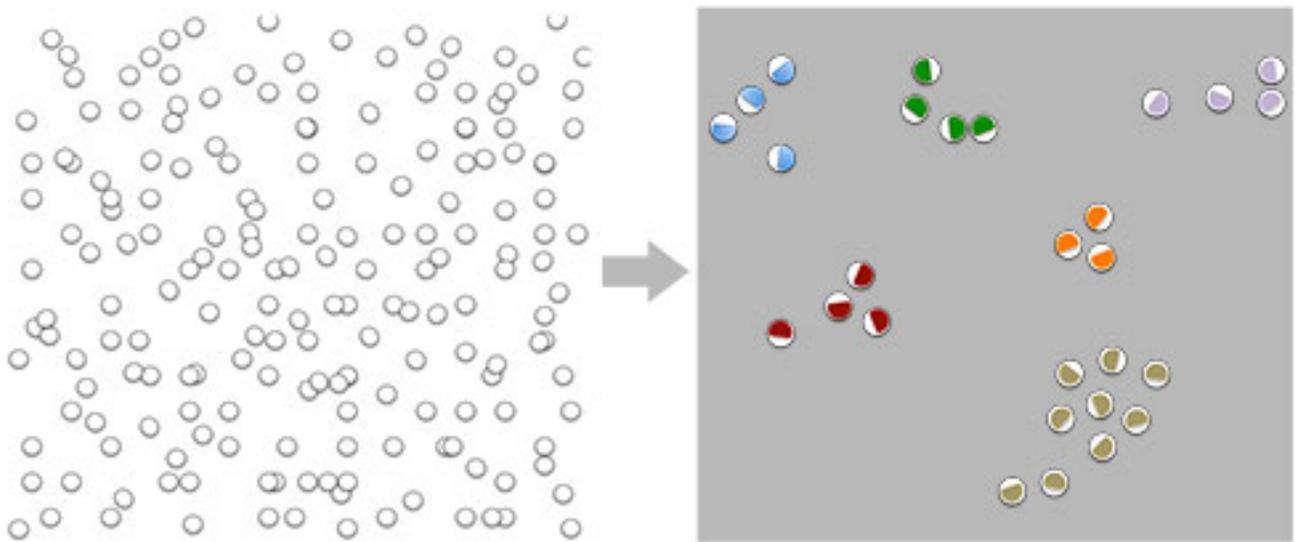


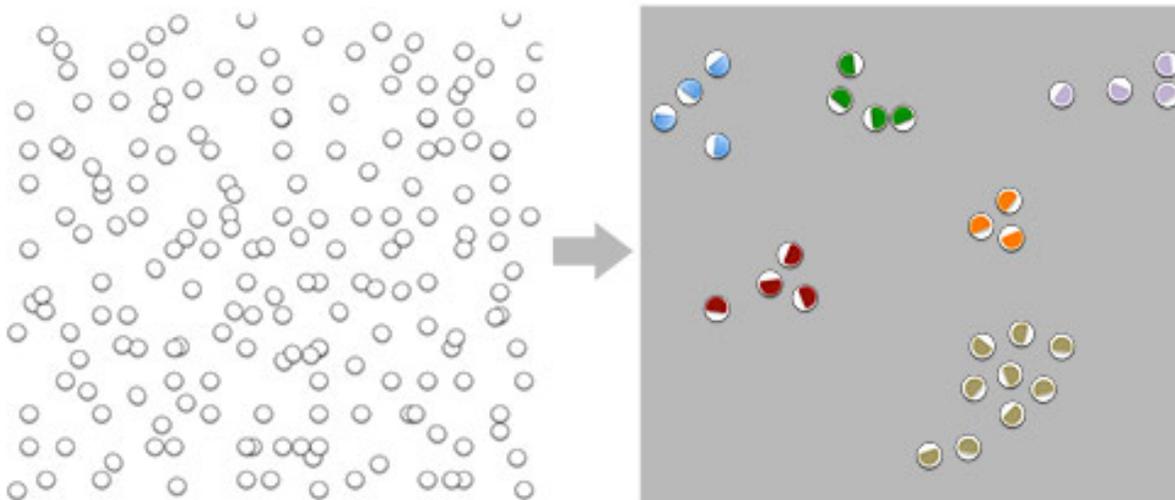
NEWS

Algorithm uncovers autism syndromes' fingerprints

BY JESSA NETTING

6 MARCH 2014





Sorting symptoms: A new algorithm can analyze behavioral data from people with autism (left) and group the cases into six behavioral clusters (right).

Autism is defined based on a wide variety of **behavioral symptoms**, but it's precisely this variation — along with a complex genetic background — that makes it tricky to **connect behavior** to the underlying genes^{1, 2}.

A new algorithm may make this challenge a bit easier to solve. The algorithm, which employs a form of artificial intelligence that learns as it goes, analyzes behavioral data and has learned to recognize six genetic disorders associated with autism, according to research published 11 February in *Molecular Autism*³.

The researchers hope to use these behavioral signatures to hone their search for the genetic underpinnings of 'idiopathic autism,' for which there is no known cause.

"There was a sort of assumption that genetic risk factors probably lead to a distinguishable set of behavioral phenotypes, but people had never really formally tested that proposition before," says lead investigator **Patrick Bolton**, professor of child and adolescent psychiatry at Kings College London. "That was the motivation for this project."

Previous studies searched for ways in which different syndromes may stem from **shared molecular or neurological pathways**, but they showed no clear consensus on how to do so⁴.

The new machine-learning technique, called support vector, sifts through large volumes of data to find identifiable patterns that can be used to subdivide a group of people. The system combed through medical data from individuals diagnosed with one of six autism-linked genetic disorders: 22q11.2 deletion syndrome, Down syndrome, **Prader-Willi syndrome**, tuberous sclerosis complex, Klinefelter syndrome and supernumerary marker chromosome 15, which is caused by a duplication of a chromosome 15 segment.

The method was able to identify behavioral signatures specific to each syndrome. And as it did so, it built an algorithm that can find the same kinds of signatures in behavioral data from people with idiopathic autism.

The researchers tested their algorithm in three blinded samples totaling 1,261 individuals with autism and found instances of all six signatures.

“The thing that I find most impressive about this study is that they are extending their classifier to idiopathic autism and they are seeing a signature there,” says **Dennis Wall**, associate professor of pediatrics at Stanford University, who has used **machine learning** to develop autism diagnostic tools but was not involved in this research.

“It suggests that you might be able to use this, or methods like this, to find a signature of genotype-phenotype correlation that can act as a gravitational force to make sense of what is currently a very heterogeneous and complex picture,” Wall says.

Exposing patterns:

The study’s results are preliminary, but it offers proof of concept that the method can link certain behavior, or a phenotype, to a specific genetic makeup, or genotype. Shared behavioral signatures may indicate shared gene pathways that lead to the behaviors, which in turn might hint at autism’s cause.

“The power of support-vector machine learning is that it can find hidden patterns — that is, patterns that would not be detected by conventional unsupervised statistical analysis,” says **Hilgo Bruining**, assistant professor of child and adolescent psychiatry at the University of Utrecht who also led the study.

The researchers plan to sift through large datasets of both behavioral and genetic data from people with idiopathic autism. If the algorithm can identify new behavioral signatures among these datasets, it may be able to cluster them into subgroups of autism and zero in on areas of the genome responsible for those subtypes of the disorder.

Some experts urge caution with this line of reasoning, however.

“While it should have been expected that autism phenotypes could be segregated around a diverse set of genetic lesions, the suggestion that these findings mean different types of autism is an overstatement,” says **Valsamma Eapen**, professor of infant child and adolescent psychiatry at the University of New South Wales in Australia, who was not involved in the study.

For their study, the researchers analyzed behavioral data from the medical records of 322 individuals at the University Medical Center in Utrecht in the Netherlands and the Institute of Psychiatry at Kings College London. Subgroups diagnosed with each of the six autism-linked syndromes included between 21 and 90 people.

In particular, the researchers fed data for 37 behaviors, including verbal rituals, imaginative play and unusual preoccupations, into their system.

When tasked with distinguishing among the six syndromes, the algorithm correctly identified a syndrome from its behavioral signature 63 percent of the time.

“Sixty-three percent is a little low, but it’s not meant to be a landslide or a slam dunk,” says Wall.

The algorithm was more accurate when comparing just two syndromes at a time. For example, it could distinguish 22q11 syndrome from supernumerary marker chromosome 15 with 97 percent accuracy, its best performance. It fared most poorly when trying to distinguish Down syndrome from the other five disorders.

To test the algorithm in cases of autism, the researchers fed it behavioral data from 1,261 people enrolled in the Autism Genetics Research Exchange (AGRE). However, these data do not include information on whether the individuals are diagnosed with one of the six syndromes, so the algorithm can only generate a probability that each identification is correct.

Within the first sample of 375 people from AGRE, the algorithm categorized 255 people with the behavioral signature of tuberous sclerosis complex, with a 61 percent probability that this analysis is correct. Figures for the other syndromes are lower, with several of them in the 40 percent range and Down syndrome once again the lowest, at 25 percent probability of accuracy.

The researchers then repeated the test with two sets of 443 people each, producing similar results.

Overall, they say, the pattern of social difficulties is the most useful in helping identify the syndromes. This suggests that social impairment, such as the extreme social avoidance seen in people with **fragile X syndrome**, may be related to particular genetic risk factors, the researchers say.

Eapen says she hopes to see studies in which the disorders’ cause and the underlying mechanisms are better characterized.

“In these gross deletion syndromes, no one is sure of which genes are linked to autism and how the size of the deletions and duplications impacts genes or phenotypes,” she says. “This study is setting the stage to look more closely at diagnosis in select cohorts based on a genetic lesion.”

References:

- 1: Eapen V. *et al. Front. Hum. Neurosci.* **7**, 567(2013) [PubMed](#)
- 2: McCray A.T. *et al. Neuroinform.* Epub ahead of print (2013) [PubMed](#)
- 3: Bruining H. *et al. Mol. Autism* **5**, 11 Epub ahead of print (2014) [PubMed](#)
- 4: Zafeiriou D.I. *et al. Am. J. Med. Genet. B Neuropsychiatr. Genet.* **162B**, 327-366 (2013) [PubMed](#)