

## Constructing the Microbial Tree of Life Using Quartet Based Approaches

Sagi Snir, Haifa University

Understanding the origin and history of life on Earth is among the most fundamental and challenging tasks in Biology. This history is usually described by a tree-like relationship. In particular, the advent of Next Generation Sequencing techniques has opened up the opportunity to combine a host of specific gene histories into the unifying big tree of life (ToL) that portrays the main trend of evolution of life.

In microbes however, this trend is largely obscured by massive activity of horizontal gene transfer (HGT) that links distantly related lineages along the ToL. Nevertheless a common hypothesis is that microbial evolution is primarily tree-like, and a routine effort is made to distil this tree-like signal from the conflicting evidence.

In a recent result, we have shown that it is possible to construct a quartet oracle such that the quartet topology of every four species is WHP correct. The next task is combining all these, possibly conflicting, quartet trees into the big ToL, a task denoted as the supertree.

This problem lies at the root of many tree reconstruction methods and theoretical as well as experimental results have been reported.

In a series of works we have developed a probabilistic, graph theoretically based, approach for the supertree task. Our approach is based on a divide and conquer algorithm where our divide step uses a semi-definite programming (SDP) formulation of MaxCut in a graph representing relationships between the organisms. We devised an extremely fast SDP-like heuristic that allows us to extend the input data from several thousands of quartet trees over few dozens of species to tens of millions of quartet trees over thousands of species. Using these tools we arrived at surprising results regarding the microbial tree of life.

The talk is self contained and assumes no prior knowledge in neither Biology nor algorithms.