

AI/ML AND VIRTUAL HUMAN PLATFORMS FOR THREAT AGENT HAZARD ASSESSMENT AND MEDICAL COUNTERMEASURE DISCOVERY AND DRUG DEVELOPMENT

Using Machine Learning To Establish A Baseline For Detection Of Epigenetic Infection Signatures

Karissa Sanbonmatsu Los Alamos National Laboratory **Nicholas Lubbers** Los Alamos National Laboratory **Christina Steadman** Los Alamos National Laboratory **Nicholas Lubbers** Los Alamos National Laboratory

Epigenetics has emerged as a critical regulatory mechanism in mammalian systems, playing important roles in cancer, development, brain function and immune response. Here, chromosomes are chemically modified, leading to silencing or activation of specific genes in specific tissues at specific times. These modifications are often triggered by changes in the environment, which may take the form of chemical agents, pathogens, or social interactions. Interestingly, the epigenetic modifications can also be highly specific, to tissue type or to the individual. In many cases, epigenetic changes can be permanent, creating an effective record of environments experienced by the individual. Because of this high degree of specificity and sensitivity to the environment, epigenetic signatures can, in principle, be used to (i) detect exposures to chemical or biological agents, and (ii) predict whether or not a particular individual is more susceptible to a certain pathogen. We apply hierarchical clustering techniques to analyze 1500 samples of epigenetic data in human chromosome 6, using data derived from the Encode database. Using dendrogram and t-distributed stochastic neighbor embedding (t-SNE) visualizations, we find the data to cluster first by epigenetic mark and then by cell type. Genes that dominate clusters tend to be related to immune response or to epigenetic machinery. The study has established a baseline to detect epigenetic signatures of infections.

We are grateful to DTRA for their generous support.