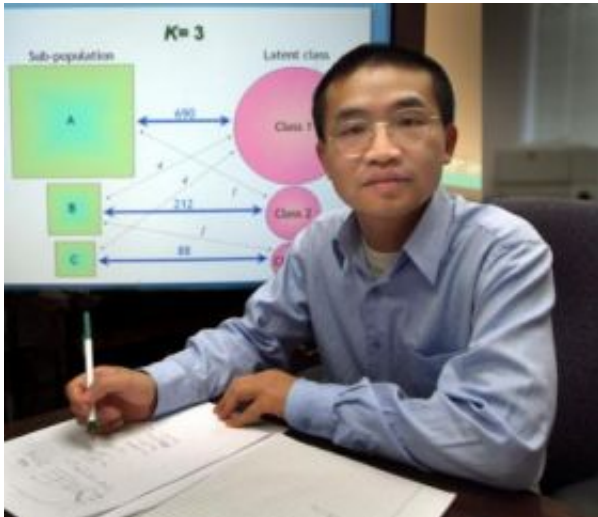


Statistical tool could explain gene study variations

28 May 2008



Dr. Hongyan Xu, a Medical College of Georgia biostatistician, is using statistics to explain variations in genome-wide association studies. Credit: Phil Jones, MCG campus photographer

While scientists are using the human genome to associate certain genes with disease, Dr. Hongyan Xu wants to ensure they are accounting for natural variations in those genes.

"These differences can create some challenges in analyzing data," says Dr. Xu, biostatistician in the Medical College of Georgia School of Graduate Studies. "There is always some difference in ethnic backgrounds across a study population."

For instance, a study looking at a population of blacks from Augusta and blacks from Chicago wouldn't necessarily take into account the difference in subpopulations, he says.

"Some groups of blacks could have different degrees of ancestry from different African groups," he says. "Some populations of blacks have different skin tones, which indicate a difference in genetic makeup. That isn't always taken into

account."

Scientists use genome-wide association studies to compare the genes of people with health conditions to the genes of healthy people, thereby better understanding basic biological processes that affect health and possibly how to better diagnose and treat disease.

Some studies account for differences by using control groups who self-report similar ethnicities. But there can be wide variations because people are not always completely aware of their ancestry, Dr. Xu says.

A computer-based statistical tool could be the answer, he says.

Dr. Xu and colleagues will start by examining an existing database from an ongoing association study of stroke risk in black children. That study, conducted by Dr. Abdullah Kutlar, hematologist/oncologist and director of the MCG Sickle Cell Center, aims to understand the genetics of stroke risk in children with sickle cell disease. With funding from the National Institutes of Health, Dr. Xu and his team will take a closer look at children already identified as high-risk because of high blood flow velocity in the brain, as measured by transcranial Doppler tests.

Previous MCG research identified high-velocity blood flow as a risk factor for stroke and regular blood transfusions as a way to reduce that risk.

"While Dr. Kutlar is looking for the underlying genetic reasons for the higher stroke risks in this sample of patients, we will be looking for ways to identify the subpopulations in that sample," Dr. Xu says. "If population structure isn't taken into account, it could affect the validity of study results."

Researchers will use a statistical approach known as coalescent theory, which traces coding

sequences of genes in a population sample to a single ancestral copy of a gene. That gene would theoretically be copied in the genetics of every member of an identical population.

For instance, two people with almost identical sets of chromosomes could differ in a very small way – by one structural unit that binds their DNA. By tracing it back, researchers would reach a point where the "copied" gene would not be present. That would indicate the point where two lineages joined, Dr. Xu says.

Genetic differences among the two populations could then be tagged, subcategorized and accounted for in study results, he says.

"With the coalescent theory, we focus on the samples rather than the whole population," Dr. Xu says. "That way, we can generate samples with various levels of population structure with great efficiency using computers, which are important for large-scale genome-wide studies. Understanding the genetic basis for disease is key to prevention, diagnosis and effective treatment. Developing a method that accounts for variations in the genetics of people who are similar but distinct is crucial to better understanding the genetics of health."

Source: Medical College of Georgia

APA citation: Statistical tool could explain gene study variations (2008, May 28) retrieved 8 December 2022 from <https://phys.org/news/2008-05-statistical-tool-gene-variations.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.