Discovery of Novel Self-Antigens in IgG4-Related Disease Using Computational Modeling and Human Proteome Screening

Guha, Mahilan (School: Montgomery Blair High School)

Autoimmune diseases affect millions worldwide, yet many are difficult to diagnose due to biological complexity and variability across patients. I developed a computational pipeline to address autoimmune heterogeneity, enabling the discovery of novel self-antigens in IgG4-related disease (IgG4-RD)—an immune-mediated disorder characterized by tumor-like lesions, fibrosis, and organ dysfunction, and lacking a defined autoimmune signature. While prior research shows that IgG4-RD patients harbor plasmablasts that recognize self-antigens, no consistent disease-specific autoantibody response has been identified. I hypothesized that proteome-wide computational screening could uncover self-antigens that define and sub-cluster IgG4-RD patients. To test this, I analyzed HuProt microarray data previously collected from 30 patients with IgG4-RD and 30 patients with Systemic Sclerosis (SSc, positive control). Using R, I used three machine learning models—LASSO regression, linear and generalized linear mixed models (LMM/GLMM), and SLIDE—to identify differentially bound proteins. Each model performed with high prediction accuracy (AUC > 0.9), and their intersection revealed seven consistently reactive autoantigens. Two, POLR3K and TOP1MT, were known targets in SSc. DDX3Y and LIMS1 emerged as strong candidates for IgG4-RD-specific autoantigens, with literature suggesting roles in immune regulation and IgG subclass expression. Gene Ontology analysis confirmed biological relevance, highlighting functions like RNA Polymerase III activity. This work advances our understanding of IgG4-RD pathophysiology and establishes a scalable framework for autoantigen discovery. It opens the door to personalized diagnostics in immunology by enabling high-resolution profiling of autoreactivity.

Awards Won: