NEWS

Mutations tied to autism rife in ‘junk’ DNA

BY **ANN GRISWOLD**

9 SEPTEMBER 2016

More than one-third of the spontaneous mutations linked to autism crop up in genomic segments that do not code for genes. That’s the upshot from the largest study to date of genomes from children with autism[1].

Most efforts to uncover spontaneous, or *de novo,* mutations tied to autism have focused on the 1.5 percent of the genome that encodes genes. Mutations in these coding regions may underlie autism

in as much as **11 percent of children** with the condition.

The new study, published 3 August in *NPJ Genomic Medicine*, suggests that *de novo* mutations in the remaining 98.5 percent of the genome also contribute to autism risk. These noncoding swaths, once considered 'junk DNA,' may play an **important role in regulating gene expression**.

The study is too small to connect specific mutations to autism risk, however.

"Even using whole-genome sequencing and throwing everything into the analysis, we're just starting to understand the genetic circuitry involved in autism," says senior researcher **Stephen Scherer**, director of the Centre for Applied Genomics at the Hospital for Sick Children in Toronto. "We need bigger numbers, better controls and more sophisticated analysis methods."

Scherer and his colleagues looked at sequences from 200 children with autism and their unaffected parents. The study is part of the **MSSNG** initiative, a project by the research and advocacy organization **Autism Speaks** that aims to sequence the genomes of 10,000 people with autism and their family members.

Crunching code:

In the new study, researchers turned up 244 spontaneous mutations in the children with autism (meaning the mutations are not present in their parents) — nearly twice as many as in 258 adults and children in the general population. Data from these controls came from a separate analysis of a Dutch population[2].

They found that 151 of the 244 mutations are in the coding parts of the genome; of the 93 mutations in noncoding sequences, 55 fall in regions just before or after a gene.

Mutations in coding regions occur most frequently in genes involved in protein translation, building connections between neurons or remodeling chromatin, the coiled complex of DNA and protein. By contrast, mutations in noncoding regions crop up largely in areas near these genes that serve as landing strips for proteins that turn the genes on or off.

However, a sample of 200 families may not have the statistical power to gauge the significance of these mutations, says **Stephan Sanders**, assistant professor of psychiatry at the University of California, San Francisco, who was not involved in the work.

The weaker a mutation's effects, he says, the larger a study must be to judge its relevance to a given condition. For severe mutations that disable a gene, researchers might need sequences from about 400 families to analyze; for mutations that are less severe, that number might need to be at least 1,000. Mutations in noncoding DNA may have weak or indirect effects, so reliably pinpointing harmful mutations there would require even larger numbers, Sanders says.

Parental input:

What's more, it is unclear whether the children with autism have an unusually high number of spontaneous mutations, because the control group was sequenced separately. Variations between the two analyses could produce misleading results, Sanders says.

More than 75 percent of the spontaneous mutations the researchers found originated in genetic material passed down from the father, and the number of these mutations increases with a father's age. In a 2012 study, researchers found that for **every year a father ages**, he passes on two additional spontaneous mutations to his children.

Nearly half of the 244 spontaneous mutations are within 200 base pairs of each other. These clusters are more likely to occur in DNA from mothers — but the researchers don't yet know why.

Despite its small size, the study is an important first step toward understanding genetic variation in autism, the researchers say. So far, more than 5,000 genome sequences are **freely available** to researchers, and an additional 2,000 sequences are pending final review by MSSNG investigators.

"We need to strive to get all of the whole-genome sequence data together to continue to move the field forward," Scherer says.

REFERENCES:

1. Yuen R.K. *et al. NPJ Genom. Med.* **1**, 160271-1602710 (2016) **PubMed**
2. Genome of the Netherlands Consortium *et al. Nat. Genet.* **46**, 818–825 (2014) **PubMed**