Biologically-informed Machine Learning for Modeling Spatiotemporal Dynamics of Cell States and Interactions in Acute Myeloid Leukemia

Shouvik Mani
Department of Computer Science
Columbia University

Abstract

To realize personalized medicine, we must appreciate that every patient is unique and develop computational methods to characterize biological differences across patients, time, and disease subtypes along with novel therapeutic strategies informed by these differences. This is particularly true in cancer immunotherapy, a modern therapeutic paradigm which has revolutionized the treatment of many cancer types, but has nonetheless led to durable responses in only a subset of patients. For instance, in the treatment of relapsed acute myeloid leukemia (AML) following stem cell transplantation, the established immunotherapy known as donor lymphocyte infusion (DLI) achieves responses in only 15-20% of patients. Why do some AML patients respond to DLI while most develop resistance? Which immune cell states and interactions form the basis of DLI response, and how do they evolve over time and treatment? To address these questions, we built DIISCO, a Bayesian machine learning framework to characterize the temporal dynamics of cell states and interactions using single-cell RNA-sequencing data. Given proportions of cell types in patient samples over time, DIISCO infers interactions between cell types based on their co-evolution and incorporates prior knowledge on receptor-ligand complexes driving interactions. In our collaboration with Dana-Farber Cancer Institute, we applied DIISCO to bone marrow samples collected longitudinally from relapsed AML patients receiving DLI. DIISCO unraveled the dynamics of the AML tumor-immune microenvironment, suggesting a coordinated inhibition of leukemia cells by expanding subpopulations of CD8+ cytotoxic T cells and naive B cells unique to responder patients post-DLI. In addition to finding differential cell states and interactions, DIISCO identified receptor-ligand genes and downstream transcriptional programs potentially mediating those interactions. By investigating differences in DIISCO's inferred cell-cell interaction networks along the axes of responders vs non-responders (patients), pre-DLI vs post-DLI (time), and leukemia subtypes (disease), we can elucidate the mechanisms underlying effective anti-leukemia responses and guide the development of next-generation immunotherapies.