Previous genetic association studies involving people with European ancestry may be inaccurate

November 7 2023

Researchers have found that previous studies analyzing the genomes of people with European ancestry may have reported inaccurate results by not fully accounting for population structure. Credit: Darryl Leja, National Human Genome Research Institute

Researchers have found that previous studies analyzing the genomes of

people with European ancestry may have reported inaccurate results by not fully accounting for population structure. By considering mixed genetic lineages, researchers at the National Human Genome Research Institute (NHGRI), part of the National Institutes of Health, demonstrated that previously inferred links between a genomic variant that helps digest lactose and traits such as a person's height and cholesterol level may not be valid.

The study, published in *Nature Communications*, shows that people with European ancestry, who were previously treated as a genetically homogenous group in large-scale genetic studies, have clear evidence of mixed genetic lineages, known as admixture. As such, the results from previous genome-wide association studies that do not account for admixture in their examinations of people with European ancestry should be re-evaluated.

"By reading population genetics papers, we realized that the pattern of genetic makeup in Europe is too detailed to be viewed on a continental level," said Daniel Shriner, Ph.D., staff scientist in the NIH Center for Research on Genomics and Global Health and senior author of the study. "What is clear based on our analysis, is when data from genetic association studies of people of European ancestry are evaluated, researchers should adjust for admixture in the population to uncover true links between genomic variants and traits."

To look at European genetic ancestry, the researchers collated data in published genetic association studies and generated a reference panel of genomic data that included 19,000 individuals of European ancestry across 79 populations in Europe and European Americans in the U.S., capturing ancestral diversity not seen in other large catalogs of human genomic variation.

As an example, the researchers investigated the lactase gene, which

encodes a protein that helps digest lactose and is highly varied across Europe. Using the new reference panel, they analyzed how a genomic variant of the lactase gene is related to traits such as height, body mass index and low-density lipoprotein cholesterol, also known as "bad cholesterol."

When the researchers considered the genetic admixture of the European population in their analysis, they found that the genomic variant that gives people the ability to digest lactose is not linked to height or level of low-density lipoprotein cholesterol. In contrast, the same variant does influence body mass index.

"The findings of this study highlight the importance of appreciating that the majority of individuals in populations around the world have mixed ancestral backgrounds and that accounting for these complex ancestral backgrounds is critically important in genetic studies and the practice of genomic medicine," says Charles Rotimi, Ph.D., NIH Distinguished Investigator, director of the Center for Research on Genomics and Global Health and senior author of the study.

While the lactase gene is one example of a gene that may be incorrectly linked to some traits based on previous analyses, the researchers say it's likely that there are other false associations in the literature and that some true associations are yet to be found. Information about how genomic variants are related to different traits helps researchers estimate polygenic risk scores and may give clues about a person's ability to respond safely to drug treatments.

While the differences in any two people's genomes are less than 1%, the small percentage of genomic variation can give clues about where a person's ancestors might have come from and how different families might be related. Information about who a person is biologically descended from, known as genetic ancestry, can give important clues

about genetic risks for common diseases.

"Finding true genetic associations will help researchers be more efficient and careful with how further research is conducted," said Mateus Gouveia, Ph.D., research fellow in the Center for Research on Genomics and Global Health and first author of the study. "We hope that by accounting for mixed ancestries in future genomic analyses, we can improve the predictive value of polygenic risk scores and facilitate genomic medicine."

The reference panel generated in this study is available to the scientific community for use in other studies, with additional information provided in the paper.

More information: Unappreciated Subcontinental Admixture in Europeans and European Americans: Implications for Genetic Epidemiology Studies, *Nature Communications* (2023). DOI: 10.1038/s41467-023-42491-0

Provided by NIH/National Human Genome Research Institute