

Title: Estimating a 'tree of life' from genomes; approaches and pitfalls

Speaker: Andrew J. Roger

Abstract

Determining the tree of life remains one of the major challenges for modern evolutionary biology. From comparisons of the sequences of genes from all major life forms on Earth over the last thirty years, we now know that life can be divided into three basic kinds of cells - complex cells that house their genome in a membrane-bounded nucleus (Eukaryotes like humans, plants and single-celled organisms called protists) and simpler single-celled organisms that do not (Eubacteria, e.g. *E. coli*, and Archaeobacteria like the salt-loving halobacteria or hyperthermophiles living in hydrothermal vents). Despite the delineation of these three major groups, the root of the tree of life and the branching order of lineages within these groups remain controversial. I will briefly review how stochastic models of gene and protein evolution are used in molecular phylogenetic estimation of the deep branches in the tree of life. I will discuss how difficulties in modeling protein evolution over deep time scales and biological processes like 'lateral gene transfer' make estimation of branching events in the tree of life that occurred between 1-4.5 billion years ago extremely difficult, if not impossible.