Project: Integration of metabolic network into interaction network for microbial communities

Name: Zhifeng Yang

Date: 10/16/2020

Introduction:

Network has been developed as a common method in microbial ecology nowadays. Metabolic network, which describes how metabolites are utilized and transformed inside or around cells, is the key to interpreting the unique functions and traits of each species. Interaction network, which describes the positive or negative interaction between paired species, determines the dynamics of microbial communities. Basically, what determined the interaction of microbes is the traits of microbes. For example, the negative interaction may result from competition for shared food sources of two species, which could share some nodes pointing to the main metabolic network. On the contrary, the positive interaction may result from mutualism of two species, one of which may provide food to another. So metabolic network can somehow imply the interaction between species.

Ideally, if we know all metabolic networks of a microbial community, which may have over thousands of species, we can build an interaction network based on edges between two metabolic networks of paired species. However, the real interaction network may be much different due to a lot of issues. For example, the metabolic network doesn't involve all microbial processes such as reproduction and death, which may influence the interaction. Besides, the interaction network generated may not reflect the real interaction, which is hard to be validated in experiment. In this project, I want to study how metabolic network can imply interactions of species and valid this with interaction network.

The work is based on the paper published in 2013 (Levy and Borenstein 2013). The paper test if the metabolic-network-informed competition or complementarity could explain the co- occurrence of species. The figure below showed how metabolic network informed competition or complementarity.

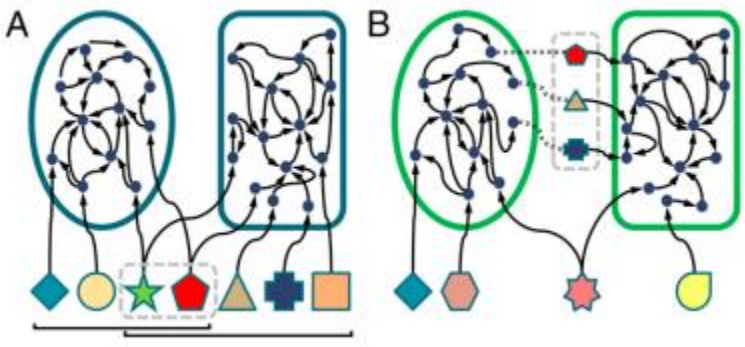

Fig. The big circle and rectangle represents metabolic network while small icons represent metabolites which are also the nodes in network. The edges represent the pathway. The A graph shows the competition for food (the green and red icons), while B graph shows complementarity because the left cell provides some metabolites to right cell.

Method and materials:

For my project, I will do the following work:

1. Simplify the individual network based on the method in reference paper (Borenstein, Kupiec et al. 2008).
2. Produce a whole network of metabolism including all species (directed)
3. map the network of each species into the whole metabolic network
4. if two species have a lot of edges; for example, if one point to another a lot, one may be benefit to another, so there is an edge pointed to another from one. If one shares a lot of nodes point to them, they may compete for the resource of nodes, there should be a negative edge between them
5. we can get a metabolic relation network of species.
6. build a microbial interaction network using abundance correlation
7. test if metabolic network has some prediction of the structure of microbial interaction network.

Data source:

1. I will use the microbial data set of my master's thesis, which includes relative abundance of microbes in front of glaciers. The abundance information of the data set is used for building the interaction network.

2. To simplify the building a metabolic network, I will use the predicted function from the taxonomy information of the data set using PICRUST 2 (Caicedo, Hashimoto et al. 2020) and reference network in KEGG (Kanehisa, Araki et al. 2007) rather than annotation from a lot of draft genomes.

Goal of this project:

The work is challenging due to the complexity of metabolic network. I guess I may finish the step 1 to step 4. For each step, I will present a detailed method and result during the process.

References:

Borenstein, E., et al. (2008). "Large-scale reconstruction and phylogenetic analysis of metabolic environments." Proceedings of the National Academy of Sciences **105**(38): 14482-14487.

Caicedo, H. H., et al. (2020). "Overcoming barriers to early disease intervention." Nat Biotechnol **38**(6): 669-673.

Kanehisa, M., et al. (2007). "KEGG for linking genomes to life and the environment." Nucleic acids research **36**(suppl_1): D480-D484.

Levy, R. and E. Borenstein (2013). "Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules." Proc Natl Acad Sci U S A **110**(31): 12804-12809.