Matthew Joel

Title: Multi-Omics Analysis of Fibrotic ILA Patients Using the DIABLO Framework: Insights from DNA Methylation, Transcriptomic, and Proteomic Data

Authors: Matthew Joel, Rachel Blumhagen, David Schwartz, Ivana Yang - University of Colorado Anschutz Medical Campus

Abstract:

Interstitial lung abnormalities (ILA) represent a spectrum of fibrotic changes, but the molecular basis remains poorly understood. Our study aims to identify multi-omics signatures in fibrotic ILA by integrating DNA methylation, transcriptomic, and proteomic data from 106 individuals (23 cases, 83 controls) in a University of Colorado ILA cohort.

We applied DIABLO (Data Integration Analysis for Biomarker discovery using Latent variable approaches for Omics) to explore correlations across DNA methylation, transcriptomic, and proteomic data blocks. Two models were tested—one with two latent components and one with four—guided by an elbow plot of classification error. We set the design matrix weight to 0.5, a threshold indicating an equal emphasis on correlation (> 0.5) versus prediction (< 0.5). Optimal variable selection was determined through parameter tuning, and 10 M-fold cross-validation assessed stability and predictive performance. Graphical outputs included circos plots to depict correlations among blocks, loading plots to highlight contributions, and ROC curves to evaluate classification performance.

Preliminary findings revealed moderate correlations between the three data types, but only modest classification accuracy. For instance, in the two-component model, the second component demonstrated a Pearson correlation of 0.64 between the RNA sequencing block and the proteomic block. The modest classification accuracy may stem from setting the design matrix value at 0.5, thereby balancing correlation and prediction rather than optimizing for one objective. Nevertheless, several candidate biomarkers were identified, warranting further validation.

These early results underscore the potential of integrative methods to reveal cross-omics relationships in fibrotic ILA, although predictive power was limited in our initial approach. Ongoing work will explore alternative data transformations, refined filtering criteria, different component numbers, and design matrix adjustments aimed at enhancing predictability. More extensive tuning, including additional variables and extended parameter ranges, may further clarify the molecular drivers of fibrotic ILA and improve predictive accuracy.