

TOOLBOX

New algorithm plots genes' relationships

BY JESSICA WRIGHT

17 APRIL 2013

A new software algorithm called GAP measures the similarities in the function of more than 19,000 human protein-coding genes, according to a study published 14 March in *BMC Systems Biology*¹. The similarity scores allowed the researchers to identify 114 potential autism candidate genes, they say.

Studies in the past few years have identified many candidate genes linked to autism. To home in on the key pathways involved in the disorder, scientists often **construct networks** that show how these genes interact with each other, based on **direct interactions** between the proteins the genes encode. However, these networks may be based only on interactions among the most abundant proteins.

What's more, protein-protein interactions tell only part of the story, the researchers note. Network analyses should include interactions between proteins and either DNA or RNA, they say.

In the new study, they instead developed an algorithm for protein function using information from various databases that list their function, disorders involving these proteins, and the drugs that target them. This approach allows them to discover more general relationships between genes that go beyond direct interactions, the researchers say. The algorithm predicts more than a million associations among the proteins.

Making connections: The GAP algorithm (left) finds more functional links between autism-associated genes than does another method, called GeneMANIA (right). [Enlarge image »](#)

To test their approach, the researchers looked at the connections between 25 of the top autism-linked genes listed in **SFARI Gene** — a comprehensive list of genes linked to autism. SFARI Gene is funded by the Simons Foundation, SFARI.org's parent organization. Because these genes are

all associated with autism, they are likely to have related functions, the researchers say.

GAP finds significantly more connections between these genes than do two other network algorithms, GeneMANIA and STRING. The other two algorithms test their predicted associations against a reference dataset, which may result in false negatives, the researchers say.

The researchers then used GAP to identify 11,215 genes that share functions with each of the 306 genes listed in SFARI gene. The top one percent of these genes may also play a role in the disorder, the researchers say.

References:

1: Vafae F. et al. *BMC Syst. Biol.* 7, 22 (2013) [PubMed](#)