DEPARTMENT OF MATHEMATICS COLLOQUIUM SERIES

FRESNO STATE
Discovery. Diversity. Distinction.

The Department of Mathematics
Presents

Mario Banuelos
Ph.D. Candidate

Department of Applied Mathematics
University of California, Merced

Friday, September 15, 2017
@ 12:00PM in PB 103

Statistical Modeling and Inference in Genomic Variants

Abstract:
The genome is the complete DNA sequence associated with an organism. Every time a DNA sequence duplicates, there is an opportunity for variation, such as insertions and deletions of short segments of DNA. The introduction and accumulation of genomic variation are important evolutionary drivers of phenomena such as the creation of new species as well as genetic diseases like cancer. Traditionally, biology has focused on genomic variation between species, but there is growing appreciation that there is substantial variation between individuals in the same species. In my research, I develop mathematical models to analyze variation both between species and within members of the same species. In this talk, I will discuss what we can learn from genomic variation in each of these contexts. First, I developed a mathematical model of the evolution of transposable elements, a major driver of genomic variation. In contrast to most approaches, I consider the complete transposable element annotation of a genome. Through comparing present day genomes of multiple related species I am able to uncover quantitative relationships of their proliferating dynamics and assess the evolutionary history of these genomic insertions.

Second, to predict genomic variation between members of the same species, I developed a constrained-optimization method to detect germline structural variants in family lineages. Unlike annotated or highly curated data, sequenced genomes within species tends to suffer from errors in both sequencing and mapping. Whereas previous methods either post-process hereditary information or rely on high-quality sequencing, I incorporate data from multiple related individuals to reduce the false positive rate of prediction. Since zygosity - the number of variant copies present in an individual - affect disease susceptibility and resistance, my method also incorporates this into variant prediction. In the context of simulated genomes, parent-child trios, and larger family lineages, my relatedness and sparsity-promoting model results in improved detection of genomic variation responsible for both diversity and disease.

If you need a disability-related accommodation or wheelchair access information, please contact Steve Chung at (559)278-2462 or e-mail schung@csufresno.edu.