

Figure 1. Elbow plot generated from CPTAC protein data. Average dispersion is on the y-axis and “K” is on the x-axis – this plot was used to determine that 3 clusters (K = 3) is the optimal choice of clusters for this set of data.

Chart, scatter chart

Description automatically generated

Figure 2. K-means plot generated from CPTAC protein data. This plot examines the clustering of patients when comparing their upregulation/downregulation of A2M and A1BG proteins.

Chart, scatter chart

Description automatically generated

Figure 3. PCA plot generated from TCGA patient/RNA data. Patients were colored by what stage of cancer they were diagnosed with, and clustered using the PCA method.

Chart, scatter chart

Description automatically generated

Figure 4. UMAP plot generated from TCGA RNA data. 3 clusters were used to group patients, as this was determined to be optimal via an elbow plot (not shown). The running k-means from Figures 1 & 2 were used to create this UMAP plot.

Chart, scatter chart

Description automatically generated

Figure 5. UMAP plot generated from the TCGA RNA counts data for A1M and A1BG RNA. 2 clusters were used in this plot to group patients as their A1M and A1BG levels were observed and plotted.