1145-AA-2567 **Mario Banuelos*** (mbanuelos22@csufresno.edu), **Suzanne Sindi** and **Roummel F Marcia**.
Mathematical -Omics Models in Error-Prone Data Regimes.

Every time a cell divides, its entire DNA sequence is duplicated, and there is an opportunity for the introduction of errors, such as insertions and deletions. The introduction and accumulation of genomic variation are important evolutionary drivers of phenomena such as the creation of new species as well as the development of genetic diseases like cancer. As genome sequencing costs decrease, the volume of sequencing data has resulted in the need for advanced mathematical and computational methods. Unlike highly annotated reference genomes, sequenced genomes from large public repositories tend to suffer from errors in both sequencing and mapping.

In this talk, I will discuss mathematical models in noisy -omic data regimes used to represent and predict genomic variation in organisms. To model the proliferation of variants, we use the full spectrum of repetitive elements and develop a fragmentation equation model which describes non-actively replicating repetitive elements in an organism's genome. Additionally, we predict genomic variation between members of the same species by developing a constrained-optimization framework using gradient-based methods and constrain our solution with a sparsity-promoting ℓ_1 penalty to detect structural variants (SVs) in family lineages. (Received September 25, 2018)