The Departments of Epidemiology and Biostatistics and Genetics and Genome Sciences
Faculty Candidate Lecture

John Capra, PhD

Assistant Professor
Department of Biomedical Informatics
Vanderbilt University Medical Center
www.capralab.org

"Comparative and Functional Genomics of Human-specific Gene Regulation"

Thursday, September 12, 2013
Wood Building – WG-73
12:00pm-1:00pm
Abstract is on the next page

Additional Information: jean.farah@case.edu

SCHOOL OF MEDICINE
CASE WESTERN RESERVE UNIVERSITY

Abstract: Many of the dramatic differences in form and function between humans and our closest living relatives, the chimpanzees, are likely the result of changes to noncoding DNA that modify patterns of gene expression. In this talk, I will describe our efforts to identify noncoding regions of the human genome that have different regulatory functions in humans than in chimps. As a first step toward this goal, we developed a machine learning algorithm to identify genomic regions that are enhancers of gene expression. The algorithm integrates the evolutionary history, DNA sequence patterns, histone modifications, chromatin state, and transcription factor (TF) binding sites for a region, and it performs very well in cross validation using known developmental enhancers. We applied the algorithm to the entire human genome to generate a set of genome-wide tissue-specific enhancer predictions. Next, we identified hundreds of potential human-specific enhancers from the genome-wide set by analyzing evolutionary patterns and TF binding site divergence between human and chimp in the predicted enhancers. We then compared the enhancer activity of the human and chimp sequences of 29 of the predicted enhancers in transgenic mice. More than 80% (24/29) of these regions proved to be enhancers at the developmental stage tested (E11.5), and several exhibited consistent human-specific enhancer activity. These regions are exciting candidates for understanding the evolution of human-specific traits. This work lays a foundation for finding generic approaches to modeling the functional effects of noncoding sequence variation within and between species.