1133-05-164 **Qijun He*** (qijun@bi.vt.edu), **Christopher Barrett**, **Fenix Huang** and **Christian Reidys**.

From genotypes to phenotypes and back: towards an information theory of DNA.

In this talk we study an information-theoretic framework rooted in the pairing of genotypes and phenotypes. We view the correlation of these as structural semantics of sequence data that allows for a different interpretation than conventional sequence alignment. This structural semantics could enable us to identify and interpret novel, embedded 'patterns' in DNA and RNA sequences. We compute the partition function of RNA sequences with respect to a fixed RNA secondary structure and connect this computation to a concept of mutual information of a sequence–structure pair for RNA secondary structures. We present a Boltzmann sampler and obtain the a priori probability of specific sequence patterns. We present a detailed analysis for several PDB-structures. We localize specific sequence patterns, contrast the energy spectrum of the Boltzmann sampled sequences versus those sequences that refold into the same structure and derive a criterion to identify native structures. We present multiple sequences in the partition function of a fixed structure, each having nearly the same mutual information, that are nevertheless poorly aligned. This indicates the possibility of the existence of relevant patterns embedded in the sequences that are not identified using alignments. (Received July 24, 2017)