

NEWS

Autism research may require tens of thousands of samples

BY EMILY SINGER

18 JUNE 2012





Population explosion: Rapid growth in the human population over the past 10,000 years has led to large numbers of rare genetic variants.

Population explosion: Rapid growth in the human population over the past 10,000 years has led to large numbers of rare genetic variants.

Rare variants make up the vast majority of human genetic variation, according to two independent papers published in May in *Science*^{1,2}. That means that genetic studies of complex diseases such as autism are likely to require tens of thousands of participants.

The researchers found that 80 to 95 percent of genetic variation is rare — present in less than 0.5 percent of the population. “That’s very striking; not many people would have predicted it,” says **Joshua Akey**, associate professor of genome sciences at the University of Washington and senior scientist on one of the studies.

Most studies, including those on autism genes, have focused on more common variants, found in

more than one percent of the population.

Akey and his collaborators sequenced nearly 16,000 protein-coding genes from 2,440 people of either European or African descent. What makes the study unique is that they sequenced each gene an average of 111 times, allowing them to distinguish errors that occur during the sequencing process from real mutations in the genome.

The second study analyzed 202 genes that encode drug targets from 14,002 people. They also found an abundance of rare variants — about 1 in every 17 bases — many of them in a single individual.

“I think we’re finally getting a handle on what rare variation looks like,” says **Daniel MacArthur**, group leader of the Analytic and Translational Genetics Unit at Massachusetts General Hospital in Boston, who was not involved in either study.

Complex causes:

Autism researchers already had hints of this complexity. Three studies published in April sequenced the **exomes** — the protein-coding regions of the genome — of about 2,000 people, including more than 600 with autism. They estimated that **as many as 1,000 genes** are linked to the disorder.

“It means we are going to have hundreds of different genetic etiologies or major risk factors for autism, which is consistent with the tremendous clinical heterogeneity,” says **David Ledbetter**, chief scientific officer at Geisinger Health System in Danville, Pennsylvania, who was not involved in the studies.

Both groups of researchers say the abundance of rare variants resulted from the explosion of the human population in the past 10,000 years: The number of people on the planet has grown by three orders of magnitude in just 400 generations³.

Much of the genomics research to date has focused on common variants. For obvious reasons, it’s much easier to find people with a common variant to study a particular disease.

However, genetic studies show that common variants tend to have a small influence on the risk of common diseases. This is not surprising: If the variants had large effects on disease, evolution would have weeded them out of the genome.

“Most of the time, evolution will stop harmful mutations from reaching a high frequency in the population,” says MacArthur. But the rare mutations identified in the new studies are probably still too young to have been weeded out, suggesting they are likely to play an important role in disease. In fact, “the more deleterious a mutation is, the rarer it should be,” says MacArthur.

The emphasis on **rare variants' role** in disease has been growing steadily as sequencing technology has grown faster and cheaper, allowing scientists to study more and more human genomes.

“Prior to the last couple of years, it was very expensive to look at large numbers of individuals, so our understanding of genetic variation was based on small sample sizes,” says Akey. “We found that understanding of genetic variation is dramatically different when you look at thousands of individuals versus hundreds.”

Based on their findings, the researchers estimate that studies designed to uncover the genetic roots of complex diseases, such as autism, will require tens of thousands of people.

“The answer will be in larger sample sizes,” says MacArthur. “The community and funders need to recognize that getting a return on investment will require sufficient sample sizes to see results.”

Large numbers won't be enough on their own, either. Researchers will also need to develop new tools for understanding the role that rare variants play.

“The biggest challenge is distinguishing variants that are functionally important from those that are benign,” says Akey. A study published in February found that major errors in the genome are **sometimes harmless**.

Scientists will soon have enough data to test out new analysis tools. Studies funded by the National Institutes of Health and others will have sequenced tens of thousands of exomes or genomes by the end of the year.

“The next 6 to 12 months will be very exciting in terms of our understanding of genetic variation,” says MacArthur. “Having all this data surging around is really exciting.”

References:

1: Tennessen J.A. *et al.* *Science* Epub ahead of print (2012) [PubMed](#)

2: Nelson M.R. *et al.* *Science* Epub ahead of print (2012) [PubMed](#)

3: Keinan A. and A.G. Clark *Science* **336**,740-743 (2012) [PubMed](#)