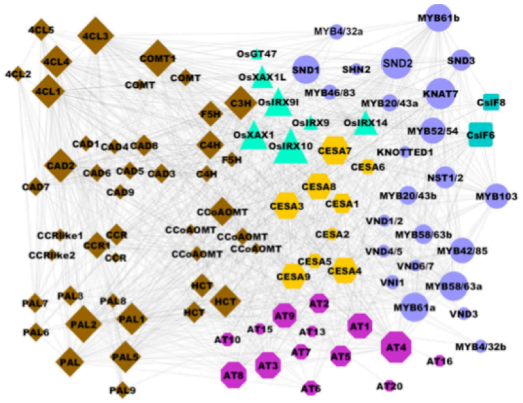
Introduction to biological gene networks and application to improving biofuel production.

Laura Bartley (lbartley@ou.edu), OU Department of Microbiology and Plant Biology
Kangmei Zhao, OU Department of Microbiology and Plant Biology
Soumen Roy, Bose Institute, Kolkata, India
Rajdeep Kaur Grewal, Bose Institute, Kolkata, India



Gene networks are often used to predict gene product biological function. However, since many gene networks rely simply on correlations in biological datasets, they are notoriously error prone and work is needed to improve the accuracy of hypotheses developed with networks. The example of biological networks I will discuss seeks hypotheses for genes that function in a particular biological process—cell wall synthesis by grasses. Grass cell wall material, including the residues of cereal crops, are an abundant source of chemical energy from which to make biofuels and results may lead to improvements in biofuel production efficiency. Furthermore, this study provides an opportunity to reveal biologically relevant features of gene networks. The rice combined inverse ranked (RCR) network uses a general linear model to build a relatively complete network of equal or greater quality to the component networks from which it was assembled. Examining the nodes connected to known cell wall synthesis genes (i.e., the guilt by association logic) reveals that genes that are unique to grass cell wall synthesis have been incorporated into the evolutionarily conserved cell wall subnetwork. The RCR network appears to have a similar true positive rate and a lower false negative rate than the best previous rice network. In addition, analysis of seven topological features of the top scored 50K edges of various rice networks identifies closeness, a measure of the accessibility among nodes, as a distinguishing feature of the RCR. This raises the possibility that close topology can be used to identify high-accuracy gene networks.