Study suggests rare genetic variants most likely to influence disease

31 March 2011

New genomic analyses suggest that the most common genetic variants in the human genome aren't the ones most likely causing disease. Rare genetic variants, the type found most often in functional areas of human DNA, are more often linked to disease, genetic experts at Duke University Medical Center report.

The study was published in the American Journal of Human Genetics on March 31.

"The more common a variant is, the less likely it is to be found in a functional region of the genome," said senior author David Goldstein, Ph.D., director of the Duke Center for Human Genome Variation. "Scientists have reported this observation before, but this study is the most comprehensive effort to date using annotations of the functional regions of the human genome and fully sequenced genomes."

Goldstein said that "the magnitude of the effect is dramatic and is consistent across all frequencies of variants we looked at." He also said he was surprised by the notable consistency of the finding. "It's not just that the most rare variants are different from the most common, it's that at every increase in frequency, a variant is less and less likely to be found in a functional region of the DNA," Goldstein said. "This analysis is consistent with what appears to be a growing consensus that common variants are less important in common diseases than many had originally thought."

The researchers established simple criteria to learn which genetic variants were functional, based on their locations. They looked at the regions in genes that make proteins and also functional regions that influence the expression of proteins.

The scientists sequenced the complete genomes of 29 people (of European origin) to assess the relationship between the functional properties of the variants and their population allele frequencies. (An allele is one member of a pair of genes, each contributed by a mother and a father, found at a specific location on a specific chromosome).

"We also asked whether we could identify any patterns by examining variants in which the derived form (i.e., the one that appeared in humans and is not seen in related primates) was the most common form," said Goldstein, the Richard and Pat Johnson Distinguished University Professor in molecular genetics and microbiology. "That is the unusual case, because when there is a mutation that changes an allele from the ancestral allele, this is usually the more rare form in the population. If you do have a mutation that is beneficial, then the variant can increase in frequency and become the common one."

When the researchers focused on the derived form that was also the common form of the allele, they didn't know quite what to expect, but one possibility was that the common alleles would be more enriched in functional locations of DNA.

"You might expect the findings to turn around in this case and go in the other direction," Goldstein said. "Instead, we observed that there was no connection at all between the frequency of these particular common variants in a population and whether they were in functional regions of the genome or not. The pattern just evaporated."

He said the obvious interpretation is that whenever a variant does become common, it does so precisely because it has no impact. "The bottom line is that we see common variants, as a rule, being neutral and not having effects," Goldstein said. "There are some big exceptions to this rule, in particular in the HLA gene region where selection for resistance to infectious disease has resulted in many common variants with major effects. But in general in the genome, we see this as very much the exception."

Goldstein noted that before scientists developed sequencing tools to look at genomewide effects, they only knew there were some common variants that caused disease, with one of the classic examples being the strong effect of Apoe4 on the risk of Alzheimer's disease. "Many people expected to see more examples of common variants having an important collective impact on common diseases, but that didn't happen," Goldstein said.

"One really interesting aspect of the work is that for every functional category we considered, there is preferentially exclusion of common variants," said Qianqian Zhu, Ph.D., lead author and postdoctoral associate in the Goldstein laboratory. "This observation helps to illustrate how well the scientific community is succeeding in identifying the functional regions of the human genome."

In the future, scientists will need to develop sophisticated ways to use population allele frequency to directly facilitate the search for disease-causing mutations, Zhu said.

"If it is true that a lot of human disease is due to relatively rare variants that are relatively harmful, then we are going to be able to find them. It may take sequencing thousands of patient genomes to track down the responsible mutations, but they will be found," Goldstein said. "I am entirely convinced that sequencing, which is becoming less expensive every month, will unlock a lot of the causes of genetic disease. What we can do clinically with that information will become the primary challenge."

Provided by Duke University Medical Center