Researchers release most detailed global study of genetic variation

February 20 2008

A schematic of worldwide human genetic variation, with colors representing different genetic types. The figure illustrates the great amout of genetic variation in Africa. Illustration by Martin Soave/University of Michigan

University of Michigan scientists and their colleagues at the National Institute on Aging have produced the largest and most detailed worldwide study of human genetic variation, a treasure trove offering new insights into early migrations out of Africa and across the globe.

Like astronomers who build ever-larger telescopes to peer deeper into

space, population geneticists like U-M's Noah Rosenberg are using the latest genetic tools to probe DNA molecules in unprecedented detail, uncovering new clues to humanity's origins.

The latest study characterizes more than 500,000 DNA markers in the human genome and examines variations across 29 populations on five continents.

"Our study is one of the first in a new wave of extremely high-resolution genome scans of population genetic variation," said Rosenberg, an assistant research professor at U-M's Life Sciences Institute and co-senior author of the study, to be published in the Feb. 21 edition of *Nature*.

"Now that we have the technology to look at thousands, or even hundreds of thousands, of genetic markers, we can infer human population relationships and ancient migrations at a finer level of resolution than has previously been possible."

The new study, led by Rosenberg and National Institute on Aging colleague Andrew Singleton, produced genetic data nearly 100 times more detailed than previous worldwide assessments of human populations. It shows that:

• A recently discovered type of human genetic variation, known as a copy-number variant or CNV, is a reliable addition to the toolkit of population geneticists and should speed the discovery of disease-related genes. Rosenberg and his colleagues discovered 507 previously unknown CNVs, which are large chunks of DNA—up to 1,000,000 consecutive "letters" of the genetic alphabet—that are either repeated or deleted entirely from a person's genome. Various diseases can be triggered by an abnormal gain or loss in the number of gene copies.

• It's sometimes possible to trace a person's ancestry to an individual population within a geographic region. While previous studies have found that broad-scale geographic ancestry could be successfully traced, the new results indicate "it's becoming increasingly possible to use genomics to refine the geographic position of an individual's ancestors with more and more precision," Rosenberg said.

• Human genetic diversity decreases as distance from Africa—the cradle of humanity—increases. People of African descent are more genetically diverse than Middle Easterners, who are more diverse than Asians and Europeans. Native Americans possess the least-diverse genomes. As a result, searching for disease-causing genes should require the fewest number of genetic markers among Native Americans and the greatest number of markers among Africans.

The results are being made available on publicly shared databases.

"I hope the study will be an invaluable resource for understanding genomic variability and investigating genetic association with disease," said the NIA's Singleton.

The researchers analyzed DNA from 485 people. They examined three types of genetic variation: single-nucleotide polymorphisms, or SNPs; haplotypes; and CNVs.

If the human genome is viewed as a 3-billion-letter book of life, then SNPs represent single-letter spelling changes, haplotype variations equate to word changes, and CNVs are wholesale deletions or duplications of full pages.

The patterns revealed by the new study support the idea that humans originated in Africa, then spread into the Middle East, followed by Europe and Asia, the Pacific Islands, and finally to the Americas.

The results also bolster the notion of "serial founder effects," meaning that as people began migrating eastward from East Africa about 100,000 years ago, each successive wave of migrants carried a subset of the genetic variation held by previous groups.

"Diversity has been eroded through the migration process," Rosenberg said.

In addition to his position at the Life Sciences Institute, Rosenberg is an assistant professor of human genetics at the Medical School; an assistant professor of biostatistics at the School of Public Health; an assistant professor of ecology and evolutionary biology at the College of Literature, Science, and the Arts; and an assistant research professor of bioinformatics at the Medical School's Center for Computational Medicine and Biology.

"This data set is so rich. It provides a much more comprehensive, cross-sectional snapshot of the human genome than previous studies," said Paul Scheet, a post-doctoral researcher in the U-M Department of Biostatistics and one of the lead authors.

"The next step for these studies is to sequence whole genomes," said Mattias Jakobsson, a post-doctoral researcher at the U-M Center for Computational Medicine and Biology and another lead author. "You would take 500 individuals, and you would just completely sequence everything, and then you'd have almost every important variant that's out there."

Source: University of Michigan