

Genetic diversity of European Americans and disease gene mapping

January 18 2008

Labels such as “European American”, “white”, or “Caucasian” are often viewed as representing a homogeneous category in gene mapping studies and census reports, but each of these labels actually groups together multiple populations, which have diverse origins due to the complex history of European immigration to the United States.

In a recent study, published in the open-access journal *PLoS Genetics*, an international team of researchers provide the first genetic dissection of the population structure of European Americans, focusing on identifying the contributions from different genetic ancestries that are important for disease gene mapping.

This is a timely issue as the last year has seen a dramatic upswing in genetic association studies and the discovery of almost a hundred new risk factors for common genetic diseases such as cancer and diabetes. If the subtle population substructure that exists within European American populations is not understood and accounted for, genetic association studies can produce incorrect findings if disease cases are compared to healthy controls that on average have different ancestry.

By systematically examining data from four actual disease association studies in European Americans, this study describes and characterizes the majority of population substructure in European Americans that could lead to spurious associations. “Although our work is far from a complete description of European American population history, for the purpose of disease gene mapping studies it is adequate to measure how

closely each person's genetic ancestry resembles three populations that can be roughly described as northwest European, southeast European, or Ashkenazi Jewish," says Dr. David Reich, one of the senior authors on the study, an Associate Professor of Genetics at Harvard Medical School and an Associate Member at the Broad Institute of Harvard and MIT. "With this approach, we can avoid most false-positive associations due to population substructure in European American disease gene mapping studies. Our previous work has addressed related challenges in studies of African Americans and Latino Americans."

Based on their discovery that ancestry from only three populations accounts for most of the potentially problematic substructure in European American disease association studies, the researchers scoured through published data sets to identify places in the genome where common DNA sequence variants differ substantially in frequency among these three ancestral populations and are therefore potentially informative for estimating genetic ancestry. The investigators then confirmed the utility of these genetic variants by testing them in DNA samples that their coauthors collected from the United Kingdom, Sweden, Poland, Spain, Italy, Greece and U.S. Ashkenazi Jews. "We identified 300 common genetic variants that have unusually different frequencies in the three ancestral populations: they are about 10 times more informative for predicting the ancestry of European Americans than random genetic variants", says lead author Dr. Alkes Price, a post-doctoral researcher at the Harvard Medical School Department of Genetics and the Broad Institute of Harvard and MIT. "We can thus correct for population substructure in European American disease association studies using just these 300 markers."

This panel of 300 markers should be valuable in targeted associated studies that follow up previously implicated candidate genes: by comparing the ancestry of disease cases to healthy controls using data from the panel of 300 markers, researchers can determine whether

observed associations are genuine, and not false-positives due to population structure. The panel can also be used to match the ancestry of cases and controls prior to more comprehensive studies.

While the technology should provide a new tool in disease gene mapping studies, the researchers caution that the ability to roughly categorize individuals into populations with a small number of genetic markers is not useful in a clinical setting, nor does it completely eliminate the utility of self-described ethnicity. “Although these 300 markers give a reasonable estimate of the major components of genetic ancestry in European Americans, self-described ethnicity can still reflect environmental, social and cultural factors that may not be captured by estimating genetic ancestry,” says Dr. Joel Hirschhorn, one of the senior authors of the study, an Associate Professor of Genetics at Children's Hospital Boston and Harvard Medical School, and a Senior Associate Member at the Broad Institute of Harvard and MIT, “Because the genetic differences between these populations are very small, the study is most important for helping in gene discovery efforts, which will lead to better understanding of human biology in health and disease, and hopefully improved care for all patients over the long term.”

Published simultaneously in *PLoS Genetics* is an independent study led by Michael Seldin, in which Chao Tian and colleagues also present panels of markers that can be used to correct for population structure in European American disease association studies. A commentary jointly authored by Michael Seldin and Alkes Price on the practical application of the panels developed by the two groups accompanies these articles.

Source: Public Library of Science

Citation: Genetic diversity of European Americans and disease gene mapping (2008, January 18)

retrieved 25 April 2024 from

<https://phys.org/news/2008-01-genetic-diversity-european-americans-disease.html>

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.