Analyzing large biological datasets with association networks

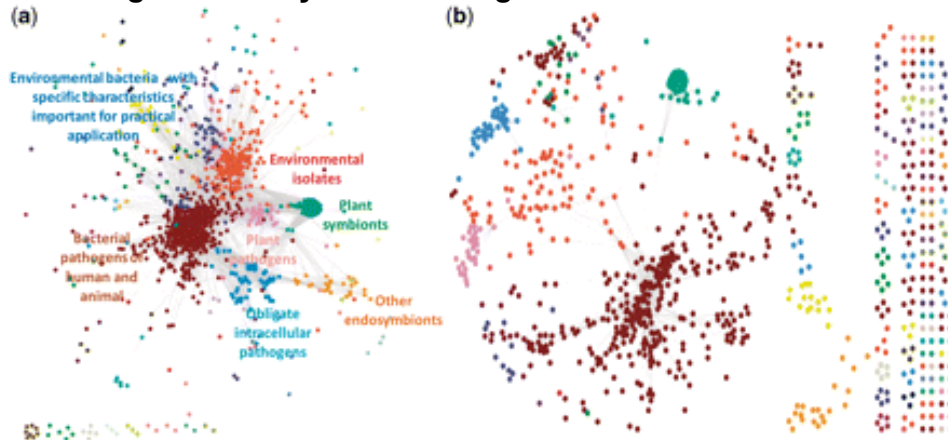*Contact:* Tatiana Karpinets (865)241-8323, karpinetstv@ornl.gov

- Due to advances in high-throughput biotechnologies, biological studies accumulate information at an amazing rate and require novel algorithms to convert the collected data into biological knowledge.

- We have developed a computational framework that utilizes classical association rule learning (Arule) and a novel approach to identify indirect associations and hidden biological regularities within a large data set.

- The framework is based on a new format for presentation of biological datasets and a novel algorithm, referred as Association NETwork or ANET, and allows discovery of significant relationships in the collected data.

- The application domain of the proposed approach is not limited to biological data. It can be applied to approximate the meaning of texts documents, to analyze social communities, to visualize results of surveys and to facilitate clustering of densely nested weighted networks.
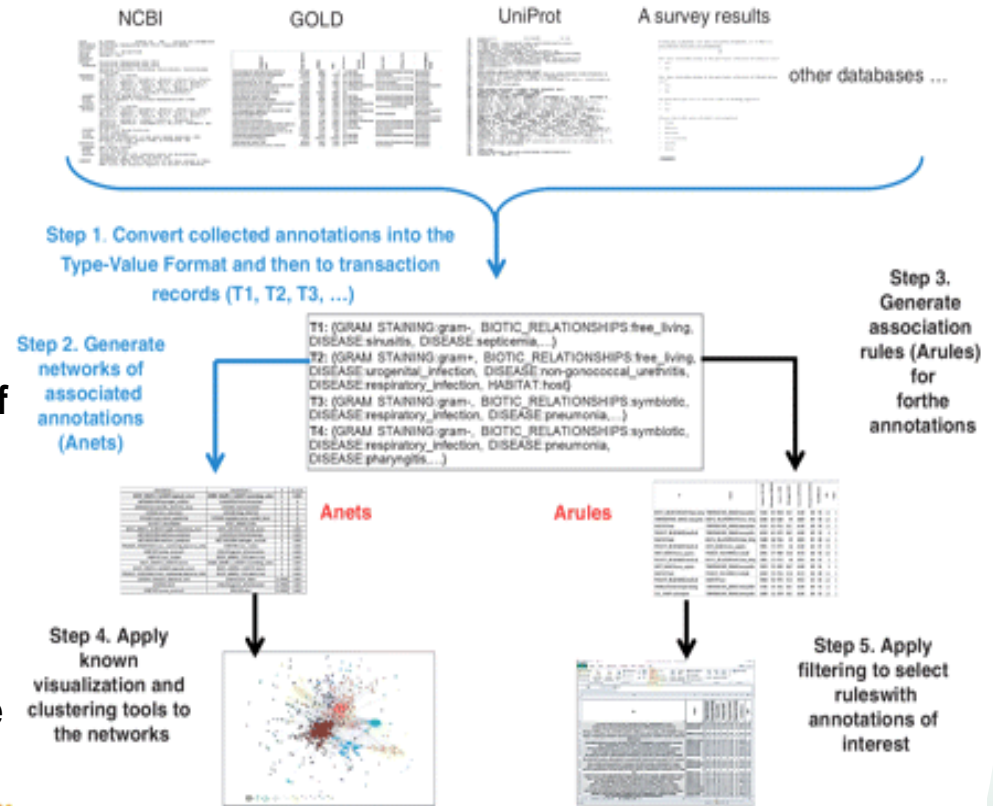
Computational framework for analysis of annotations collected in biological databases

Biological maps of sequenced prokaryotic organisms based on their metadata collected in the GOLD

Karpinets TV, Park BH, Uberbacher EC. Analyzing large biological datasets with association networks. Nucleic Acids Res. 2012, 1-8. doi:10.1093/nar/gks403

OAK RIDGE National Laboratory