assays. Assays were conducted using HuProt arrays (version 3.1) provided by CDI Laboratories, Inc¹¹. Each array comprised 22,817 human proteins, isoform variants, or protein fragments (hereafter referred to as autoantigens). Sera were diluted 1:1,000 and reacted with individual HuProt protein microarrays. Briefly, the HuProt arrays were blocked with blocking solution (5% BSA/1 × TBS-T) at room temperature for 1 h, and then probed with serum samples (diluted 1:1000) at 4°C overnight. The arrays were then washed with 1 × TBS-T for 3 times, 10 min each, and probed with Alexa-647 labeled anti-human IgG Fc gamma fragment specific and Cy3 labeled anti-human IgM Fc5 μ fragment specific secondary antibodies (Jackson ImmunoResearch, West Grove, PA) at room temperature for 1 h, followed by three washes of 1 × TBS-T, 10 min each, and then dried and scanned using a GenePix 4000B scanner (Molecular Devices, Sunnyvale, CA) to create raw TIF images. Spots were then aligned to HuProt array list (.gal) files, and spot readings captured as raw Genepix results (.gpr) files used for data processing. In total, 21,211 autoantigen spots (93%) passed laboratory quality control criteria for spot quality and had their identities confirmed by the laboratory. We focus on that subset for this study.

Eight quality control replicates from a common pool generated using samples from eight healthy NYU Women's Health Study participants were included to assess assay variability. Samples were randomly ordered for laboratory analyses. Samples were identified only by a sample number, so that laboratory personnel could not identify which samples were from the same woman nor which were QC samples.

Data processing. For each probe, the signal intensity is given by the median foreground by Genepix. Because each protein was printed in duplicate, the mean of the two readings was used. Signal intensity values were log₂-transformed (base two).

Statistical analysis. Laboratory variability was assessed by means of the coefficient of variation (CV) using data from the eight blinded replicate quality control samples. All subsequent analyses were limited to AABs with CV < 20%. The intraclass correlation coefficient (ICC) was used to examine temporal reproducibility of the log₂-transformed signal intensity for all AABs. The ICC estimates the fraction of the total variation (within-subject plus between-subject) that is due to between-subject variation^{12,13}. Heatmaps were created using log₂-transformed non-normalized data to examine clustering of samples across AABs using the R package “ComplexHeatmap” and Pearson distance hierarchical clustering of samples and AABs. We also conducted analyses using quantile normalization in the R package “normalize.quantiles” on the log₂-transformed data because this method is commonly used to reduce assay variability¹⁸. To assess the correlation between AAB measurements and the covariates age and BMI, we calculated Pearson correlation coefficients of the average value of the two log₂-transformed repeat AAB measurements with the average log₂-transformed age and BMI from the two visits.

Data availability. The data generated in this study are available within the article and its supplementary data files.

Results

The 46 women included in this study were between the ages of 38 and 68 (mean age 51.7), and 50% were premenopausal. BMI ranged from 19.2 to 35.4 kg/m² (mean BMI 24.1). Twenty-nine (63%) women were never smokers, 10 (22%) were former smokers, and 7 (15%) were current smokers at both blood donations. Consistent with the distributions observed in the full NYUWHS cohort, many of the women included in this analysis reported their race/ethnicity as White (70%, $n = 32$ women), 15% Black ($n = 7$), 4% Asian, and 4% Hispanic (6% of women did not report).

Table 1 reports CVs estimated using the eight blinded replicates from a common sample pool, randomly interspersed among the study samples. A high percentage of the IgG AABs (86%) showed acceptable lab reproducibility (CV < 20%) and 25% showed high lab reproducibility (CV < 10%). However, less than half (34%) of the IgM AABs had a CV < 20% and only 0.4% a CV < 10%.

ICCs were estimated, and heatmaps generated, for the 18,303 IgG and 7242 IgM AAB with acceptable lab reproducibility (CV < 20%). A very high proportion of both of IgG and IgM AABs (76% of IgG and 98% of IgM) showed excellent temporal reproducibility (ICCs > 0.8, Table 2). ICCs computed using quantile-normalized data were lower than the ICCs computed from the raw data (log₂-transformed intensities) (Table 2). However, a high proportion of IgG (69%) and a substantial proportion of IgM (40%) AABs still had an ICC ≥ 0.6 (Table 2).

Heatmaps showing sample clustering (columns) and probe clustering (rows) based on log₂-transformed signal intensity (left) and quantile normalized log₂-transformed signal intensity (right) for AABs with CV < 20%

	Total N	CV, <i>n</i> (%)			
		<10%	10–15%	15–20%	≥20%
IgG	21,200	5344 (25.2)	8560 (40.4)	4399 (20.8)	2897 (13.7)
IgM	21,209	83 (0.4)	1762 (8.3)	5397 (25.4)	13,967 (65.9)

Table 1. Coefficients of variation, $n=8$ replicate QC samples (raw data). 11 IgG autoantigen probes and 2 IgM had no variation and were not included in the above table.

	N		ICC, <i>n</i> (%)					
			≤0.59		0.6–<0.79		0.8–1.0	
IgG	18,303	Log ₂ scale	871	(4.76)	3465	(18.9)	13,967	(76.3)
		Quantile normalized	5736	(31.3)	5460	(29.8)	7107	(38.8)
IgM	7242	Log ₂ scale	14	(0.19)	131	(1.81)	7097	(98.0)
		Quantile normalized	4395	(60.7)	1861	(25.7)	986	(13.6)

Table 2. Intra-class correlation coefficients (limited to autoantibodies with CV <20%).

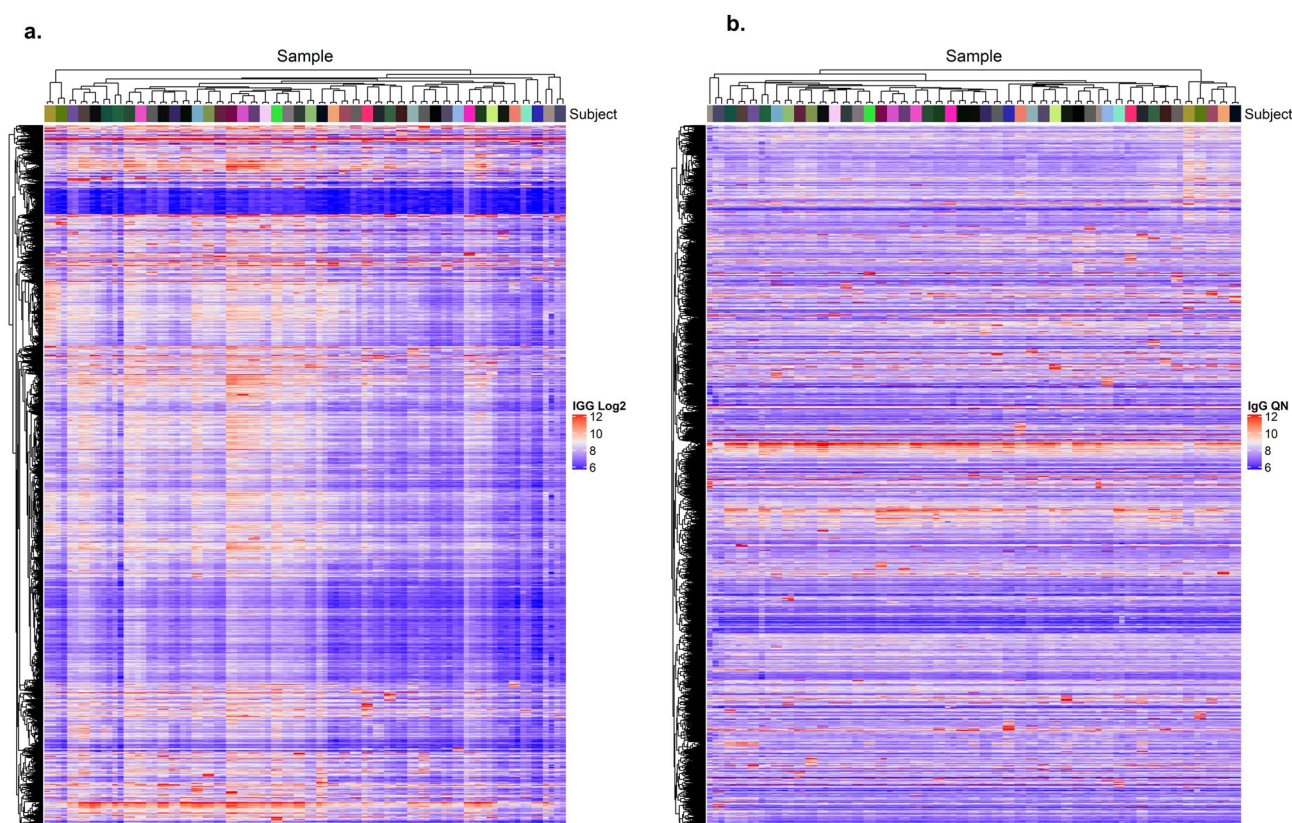


Figure 1. Heatmap of IgG autoantibodies, log₂ transformed (left) and quantile normalized (right) ($n=18,303$ with CV <20%). IgG autoantibody profiling using HuProt arrays of repeat samples collected one year apart from $n=46$ healthy subjects ($n=92$ samples total). Raw log₂ transformed data (a) and quantile normalized data (right) was used to cluster AAbs and samples using Pearson distance hierarchical clustering for 18,303 probes with CV <20%. Top bar annotation colors correspond to individual subjects. (Two samples from each subject clustered together for all 46 subjects using log₂ transformed data and for all but one subject using quantile normalized data).

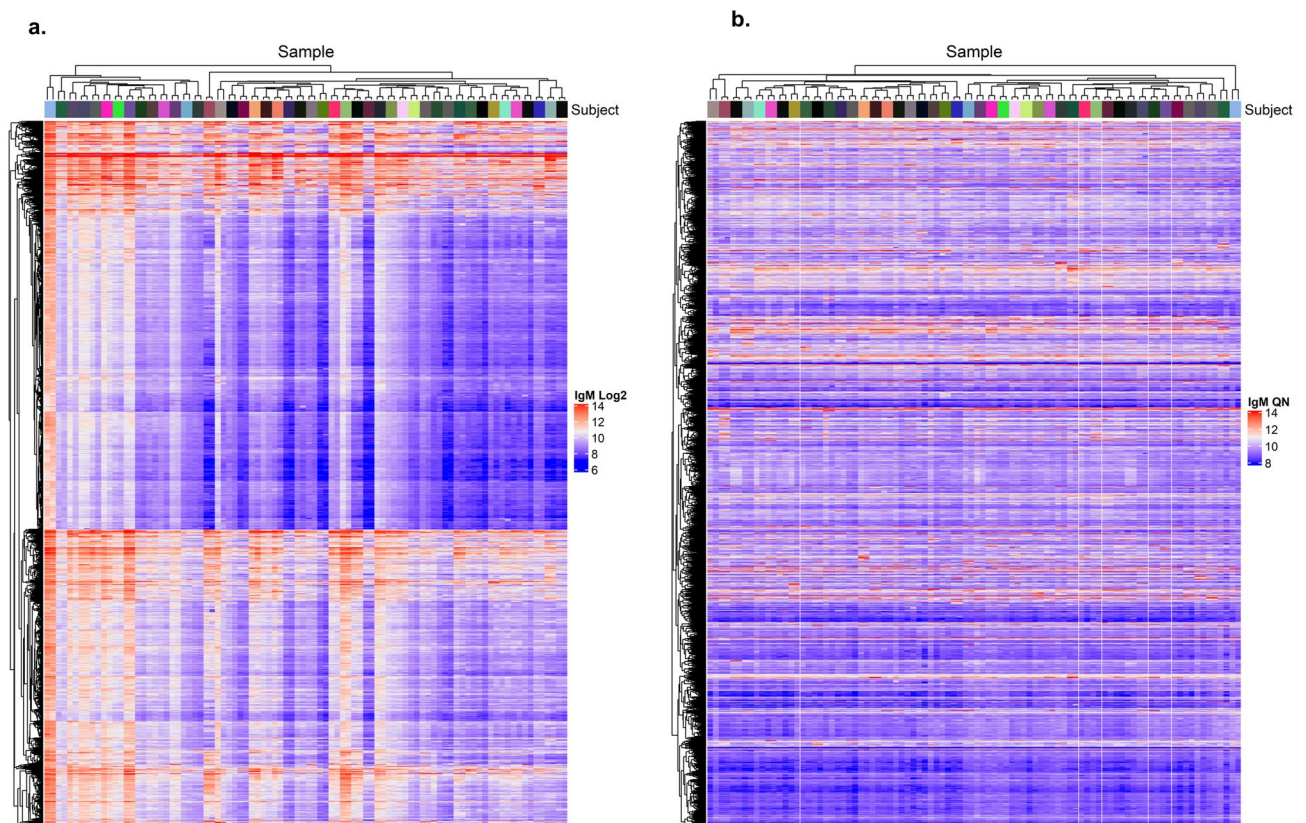


Figure 2. Heatmap of IgM autoantibodies, log₂ transformed (left) and quantile normalized (right) ($n=7,242$ with $CV < 20\%$) IgM autoantibody profiling using HuProt arrays of repeat samples collected one year apart from $n=46$ healthy subjects ($n=92$ samples total). Raw log₂ transformed data (a) and quantile normalized data (right) was used to cluster AAbs and samples using pearson distance higherarchical clustering for 7,242 probes with $CV < 20\%$. Top bar annotation colors correspond to individual subjects. (Two samples from each subject clustered together for all 46 subjects using both log₂ transformed data and quantile normalized data).

are shown in Fig. 1 for IgG and Fig. 2 for IgM. The heatmaps show that IgG and IgM AAb profiles were much more similar for samples collected from the same woman one year apart than for samples collected from different women. This was true both before and after quantile normalization, though clustering was perfect before normalization using both IgG and IgM while one set of paired samples did not cluster together after quantile normalization in the IgG analysis.

We assessed the correlation of AAb measurements with age and BMI. Most of the IgG and IgM AAbs were negatively correlated with age, and most IgM AAbs were negatively correlated with BMI. IgG AAbs were evenly spread between negative and positive correlations with BMI. However, none of the correlations of IgG or IgM AAbs with age or BMI were strong (correlation coefficients $< |0.45|$, Fig. 3).

We assessed whether samples would cluster according to smoking status using heatmaps of AAbs with $CV < 20\%$ and $ICC \geq 0.6$. Smoking status (never, former, current) was not associated with AAb profiles (Fig. 4).

AAb-specific reproducibility estimates and average intensities for IgG and IgM are available in Appendices (see Supplementary IgG and IgM Files, Sheet 1). Those with $CV < 20\%$ and high $ICC (> 0.8)$ both before and after quantile normalization are also shown along with information about their functional role (see Supplementary IGG File, Sheet 2). These AAbs showed a range of functions; for example, IgG AAbs were involved in signal transduction, energy production, cellular mobility, and gene expression regulation. One hnRNP (hnRNPD) and two histone (H2BC and H3A) IgGs were in the top 500 highest AAbs (as measured by mean intensity after log₂ transformation), while most other hnRNPs, histones, GPI, SSAs, Sm, and other commonly expressed autoimmune-disease associated autoantibodies did not have high mean intensities. IgG AAbs that were commonly observed in previous studies of healthy subjects are shown along with our estimates in sheet 3 (IgM AAbs have not yet been studied in healthy individuals). We examined the temporal reproducibility of AAbs frequently observed in healthy individuals in other studies. 38 of the 66 IgG AAbs that were observed in at least 60% of participants in one study 4 had overlapping autoantigen probes on HuProt, and all but three had $ICC > 0.6$. We measured IgGs for eight of 12 other AAbs frequently observed in another study 19, and each of them had at least one probe with $ICC > 0.6$.

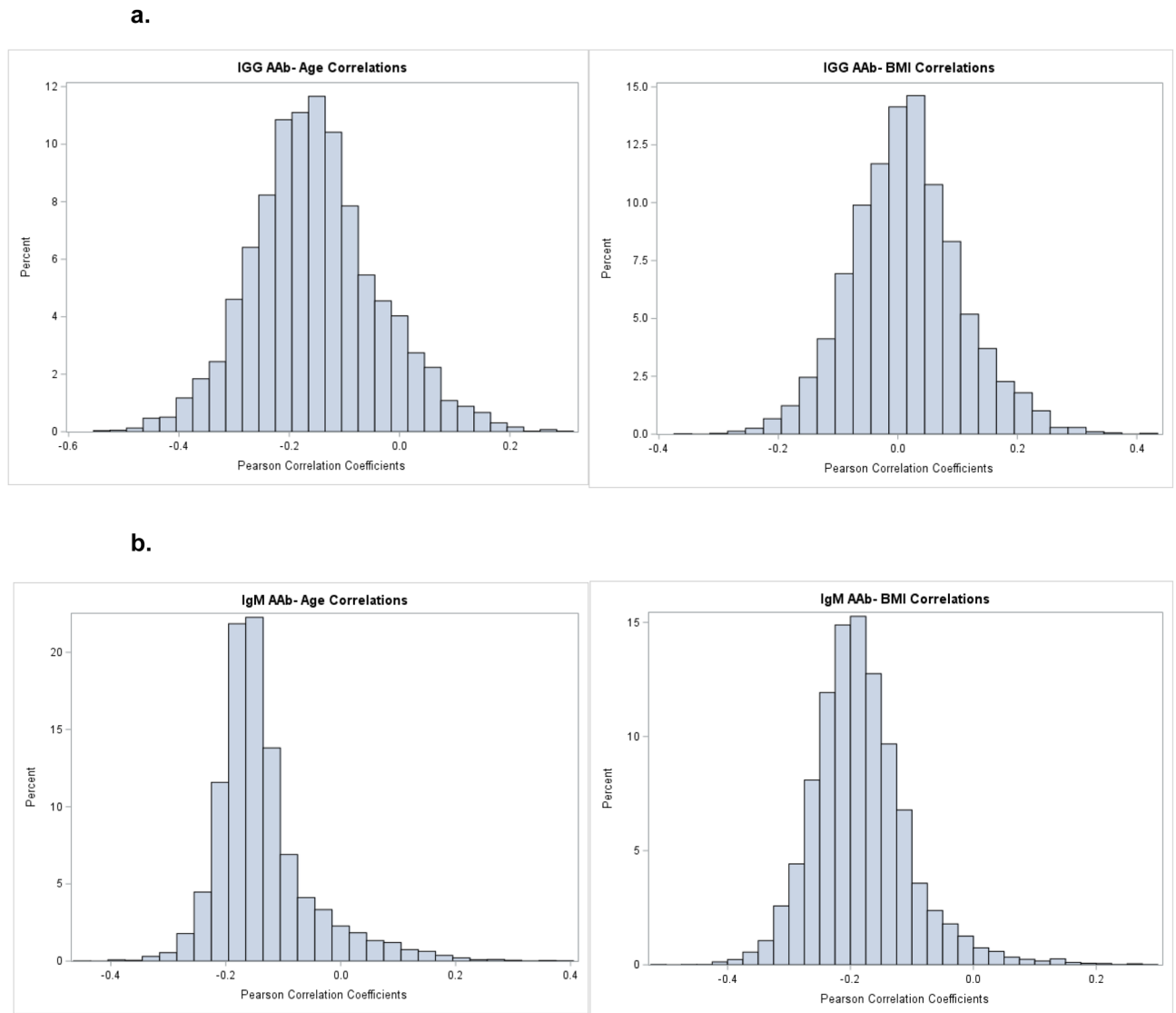


Figure 3. Distribution of Pearson correlation coefficients with age and BMI, restricted to autoantibodies with $CV < 20\%$ and $ICC \geq 0.6$ (a) IgG correlation with age (left) and BMI (right). (b) IgM correlation with age (left) and BMI (right). Histograms of Pearson correlation coefficients for IgG (a) and IgM (b) AAbs with age (left panels) and BMI (right panels). Log₂-transformed raw AAb concentrations correlations with age and BMI were between -0.4 to 0.2 for nearly all IgG and IgM AAbs.

Discussion

Most AAbs in serum samples stored at $-80\text{ }^{\circ}\text{C}$ for over 30 years could be measured with good assay precision using the HuProt array. We also found that samples collected one year apart from the same woman showed more similar AAb profiles than samples collected from different women, indicating that there is AAb repertoire stability in healthy individuals. When considered individually, a majority ($>70\%$) of IgG AAbs and a third of IgM AAbs had low laboratory variability ($CV < 20\%$) and ICCs consistent with adequate to excellent temporal reproducibility among healthy women over a one-year period. Finally, clustering of samples from the same woman was better achieved, and proportion of AAbs with good temporal reliability higher, in analyses of the raw (log₂-transformed) data than the quantile-normalized data.

Autoimmune and chronic diseases have been associated with specific AAb profiles^{8,14,15}. As AAb profiles specific to autoimmune and other chronic diseases are identified and then refined to select the most informative panels associated with disease, ICCs should be considered as AAbs included in such panels should be stable in healthy individuals.

Comparing results from studies using different protein arrays is difficult because of the lability and 3-dimensional nature of proteins, differences in the methodology used to create the array, and in AAb nomenclature. However, our observations that many autoantibodies are detectable in healthy individuals and that an individual's AAb profile is highly reproducible over time is consistent with the few other studies conducted to date. Nagele et al.³ examined about half as many AAbs (9486 IgG) using Invitrogen microarrays among 166 individuals, 57

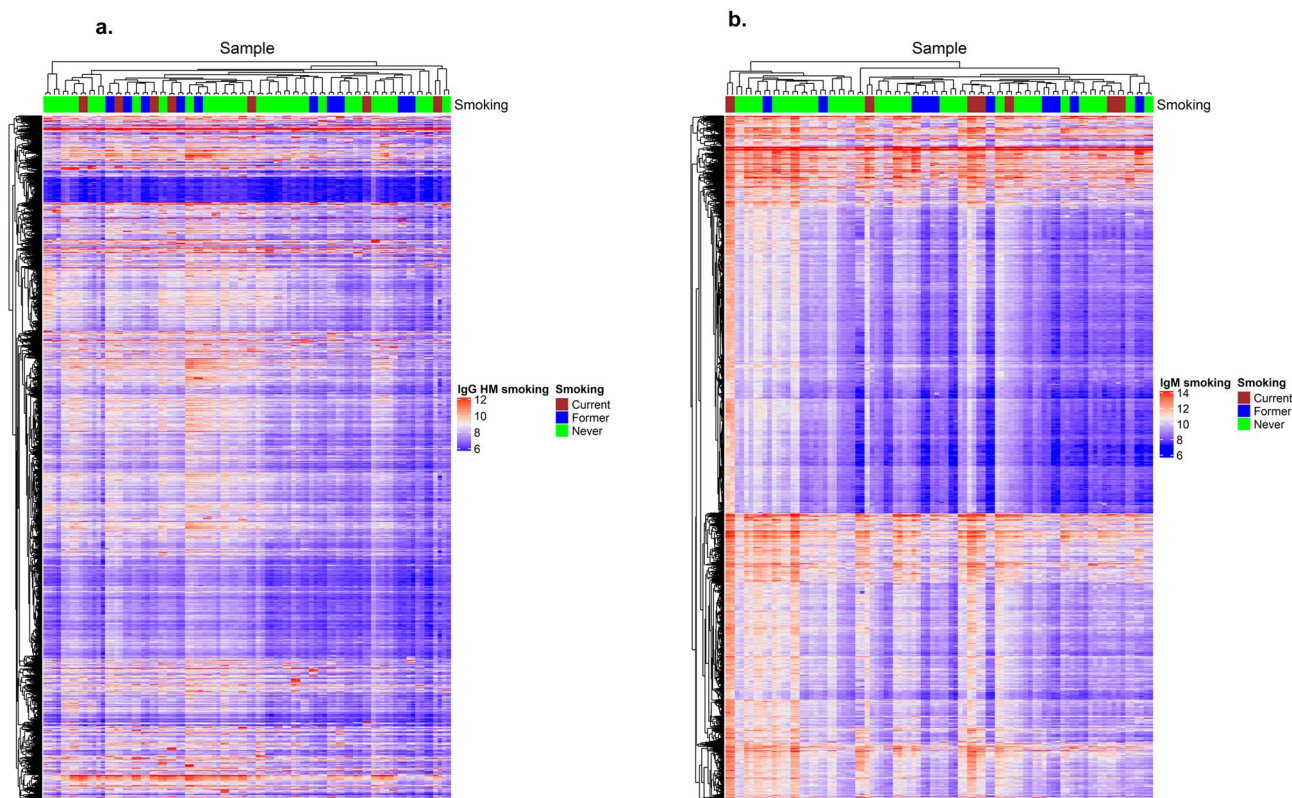


Figure 4. IgG (left) and IgM (right), by smoking status (never, former, current). IgG (a) and IgM (b) autoantibody profiling using HuProt arrays of repeat samples collected one year apart from $n = 46$ healthy subjects ($n = 92$ samples total). Raw log₂ transformed data was used to cluster AAbs and samples using Pearson distance hierarchical clustering for 18,303 IgG and 7,242 IgM AAbs with CV < 20%. Top bar annotation colors correspond to smoking status at blood donation ($n = 7$ current, 10 former, 29 never smokers). (Two samples from each subject clustered together for all 46 subjects, but did not cluster by smoking status).

without chronic disease, and observed the presence of hundreds to thousands of IgG AAbs, even among healthy individuals. Neiman et al.¹⁶ examined 335 protein fragments (204 proteins) selected for their known role in autoimmune disorders in 193 healthy individuals' blood samples taken 4 times over a single year period. An AAb fingerprint was created for each sample, which was consistent over repeat timepoints. Four individual's serum samples were further profiled using an array containing 23,000 protein fragments. For each of the 49 AAbs in these two studies considered to be frequently observed^{3,16} that were also measured in our study, all but three had at least one probe with high temporal reproducibility (see Supplementary IgG File, Sheet 3). Further, consistent with a previous study that examined temporal reproducibility of candidate AAbs individuals and those with cardiovascular disease¹⁷, we observed high ICCs for at least one probe for each of the seven AAbs common to our array that had reported high reproducibility (bottom of Supplementary IgG File, Excel Sheet 3).

IgG AAb differences among cancer cases vs. controls have been assessed using arrays, including HuProt¹⁸. Specifically, AAb biomarkers have been identified for lung, breast, esophageal, gastric, ovarian, colorectal, and hepatic cell cancers^{19–36}. The laboratory variability and temporal reproducibility estimates that we observed for these cancer markers are shown in the Supplementary IgG File (Sheet 4). Of the autoantigen probes for 34 HuProt-based markers from seven studies, 32 markers in our study had at least one autoantigen probe that showed high temporal reproducibility (all except for CLDN ICC: 0.40, PDGFRA ICC: 0.33, and one of five GNAS autoantigen probes ICC: 0.41). In 11 other studies of AAbs and cancer (not using HuProt), 50 AAbs were associated with one or more cancers. Of those, 45 were included in the array used in the present study, met laboratory quality control criteria, and were detected in the NYUWHS samples. Each of the 45, except four (TSC1 ICC: 0.47, two of four autoantigen probes for BDNF ICCs: 0.39, 0.45, one of five for FGFR2 ICC: 0.46, and one of two for PSRC1 ICC: 0.58) had ICCs > 0.6. Four of the markers identified from previous AAb studies that had low ICCs in our study, also had autoantigens for the same target with ICCs > 0.6. A strength of having more than one probe per marker on the array is that many markers are covered with redundancy, which may be important to consider in studies selecting optimal probe sets in future etiologic and early detection studies.

Normalization is commonly used in proteomic analyses, as array variability (also referred to as batch variability) is of concern^{37,38}. Therefore, we conducted an analysis of the quantile-normalized data, in addition to the analysis of the raw (log₂-transformed) data. Though it was not our expectation, clustering of samples from the same woman was not as good using quantile-normalized data as raw data, suggesting that batch variability is not as much of a concern, at least for the HuProt assay, as previously thought. While normalization should

be assessed as part of a standard analysis to address concerns of batch variability, it may obscure some true biological differences.

Our study has several strengths. To our knowledge, it is the largest study to assess the occurrence and temporal reproducibility of AAbs among individuals who were free of cancer and cardiovascular disease not only at the two blood draws, but also in the following 35 years. For many proteins, the HuProt array includes several antigens, providing redundancy and capturing some AAbs to fragmented or variant antigens. This is also the first study that assessed reproducibility of an array of IgM AAbs. The inclusion of eight replicates of a pooled sample interspersed in a blinded fashion within the paired samples allowed us to accurately assess assay precision. CVs were good for most AAbs and suggest that batch effects were not a major concern, particularly for IgG assays. Further, despite not controlling for non-specific binding, we still observed high intra-subject stability in samples collected over time for most AAbs, indicating that noise from sticky binding did not obscure autoantibody profiles.

A limitation of our study is that the sample size was adequate to estimate temporal reproducibility but was insufficient to reliably determine the impact of subject characteristics, such as demographic and lifestyle variables, on the AAb profile. It is not clear why a high proportion of IgM AAbs had high laboratory variability (CV > 20%) despite being assayed simultaneously in the same array. It is possible that both biological and technical factors contributed, including lower specificity binding of most IgM AAbs, the larger size of IgM vs. IgG AAbs, and the relationship of these factors to array characteristics (e.g. antigen spacing)³⁹. Further, all of the antigens spotted on the array are proteins and protein fragments. Because these proteins are expressed in yeast, some post-translational modifications (e.g., glycosylation and phosphorylation) will be present and printed on the array. However, the arrays we used would not have captured autoantibodies directed at non-protein targets, such as DNA, oxidized lipids and other frequent non-protein targets of autoantibodies as well as an extensive array of citrullinated, acetylated, and other post-translationally modified proteins. The number and types of AAb targets are vast and may be included on arrays specific to the target type in the future⁴⁰.

In conclusion, our results show that many AAbs are present and detectable in healthy women and that their levels are stable over a one-year period. AAbs in the blood reflect systemic autoimmunity and can be measured rapidly in large numbers using arrays. Our observations, along with the established role of AAbs in health and disease, suggest that they may be good candidate biomarkers of risk and early detection of cancer and other chronic diseases.

Received: 29 December 2021; Accepted: 30 March 2022

Published online: 13 April 2022

References

- Avrameas, S. & Selmi, C. Natural autoantibodies in the physiology and pathophysiology of the immune system. *J. Autoimmun.* **41**, 46–49. <https://doi.org/10.1016/j.jaut.2013.01.006> (2013).
- Merbl, Y., Zucker-Toledano, M., Quintana, F. J. & Cohen, I. R. Newborn humans manifest autoantibodies to defined self molecules detected by antigen microarray informatics. *J. Clin. Invest.* **117**, 712–718. <https://doi.org/10.1172/JCI29943> (2007).
- Nagele, E. P. *et al.* Natural IgG autoantibodies are abundant and ubiquitous in human sera, and their number is influenced by age, gender, and disease. *PLoS ONE* **8**, e60726. <https://doi.org/10.1371/journal.pone.0060726> (2013).
- Grönwall, C. & Silverman, G. J. Natural IgM: beneficial autoantibodies for the control of inflammatory and autoimmune disease. *J. Clin. Immunol.* **34**(Suppl 1), S12–S21. <https://doi.org/10.1007/s10875-014-0025-4> (2014).
- Panda, S. & Ding, J. L. Natural antibodies bridge innate and adaptive immunity. *J. Immunol.* **194**, 13–20 (2015).
- Zavdy, O., Shoenfeld, Y. & Amital, H. In *Autoantibodies (Third Edition)* (eds Yehuda Shoenfeld, Pier Luigi Meroni, & M. Eric Gershwin) 21–33 (Elsevier, 2014).
- Gibson, L. L. *et al.* Neuronal surface autoantibodies in dementia: a systematic review and meta-analysis. *J. Neurol.* <https://doi.org/10.1007/s00415-020-09825-0> (2020).
- Kobayashi, M., Katayama, H., Fahrman, J. F. & Hanash, S. M. Development of autoantibody signatures for common cancers. *Semin. Immunol.* **47**, 101388. <https://doi.org/10.1016/j.smim.2020.101388> (2020).
- Vollmers, H. P. & Brändlein, S. Natural antibodies and cancer. *New Biotechnol.* **25**, 294–298. <https://doi.org/10.1016/j.nbt.2009.03.016> (2009).
- Tebani, A. *et al.* Integration of molecular profiles in a longitudinal wellness profiling cohort. *Nat. Commun.* **11**, 4487–4487. <https://doi.org/10.1038/s41467-020-18148-7> (2020).
- Jeong, J. S. *et al.* Rapid identification of monospecific monoclonal antibodies using a human proteome microarray. *Mol. Cell Proteom.* <https://doi.org/10.1074/mcp.O111.016253> (2012).
- Shrout, P. E. & Fleiss, J. L. Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* **86**, 420–428. <https://doi.org/10.1037//0033-2909.86.2.420> (1979).
- Pleil, J. D., Wallace, M. A. G., Stiegel, M. A. & Funk, W. E. Human biomarker interpretation: the importance of intra-class correlation coefficients (ICC) and their calculations based on mixed models, ANOVA, and variance estimates. *J. Toxicol. Environ. Health B Crit. Rev.* **21**, 161–180. <https://doi.org/10.1080/10937404.2018.1490128> (2018).
- Zhang, Y. *et al.* Low level antibodies against alpha-tropomyosin are associated with increased risk of coronary heart disease. *Front. Pharmacol.* **11**, 195–195. <https://doi.org/10.3389/fphar.2020.00195> (2020).
- Livingston, B., Bonner, A. & Pope, J. Differences in autoantibody profiles and disease activity and damage scores between childhood- and adult-onset systemic lupus erythematosus: a meta-analysis. *Semin. Arth. Rheum.* **42**, 271–280. <https://doi.org/10.1016/j.semarthrit.2012.05.001> (2012).
- Neiman, M. *et al.* Individual and stable autoantibody repertoires in healthy individuals. *Autoimmunity* **52**, 1–11. <https://doi.org/10.1080/08916934.2019.1581774> (2019).
- Li, X. *et al.* Autoantibody profiling on a plasmonic nano-gold chip for the early detection of hypertensive heart disease. *Proc. Natl. Acad. Sci. USA* **114**, 7089–7094. <https://doi.org/10.1073/pnas.1621457114> (2017).
- Beutgen, V. M., Perumal, N., Pfeiffer, N. & Grus, F. H. Autoantibody biomarker discovery in primary open angle glaucoma using serological proteome analysis (SERPA). *Front. Immunol.* **10**, 381. <https://doi.org/10.3389/fimmu.2019.00381> (2019).
- Yang, L. *et al.* Identification of serum biomarkers for gastric cancer diagnosis using a human proteome microarray. *Mol. Cell Proteom.* **15**, 614–623. <https://doi.org/10.1074/mcp.M115.051250> (2016).

20. Pan, J. *et al.* Integration of IgA and IgG autoantigens improves performance of biomarker panels for early diagnosis of lung cancer. *Mol. Cell Proteom.* **19**, 490–500. <https://doi.org/10.1074/mcp.RA119.001905> (2020).
21. Pan, J. *et al.* Identification of serological biomarkers for early diagnosis of lung cancer using a protein array-based approach. *Mol. Cell Proteom.* **16**, 2069–2078. <https://doi.org/10.1074/mcp.RA117.000212> (2017).
22. Pan, J. *et al.* Discovery and validation of a serologic autoantibody panel for early diagnosis of esophageal squamous cell carcinoma. *Cancer Epidemiol. Biomark. Prev.* **28**, 1454–1460. <https://doi.org/10.1158/1055-9965.EPI-18-1269> (2019).
23. Zhang, S. *et al.* Autoantibody signature in hepatocellular carcinoma using seromics. *J. Hematol. Oncol.* **13**, 85. <https://doi.org/10.1186/s13045-020-00918-x> (2020).
24. Jett, J. R. *et al.* Audit of the autoantibody test, EarlyCDT(R)-lung, in 1600 patients: an evaluation of its performance in routine clinical practice. *Lung Cancer* **83**, 51–55. <https://doi.org/10.1016/j.lungcan.2013.10.008> (2014).
25. Gonzalez Maldonado, S., Johnson, T., Motsch, E., Delorme, S. & Kaaks, R. Can autoantibody tests enhance lung cancer screening? An evaluation of EarlyCDT(R)-Lung in context of the German Lung Cancer Screening Intervention Trial (LUSI). *Transl. Lung Cancer Res.* **10**, 233–242. <https://doi.org/10.21037/tlcr-20-727> (2021).
26. Tang, Z. M., Ling, Z. G., Wang, C. M., Wu, Y. B. & Kong, J. L. Serum tumor-associated autoantibodies as diagnostic biomarkers for lung cancer: a systematic review and meta-analysis. *PLoS ONE* **12**, e0182117. <https://doi.org/10.1371/journal.pone.0182117> (2017).
27. Jiang, D. *et al.* Discovering panel of autoantibodies for early detection of lung cancer based on focused protein array. *Front. Immunol.* **12**, 658922. <https://doi.org/10.3389/fimmu.2021.658922> (2021).
28. Sun, G. *et al.* Identification of novel autoantibodies based on the protein chip encoded by cancer-driving genes in detection of esophageal squamous cell carcinoma. *Oncoimmunology* **9**, 1814515. <https://doi.org/10.1080/2162402X.2020.1814515> (2020).
29. Cui, C. *et al.* Identification of novel autoantibodies based on the human proteomic chips and evaluation of their performance in the detection of gastric cancer. *Front. Oncol.* **11**, 637871. <https://doi.org/10.3389/fonc.2021.637871> (2021).
30. Wang, S. *et al.* Using a panel of multiple tumor-associated antigens to enhance autoantibody detection for immunodiagnosis of gastric cancer. *Oncoimmunology* **7**, e1452582. <https://doi.org/10.1080/2162402X.2018.1452582> (2018).
31. Wang, P. *et al.* Using a panel of multiple tumor-associated antigens to enhance the autoantibody detection in the immunodiagnosis of ovarian cancer. *J. Cell Biochem.* **120**, 3091–3100. <https://doi.org/10.1002/jcb.27497> (2019).
32. Kobayashi, M. *et al.* Proteome profiling uncovers an autoimmune response signature that reflects ovarian cancer pathogenesis. *Cancers (Basel)* <https://doi.org/10.3390/cancers12020485> (2020).
33. Ma, Y. *et al.* Using protein microarray to identify and evaluate autoantibodies to tumor-associated antigens in ovarian cancer. *Cancer Sci.* **112**, 537–549. <https://doi.org/10.1111/cas.14732> (2021).
34. Katchman, B. A. *et al.* Autoantibody biomarkers for the detection of serous ovarian cancer. *Gynecol. Oncol.* **146**, 129–136. <https://doi.org/10.1016/j.ygyno.2017.04.005> (2017).
35. Wang, H. *et al.* Identification and validation of novel serum autoantibody biomarkers for early detection of colorectal cancer and advanced adenoma. *Front. Oncol.* **10**, 1081. <https://doi.org/10.3389/fonc.2020.01081> (2020).
36. Wang, J. *et al.* Plasma autoantibodies associated with basal-like breast cancers. *Cancer Epidemiol. Biomark. Prev.* **24**, 1332–1340. <https://doi.org/10.1158/1055-9965.EPI-15-0047> (2015).
37. Syed, P. *et al.* Autoantibody profiling of glioma serum samples to identify biomarkers using human proteome arrays. *Sci. Rep.* **5**, 13895. <https://doi.org/10.1038/srep13895> (2015).
38. Välikangas, T., Suomi, T. & Elo, L. L. A systematic evaluation of normalization methods in quantitative label-free proteomics. *Brief. Bioinform.* **19**, 1–11. <https://doi.org/10.1093/bib/bbw095> (2016).
39. Rechavi, E. *et al.* Timely and spatially regulated maturation of B and T cell repertoire during human fetal development. *Sci. Transl. Med.* <https://doi.org/10.1126/scitranslmed.aaa0072> (2015).
40. Eisenberg-Lerner, A., Regev, I. & Merbl, Y. Post-translational modification profiling—functional proteomics for the analysis of immune regulation. *Methods Mol. Biol. (Clifton, NJ)* **1647**, 139–152. doi:https://doi.org/10.1007/978-1-4939-7201-2_9 (2017).
41. Avrameas, S. Natural autoantibodies: from “horror autotoxicus” to “gnothi seauton”. *Immunol. Today* **12**, 154–159. [https://doi.org/10.1016/s0167-5699\(05\)80045-3](https://doi.org/10.1016/s0167-5699(05)80045-3) (1991).
42. Wardemann, H. *et al.* Predominant autoantibody production by early human B cell precursors. *Science* **301**, 1374–1377. <https://doi.org/10.1126/science.1086907> (2003).
43. Lobo, P. I. Role of natural autoantibodies and natural IgM anti-leucocyte autoantibodies in health and disease. *Front. Immunol.* <https://doi.org/10.3389/fimmu.2016.00198> (2016).
44. Brezina, S. *et al.* Immune-signatures for lung cancer diagnostics: evaluation of protein microarray data normalization strategies. *Microarrays (Basel)* **4**, 162–187. <https://doi.org/10.3390/microarrays4020162> (2015).

Author contributions

T.V.C., Y.A., K.L.K., F.W., A.Z.J., and Y.C. collected and/or maintained biospecimens and data on which the paper is based. S.H., T.H., S.L. designed the assays and collected the autoantibody data. T.V.C., Y.A., M.A., and M.L. oversaw the statistical analysis. T.V.C., S.H., K.L.K., Y.A., M.A., A.Z.J., and Y.C. wrote the manuscript. All authors verified the analysis, tables, figures, and interpretations of the data and read the manuscript.

Funding

The NYU Women’s Health Study is supported by funding from the NCI: U01CA182934 (mPI: Zeleniuch-Jacquotte and Chen).

Competing interests

S.H., T.H., and S.L. are employed by CDI laboratories, which created the HuProt™ platform for autoantibody assays used in this study. M.A. is a consultant for CDI Laboratories and the owner of Biomedical Hosting, LLC. T.V.C., Y.A., K.L.K., M.L., F.W., A.Z.J., and Y.C. declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-10174-3>.

Correspondence and requests for materials should be addressed to Y.C.

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

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



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

© The Author(s) 2022