

NEWS

# New approach mines multiple sources to link genes to disease

BY JESSICA WRIGHT

14 NOVEMBER 2011

Dance card: Maps of interactions among candidate genes for a disorder can point researchers towards key pathways.

**Dance card:** Maps of interactions among candidate genes for a disorder can point researchers towards key pathways.

Researchers have developed a computational technique that allows them to simultaneously use multiple types of information — including gene expression, ratings and associations — to identify candidate genes for a disorder. The unpublished results were presented Sunday at the **2011 Society for Neuroscience annual meeting** in Washington, D.C.

Computational scientists typically begin with a single dataset and then add other data to the model to refine the results, says Robert Peitzsch, head of **DKP Genomics**, a Connecticut-based biotechnology company with a focus on data analysis.

"The problem with this is that it inherently builds in some bias from the beginning because we focused on a particular dataset," says Peitzsch. Because the added data filters the results of the

original dataset, some of the original information may also become lost.

To resolve this issue, Peitzsch and his colleagues have developed a method called multiple parameter optimization, which runs the initial analysis simultaneously using all of the available data.

The researchers have already applied this method to Alzheimer's disease to identify 25 candidate genes. The approach can be used similarly for many disorders, including autism, says Peitzsch.

Because multiple parameter optimization is based on a simple linear scoring system, with a '1' for the most desirable response and '0' for the least, the analysis can be applied to multiple scoring systems, the researchers say.

For example, the Alzheimer's model includes data from **AlzGene**, a curated database that rates the association of a given gene to Alzheimer's as 'A,' 'B' or 'C.' In that case, A is given a desirability score of 1, B is 0.75 and C is 0.5. Genes not included in AlzGene are given a score of 0.

The analysis also includes results from gene interactions with the Alzheimer's candidate gene apolipoprotein E or APOE, and the **Online Mendelian Inheritance in Man** database.

The 25 genes identified include known candidates such as APOE and DNML1, as well as some genes, such as SPRY4, that the researchers didn't expect to score so high. "Simply because it incorporates all of the data in the original analysis, we were able to catch genes that we probably would have missed otherwise," says Peitzsch.

Mapping the interactions among these top genes implicated pathways such as metabolism and cell signaling.

One shortcoming of this approach is that genes missing data for one or several of the measures could skew the results. Still, the method is promising because the pathways identified match those linked to Alzheimer's strikingly well, notes Peitzsch. "I forgot to tell my colleague that this was Alzheimer's data, and he came back to me and said, "Is this Alzheimer's data?"

For more reports from the 2011 Society for Neuroscience annual meeting, please [click here](#).