Time & Location: July 1, 2015 at 12:00 PM in HEC 356
Title: Finding Consensus Energy Folding Landscapes Between Multiple RNA Sequences

In molecular biology, the secondary structure of a ribonucleic acid (RNA) molecule is closely related to its biological function. One problem in structural bioinformatics is to determine the two- and three-dimensional structure of RNA using only sequencing information, which can be obtained at low cost. This entails designing sophisticated algorithms to simulate the process of RNA folding using detailed sets of thermodynamic parameters.

The set of all chemically feasible structures an RNA molecule can assume, as well as the energy associated with each structure, is called its energy folding landscape. This research focuses on defining and solving the problem of finding the consensus landscape between multiple RNA molecules. Specifically, we discuss how this problem is equivalent to the problem of Balanced Global Network Alignment, and what effect a solution to this problem would have on our understanding of RNA.

Because this problem is known to be NP-hard, we instead define an approximate consensus on a landscape of reduced size, which dramatically reduces the searching space associated with the problem. We use the program RNASLOpt to enumerate all stable local optimal secondary structures in multiple landscapes within a certain energy and stability range of the minimum free energy (MFE) structure. We then encode these using an extended structural alphabet and perform sequence alignment using a structural substitution matrix to find and rank the best matches between the sets based on stability, energy, and structural distance. We apply this method to twenty landscapes from four sets of riboswitches from Bacillus subtilis in order to predict their native “on” and “off” structures. We find that this method significantly reduces the size of the list of candidate structures, as well as increasing the ranking of previously obscure secondary structures, resulting in more accurate predictions overall. Advances in the field of structural bioinformatics can help elucidate the underlying mechanisms of many genetic diseases.

Major: Computer Engineering

Educational Career:
Bachelor's of Computer Engineering, BS, 2013, University of Central Florida

Committee in Charge:
Shaojie Zhang, Chair, EECS
Haiyan Hu, EECS
Sumit Jha, EECS

Approved for distribution by Shaojie Zhang, Committee Chair, on June 15, 2015.

The public is welcome to attend.