an unbiased proteomics approach to identify candidate urine biomarkers (CUBMs) predictive of LN chronicity and further pursued their validation in a larger cohort.

Methods In this cross-sectional pilot study, we selected urine collected at kidney biopsy from 20 children with varying levels of LN damage (discovery cohort) and performed proteomic analysis using iTRAQ. We identified differentially excreted proteins based on degree of LN chronicity and sought to distinguish markers exhibiting different relative expression patterns using hierarchically-clustered log10-normalized relative abundance data with linked and distinct functions by biological network analyses. For each CUBM, we performed specific enzyme-linked immunosorbent assays (ELISAs) on urine from a validation cohort (n=41) and analysis of variance (ANOVA) to detect differences between LN chronicity, with LN activity adjustment. We evaluated for CUBM expression in LN biopsies with immunohistochemistry.

Results iTRAQ detected 112 proteins from urine samples in the discovery cohort, 51 of which were quantifiable in all replicates. Simple ANOVA revealed four differentially expressed, chronicity-correlated proteins (p-values<0.05). Further correlation and network analyses led us to select a total of seven CUBMs for LN chronicity: afamin (AFM), immunoglobulin heavy constant alpha 1 (IGHA1), alpha-1-antichymotrypsin (SERPINA3), transthyretin (TTR), retinol binding protein 4 (RBP4), alpha-1-acid glycoprotein, type 2 (ORM2) and transferrin (TF). In the validation cohort, urine SERPINA3 elevation was strikingly robust with respect to validation in high chronicity LN samples (3.3-fold change, p-value 0.012). Immunohistochemistry further demonstrated SERPINA3 staining in both endothelial and proximal tubular epithelial cells.

Conclusions Using advanced proteomic techniques followed by confirmation using specific ELISAs, we identified SERPINA3, a known inhibitor of neutrophil cathepsin G and angiotensin II production, as a potential urine biomarker to help quantify LN damage. SERPINA3 expression may be a protective mechanism from further kidney damage. Further validation of SERPINA3 as an LN damage biomarker in an independent cohort is needed to determine its ability to guide treatment and predict prognosis.

Acknowledgements This study was supported by grants from the National Institutes of Health [P50 DK 096418, U01 AR065098, T32 AR069512, P30 AR070549] and a Lupus Foundation Career Development Award to Dr. Turnier. Mass spectrometry data were collected on a system funded through an NIH shared instrumentation grant (S10 RR027015–01; KD Greis-PI).

BD-06 IDENTIFICATION OF SYSTEMIC LUPUS ERYTHEMATOSUS SUBGROUPS USING ELECTRONIC HEALTH RECORD AND GENETIC DATABASES

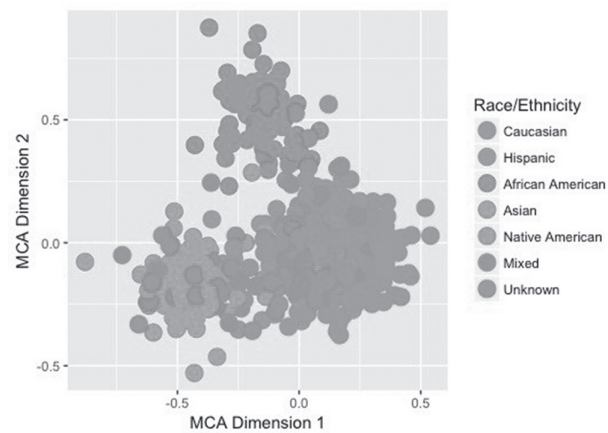
[1]Milena A Gianfrancesco*, [2]Ishan Paranjpe, [1]Julia Kay, [1]Joanne Nititham, [1]Kimberly Taylor, [1]Lindsey Criswell, [2]Marina Sirota, [1]Gabriela Schmajuk, [1]Jinoos Yazdany. *[1]Division of Rheumatology, Department of Medicine, University of California, San Francisco; [2]Institute for Computational Health Sciences, University of California, San Francisco*

10.1136/lupus-2018-lsm.30

Background Systemic lupus erythematosus (SLE) is a multifactorial autoimmune disease with genetic and environmental risk factors, as well as heterogeneous manifestations that encompass a wide range of disease severity. Long-term outcomes for individual patients are therefore difficult to predict, as is the scope of organ-system involvement. Little is known about why an affected individual might develop a particular SLE phenotype. A number of studies have used phenotype-mapping approaches to identify subgroups of SLE using genome-wide association studies and gene expression data; however, studies integrating the contribution of both genetic and clinical factors to identify SLE phenotypes in large, comprehensive data sources using bioinformatics analyses remain limited.

Methods We characterized subgroups of SLE patients using genetic, sociodemographic, and clinical variables from previously collected genetic cohorts and electronic health record (EHR) data for 712 individuals with SLE. Genetic data included 95 single nucleotide polymorphisms (SNPs) associated with SLE in the literature genotyped on the Immunochip platform. Variables extracted from the EHR included age, sex, race/ethnicity, age at disease onset, Charlson comorbidity index, and various disease-associated laboratory measures. Preliminary clustering was conducted using multiple correspondence analysis.

Results Approximately 90% of SLE patients represented in genetic and EHR databases were female, and 50% self-identified as Caucasian, 12% Hispanic, 10% African-American, 14% Asian, and 14% Other/Missing. Preliminary results showed distinct clustering by race/ethnicity amongst the 95 SNPs associated with SLE (figure 1). Further clustering and network approach analyses incorporating both genetic and clinical variables, such as laboratory measures, are ongoing and will explore whether distinct SLE subgroups can be identified.



Abstract BD-06 Figure 1 Ninety-five SNPs associated with SLE demonstrate clustering by race/ethnicity

Conclusion This project is a first step towards identifying subgroups of SLE patients through clinical and genetic databases. These findings will contribute to our understanding of SLE and illustrate how combining big data in both genetics and EHR has the potential to further define this heterogeneous disease.

Acknowledgements This work was supported by NIH/NIAMS (grant number F32 AR070585 to M.A.G. and K23 AR063770 to G.S.); AHRQ [grant number R01 HS024412 to J.Y.]; and PREMIER, a NIH/NIAMS P30 Center for the Advancement of Precision Medicine in Rheumatology at UCSF (AR040155). Drs. Yazdany and Schmajuk are also supported by the Russell/Engleman Medical Research Center for Arthritis. The content is solely the responsibility of the authors and does not necessarily represent the official views of the AHRQ or NIH.