



Friday, March 7, 2008
2:00 pm – 3:00 pm
COOK 3118 A&B

An Investigation of the Oncomine User Interface as a Model for Creating New Microarray Data Mining Tools

Cheryl Lee – Masters Student

ABSTRACT:

Extremely large amounts of microarray data have been produced and accumulated. The databases of the two largest repositories for microarray data have grown exponentially each year. As of October 2007, 173,486 arrays have been deposited into the NCBI GEO database and 97,635 arrays deposited into the ArrayExpress (EMBL-EBI). This data will gain additional value using meta-analysis, combining the results of several different studies which address related research hypotheses. A user-friendly web-based interface is an essential part of a data mining tool to help users effectively explore thousands of experiments and tens of millions of gene expression profiles stored in databases. The study will investigate the user interface of Oncomine, in terms of the principles that may be considered during the design of a more user-friendly interface.

Research Advisor:

Simon Lin, Feinberg School of Medicine

Motif Discovery in DNase Hypersensitive Sites

Jairav Desai – Masters Student

ABSTRACT:

Cystic fibrosis (CF) is a hereditary disease of complex pathology that affects mainly the lungs and digestive system and results in progressive, life shortening disability. The cystic fibrosis transmembrane conductance regulator gene (CFTR), when mutated, causes cystic fibrosis and is thus has been highly investigated. The gene shows a carefully regulated pattern of expression; the molecular basis for which is not understood. It is clear, however, that important *cis* regulatory elements are not just restricted to the gene's promoter region. The Harris lab has identified regulatory elements relevant to the CFTR gene and mapped the function of these elements using comparative genomics; in particular data obtained from the ENCODE (ENcyclopedia of DNA Elements) database was used to identify potential DNA regulatory regions in CFTR using data from multiple experiments in a number of different cell lines. One particularly powerful line of investigation, DNase CHIP, identifies DNase Hypersensitive Sites (DHS) in DNA, which indicate nucleosome-depleted regions of chromatin and thus potential sites for the binding of regulatory elements. DHS CHIP data using DNA regions of the CFTR gene identified from the ENCODE database will be used to elucidate motifs for potential transcription factor binding sites and *de novo* motifs within DHS.

Research Advisor:

Ann Harris, Department of Pediatrics