

Global genome effort seeks genetic roots of disease

October 31 2012



By decoding the genomes of more than 1,000 people whose homelands stretch from Africa and Asia to Europe and the Americas, scientists have compiled the largest and most detailed catalog yet of human genetic variation. The massive resource will help medical researchers find the genetic roots of rare and common diseases in populations worldwide. Credit: National Human Genome Research Institute

By decoding the genomes of more than 1,000 people whose homelands stretch from Africa and Asia to Europe and the Americas, scientists have compiled the largest and most detailed catalog yet of human genetic variation. The massive resource will help medical researchers find the genetic roots of rare and common diseases in populations worldwide.

The 1000 Genomes Project involved some 200 scientists at Washington University School of Medicine in St. Louis and other institutions. Results detailing the DNA variations of individuals from 14 ethnic groups are published Oct. 31 in the journal *Nature*. Eventually, the initiative will involve 2,500 individuals from 26 populations.

"With this resource, researchers have a roadmap to search for the [genetic origins](#) of diseases in populations around the globe," says one of the study's co-principal investigators, Elaine Mardis, PhD, co-director of The [Genome](#) Institute at Washington University. "We estimate that each person carries up to several hundred rare DNA variants that could potentially contribute to disease. Now, scientists can investigate how detrimental particular rare variants are in different ethnic groups."

At the [genetic level](#), any two people are more than 99 percent alike. But rare variants – those that occur with a frequency of 1 percent or less in a population – are thought to contribute to [rare diseases](#) as well as common conditions like cancer, heart disease and diabetes. Rare variants may also explain why some medications are not effective in certain people or cause side effects such as nausea, vomiting, insomnia and sometimes even [heart problems](#) or death.

Identifying rare variants across different populations is a major goal of the project. During the pilot phase of the effort, the researchers found that most rare variants differed from one population to another, and that they developed recently in human evolutionary history, after populations in Europe, Africa, Asia and the Americas diverged from a single group. The current study bears this out.

"This information is crucial and will improve our interpretation of individual genomes," says another of the study's co-[principal investigators](#), Richard K. Wilson, PhD, director of The Genome Institute and a pioneer in cancer genome sequencing. "Now, if we want to study

cancer in Mexican Americans or Japanese Americans, for example, we can do so in the context of their diverse geographic or ancestry-based genetic backgrounds."

Results of the new study are based on DNA sequencing of the following populations: Yoruba in Nigeria; Han Chinese in Beijing; Japanese in Tokyo; Utah residents with ancestry from northern and western Europe; Luhya in Kenya; people of African ancestry in the southwestern United States; Toscani in Italy; people of Mexican ancestry in Los Angeles; Southern Han Chinese in China; Iberian from Spain; British in England and Scotland; Finnish from Finland; Colombians in Columbia; and Puerto Rican in Puerto Rico.

All study participants submitted anonymous DNA samples and agreed to have their genetic data included in an online database. To catalog the variants, the researchers first sequenced the entire genome – all the DNA – of each individual in the study about five times. Surveying the genome in this way finds common DNA changes but misses many rare variants.

Then, to find rare variants, they repeatedly sequenced the small portion of the genome that contains genes – about 80 times for each participant to ensure accuracy – and they looked closely for single letter changes in the DNA sequence called SNPs (for single-nucleotide polymorphisms).

Using special tools developed to analyze and integrate the data, the researchers discovered a total of 38 million SNPs, including more than 99 percent of the variants with a frequency of at least one percent in the participants' DNA samples. They also found numerous structural variations, including 1.4 million short stretches of insertions or deletions and 14,000 large DNA deletions.

SNPs and structural variants can help explain an individual's susceptibility to disease, response to drugs or reaction to environmental

factors such as air pollution or stress. Other studies have found an association between small insertions and deletions and diseases such as autism and schizophrenia.

The 1000 Genomes Project has generated massive amounts of genomic data. Simply recording the raw information took up some 180 terabytes of hard-drive space, enough to fill more than 40,000 DVDs. All of the information is freely available on the Internet through public databases.

"This tremendous resource builds on the knowledge of the Human Genome Project," says co-author George Weinstock, PhD, associate director of The Genome Institute. "Scientists and, ultimately, patients worldwide will benefit from the extensive effort to understand the shared features and geographic diversity of the human genome."

More information: The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature*. Oct. 31, 2012.

Provided by Washington University School of Medicine

Citation: Global genome effort seeks genetic roots of disease (2012, October 31) retrieved 20 March 2024 from

<https://medicalxpress.com/news/2012-10-global-genome-effort-genetic-roots.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--