Reconstructing hybridization networks
Elizabeth Allman, Univeristy of Alaska Fairbanks

An important goal in evolutionary biology is to understand historical relationships between extant species. Traditionally, a tree has been used to relate species so that, for example, the tree

(((Human,Chimp),Gorilla),Orangutan)

expresses that Human and Chimpanzees are the most closely related among the four great ape species, and that the orangutan species split at some distant time in the past from some ancestral species to modern day humans, chimps and gorillas.

However, because of a number of biological processes including, for example, gene duplication, incomplete lineage sorting, and hybridization, it is well know that species-level relationships may be best described with a network, rather than a tree. Molecular sequences for a large number of genes sampled from individuals in species are now routinely collected, and gene trees are constructed from these data. Because of processes like hybridization, the topologies of these gene trees exhibit discord (they do not agree with one another), yet such `multi-locus' datasets are often used as input into a method for constructing a species network (or tree).

In this talk, we discuss a new statistically consistent method to reconstruct a topological species network from a large sample of gene trees. The network estimator is consistent under the Network Multispecies Coalescent model and is based on the assumption that the discordance in gene trees arises from incomplete lineage sorting. We also discuss a hypothesis test for tree-like evolution.

This talk will not be technical in nature and will begin with an overview of the species tree/network inference problem; the level of the talk is appropriate for graduate students.

This is joint work with Hector Banos, Jonathan Mitchell, and John Rhodes.