Phylogenetically-Informed Distance Methods for Microbiome Data Analysis

Phylogenetically-informed distances are widely used in ecology, often in conjunction with multi-dimensional scaling, to describe the relationships between communities of organisms and the taxa they comprise. A large number of such distances have been developed, each leading to a different representation of the communities. The ecology literature often tries to interpret the differences between representations given by different distances, but without a good understanding of the properties of the distances it is unclear how useful these interpretations are. I give an overview of some of these distances, describe the interpretational challenges they pose, develop some interesting properties, and comment on opportunities for improvement.