Computational tool offers new insight into key biological processes

March 6 2014, by Matt Shipman

Researchers from North Carolina State University have developed a computational tool designed to guide future research on biochemical pathways by identifying which components in a biological system are related to specific biochemical processes, including those processes responsible for gene expression, cell signaling, stress response, and metabolism.

"Our goal was to identify modules, or functional units, which are critical to the performance of the biochemical pathways that govern a host of biological processes," says Dr. Cranos Williams, an assistant professor of electrical and computer engineering at NC State and senior author of a paper describing the work.

"For example, a car has lots of modules – the parts that make it go, the parts that make it stop, the parts that let you steer, etc. If you understand those modules, you understand how the car works. But if you just have a list of parts, that's not very helpful.

"And what we have right now for many biochemical pathways is essentially just a list of parts – metabolites, biochemical reactions and enzymes that facilitate those reactions – and, in some cases, how those parts change over time. What we need is a clear understanding of which parts work together. That's where our new algorithm comes in."

The researchers developed an algorithm that allows them to identify which parts – the metabolites, reactions and enzymes – are related to

each other and can be grouped into functional modules. The algorithm also identifies whether an individual component plays a role in multiple modules. For example, an enzyme may play a primary role in critical stress response pathways and a secondary role in processes associated with programmed cell maintenance or death.

The algorithm also characterizes how the relationships between different modules and individual components may change over time and under different internal and external conditions.

The input for the algorithm comes from using well-established dynamic models to observe changes in concentrations of metabolites, reactions and enzymes under various conditions. The algorithm then processes that data to establish primary and secondary relationships between all of the constituent parts.

"When modifying biological processes, there are thousands of possible combinations of metabolites, reactions and enzymes for any given biochemical pathway," Williams says. "Our work should help life scientists narrow down the list of key players in order to target their research efforts on functional groups that are most likely to improve our ability to understand and control important biological processes. This has applications in everything from biomedical research to agriculture to biofuels."

More information: The paper, "Hierarchical Modularization Of Biochemical Pathways Using Fuzzy-C Means Clustering," is forthcoming from *IEEE Transactions on Cybernetics*. ieeexplore.ieee.org/xpl/articl … jsp?arnumber=6651824

Abstract
Biological systems that are representative of regulatory, metabolic, or signaling pathways can be highly complex. Mathematical models that

describe such systems inherit this complexity. As a result, these models can often fail to provide a path toward the intuitive comprehension of these systems. More coarse information that allows a perceptive insight of the system is sometimes needed in combination with the model to understand control hierarchies or lower level functional relationships. In this paper, we present a method to identify relationships between components of dynamic models of biochemical pathways that reside in different functional groups. We find primary relationships and secondary relationships. The secondary relationships reveal connections that are present in the system, which current techniques that only identify primary relationships are unable to show. We also identify how relationships between components dynamically change over time. This results in a method that provides the hierarchy of the relationships among components, which can help us to understand the low level functional structure of the system and to elucidate potential hierarchical control. As a proof of concept, we apply the algorithm to the epidermal growth factor signal transduction pathway, and to the C3 photosynthesis pathway. We identify primary relationships among components that are in agreement with previous computational decomposition studies, and identify secondary relationships that uncover connections among components that current computational approaches were unable to reveal.

Provided by North Carolina State University