Using Machine Learning to Repurpose FDA-Approved Drugs to Treat Cancers and Inflammatory Diseases

Vermani, Aakarsh (School: Westview High School)

P38 alpha (MAPK 14) is a protein kinase implicated in the pathological mechanisms of myofibrillar myopathy, cancers, and inflammatory diseases like Alzheimer's and rheumatoid arthritis. Although p38 inhibitors have shown promise as treatments, traditional drug discovery methods are slow, costly, and have failed to create effective and safe inhibitors to treat these diseases. I addressed these shortcomings by developing a computational framework that combines deep learning with structure-based virtual screening to identify potential p38 blockers from FDA-approved drugs. I hypothesized that the predicted inhibitors would show a significantly higher binding affinity for p38 than a control group of random FDA-approved drugs. The best known p38 inhibitors were determined based on available bioactivity data, and fingerprint clustering was applied to isolate the compounds with similar structures. Descriptors were calculated for these selected compounds and the most important descriptors were isolated through machine learning-based feature selection. These descriptors were the training data for a deep neural network that was applied to 2,151 FDA-approved drugs. My deep learning model demonstrated a 92% validation accuracy and 0.97 AUROC, predicting 149 potential p38 inhibitors whose efficacies were confirmed by ligand-docking simulations to be significantly higher than the control group and slightly higher than even the known inhibitors, confirming the hypothesis. This project not only reveals potential treatments for p38-mediated diseases but also demonstrates the capability of this computational pipeline to predict novel functions of existing pharmaceuticals. In the future, experimental trials must be performed to confirm the efficacy of the predicted inhibitors.

Awards Won:
Third Award of $1,000
University of Arizona: Renewal Tuition Scholarship