

PMI 3-way Networks: Application of Hypergraphs for Modelling Increased Complexity in Comparative Genomics

Contact: Dan Jacobson (jacobsonda@ornl.gov, 865-574-8212)

Funding Source: DOE Office of Biological and Environmental Research, Genomic Science Program and BioEnergy Science Center

Background

- In order to address the underlying complexity of phylogenomic relationships between organisms, we have developed a new 3-way similarity metric and constructed 3-way networks modelling the relationships between 211 bacterial genomes.

Science

- We used 3-way networks to model the relationships between triplets of objects instead of pairs of objects and used a hypergraph, which is a generalized network, in which an edge can model the relationship between an arbitrary number of objects.
- In order to quantify ternary relationships, a 3-way Sørensen Index and 3-way Czekanowski Index was used to quantify the similarity between all triplets of bacterial species, based on their gene family content. This allowed for similarities between more than two species to contribute to the interpretation.

Significance

- We have developed a new 3-way metric for the weighting of 3-way edges. This is the first time that the concept of 3-way networks has been applied in the field of comparative genomics.
- These networks will be a useful tool to model and reveal complex relationships between bacterial species, which are not found using the conventional 2-way network models.

Subnetworks containing the *Clostridium* and *Bacillus* species selected from: **A:** 3-way Sørensen Network; **B:** 3-way Czekanowski Network; **C:** Gene family enrichment network

