Dynamic model helps understand healthy lakes to heal sick ones

April 7 2016

Development of a dynamic model for microbial populations in healthy lakes could help scientists understand what's wrong with sick lakes, prescribe cures and predict what may happen as environmental conditions change. Those are among the benefits expected from an ambitious project to model the interactions of some 18,000 species in a well-studied Wisconsin lake.

The research produced what is believed to be largest dynamic model of microbial species interactions ever created. Analyzing long-term data from Lake Mendota near Madison, Wisconsin, a Georgia Tech research team identified and modeled interactions among 14 sub-communities, that is, collections of different species that become dominant at specific times of the year. Key environmental factors affecting these sub-communities included water temperature and the levels of two nutrient classes: ammonia/phosphorus and nitrates/nitrites. The effects of these factors on the individual species were, in general, more pronounced than those of species-species interactions.

Beyond understanding what's happening in aquatic microbial environments, the model might also be used to study other microbial populations - perhaps even human microbiomes. The research was reported on March 24 in the journal *Systems Biology and Applications*, a Nature partner journal. The work was sponsored by the National Science Foundation's Dimensions of Biodiversity program.

"Ultimately, we want to understand why some microbial populations are

declining and why some are increasing at certain times of the year," said Eberhard Voit, the paper's corresponding author and The David D. Flanagan Chair Professor in the Wallace H. Coulter Department of Biomedical Engineering at Georgia Tech and Emory University. "We want to know why these populations are changing - whether it is because of environmental conditions alone, or interactions between the different species. Importantly, we also look at the temporal development: how interactions change over time."

Because of the large number of different microorganisms involved, creating such a model was a monumental task. To make it more manageable, the researchers segmented the most abundant species into groups that had significant interactions at specific times of the year. Georgia Tech Research Scientist Phuongan Dam created 14 such categories or sub-communities - corresponding to roughly one per month - and mapped the relationships between them during different times of the year. Two of the 14 groups had two population peaks per year.

"The exciting part about this work is that we are now able to model hundreds of species," said Kostas Konstantinidis, a co-author on the paper and the Carlton S. Wilder associate professor in Georgia Tech's School of Civil and Environmental Engineering. "The ability to dynamically model microbial communities containing hundreds or even thousands of species as those interactions change over time or after environmental perturbations will have numerous implications and applications for other research areas."

In the past, researchers have created static models of interactions between large numbers of microorganisms, but those provided only snapshots in time and couldn't be used to model interactions as they change throughout the year. Scientists might want to know, for example, what would happen if a community lost one species, if a flood of nutrients hit the lake or if the temperature rose.

As with many communities, the lake includes organisms from different species and families that are highly interconnected, playing a variety of interrelated roles, such as fixing nitrogen, carrying out photosynthesis, degrading pollutants and providing metabolic services used by other organisms. Information about the microbes came from a long-term data set compiled by other scientists who study the lake on a regular basis.

Voit, a bio-mathematician, said the model, although itself nonlinear, uses algorithms based on linear regression, which can be analyzed using standard computer clusters. Using their 14 sub-communities, the researchers found 196 interactions that could describe the species interactions - a far easier task than analyzing the 300 million potential interactions between the full 18,642 species in the lake. Reducing the number of potential interactions was possible only due to the strategy of defining sub-communities and a clever modeling approach.

The researchers initially tried to organize the microbes into genetically related organisms, but that strategy failed.

"At any time of the year, the lake needs species that can do certain tasks," said Voit. "Closely-related species tend to play essentially the same roles, so that putting them all together into the same group results in having many organisms doing the same things - but not executing other tasks that are needed at a specific time. By looking at the 14 sub-communities, we were able to get a smorgasbord of every task that needed to be done using different combinations of the microorganisms at each time."

By looking at sub-communities present at specific times of the year, the research team was able to study interactions that occurred naturally - and avoided having to study interactions that rarely took place. The model examines interactions at two levels: among the 14 sub-communities, and between the sub-communities and individual species.

The research depended heavily on metagenomics, the use of genomic analysis to identify the microorganisms present. Only 1 percent of microbial species can be cultured in the laboratory, but metagenomics allows scientists to obtain the complete inventory of species present by identifying specific sections of their DNA. Because they are not fully characterized species, the components of genomic data are termed "operational taxonomic units" (OTUs), which the team used as a "proxy" for species.

The next step in the research will be to complete a similar study of Lake Lanier, located north of Atlanta. In addition to the information studied for Lake Mendota, that study will gather data about the enzymatic and metabolic activities of the microorganism communities. Lake Lanier feeds the Chattahoochee River and a series of other lakes, and the researchers hope to study the entire river system to assess how different environments and human activities affect the microbial populations.

The work could lead to a better understanding of what interactions are necessary for a healthy lake, which may help scientists determine what might be needed to address problems in sick lakes. The modeling technique might also help scientists with other complex microbial systems.

"Our work right now is with the lake community, but the methods could be applicable to other microbial communities, including the human microbiome," said Konstantinidis. "As with sick lakes, understanding what is healthy might one day allow scientists to diagnose microbiome-related disease conditions and address them by adjusting the populations of different microorganism sub-communities."