

Integration and Dimensionality Reduction of CMAP and ChEA

Fred Clark Jr.^{1,3}, Lawrence Sirovich^{1,2}, Avi Ma'ayan^{1,2}

¹*Systems Biology Center New York (SBCNY)*, ²*Department of Pharmacology and Systems Therapeutics, Mount Sinai School of Medicine, New York, NY USA*; ³*Electrical Engineering Department, Fairleigh Dickinson University, Teaneck, NJ USA*

For this project we attempt to integrate and analyze two datasets: CMAP and ChEA. The Connectivity MAP (CMAP) project uses microarrays to profile the genome-wide expression changes effects of drugs on human cancer cell lines. The effect of ~6000 drug perturbations using ~1300 different drugs on gene expression is recorded in the CMAP database. Correspondingly, the ChIP Enrichment Analysis (ChEA) project is a collection of ChIP experiments profiling transcription factors and their putatively identified gene targets. In this project we utilize singular value decomposition (SVD) to assess the dimensionality of each dataset and examine how these two dataset correlate. Reconciling these resources using dimensionality reduction analysis will allow us to discover specific molecular regulatory mechanisms responsible for drug induced changes in gene expression in human cells as well as design drug combinations that could be used to control specific transcription factor activity.