Machine-learning methods for finding HIV epitopes
David Heckerman

I will discuss machine-learning methods for identifying HIV epitopes.

Current thinking among HIV-vaccine researchers is that HIV vaccines must prime the cellular immune response rather than or in addition to the humoral response in order to be successful. At the core of the cellular response is the ability of cells to ingest and digest viral proteins into smaller peptides, and then to present these peptides, known as epitopes, at the surface of the cell. This process is mediated by HLA (Human Leukocyte Antigen) molecules which form a complex with the epitope as it is presented. The epitope/HLA complexes can then be recognized by a T-cell, thereby activating the T-cell to subsequently recognize and kill virally infected cells. Each host has a different set of HLA alleles and presents a different set of epitopes.

A good vaccine will likely include a robust set of epitopes--robust in the sense of broad coverage and of covering regions that are essential for viral fitness in a given population characterized by a particular distribution of HLA alleles.

An important task in vaccine design is therefore the identification of epitopes and the HLA molecules that present them. Unfortunately, lab experiments required to prove that a peptide is an epitope for a particular HLA molecule are time-consuming and expensive. Consequently, alternative methods for identifying epitopes (including those that I will discuss) can be of tremendous help in vaccine design.