An Examination of Genome-Wide Association Studies Finds a European Bias

If you're a person of European ancestry from the United States, United Kingdom, or Iceland, congratulations! A recent study review published in Communications Biology examined 3,369 genome-wide association studies (GWAS) conducted from 2005 to 2018 and found that more than 80 percent of the people studied had a European background, and 72 percent of discoveries were made from studies examining individuals from those three countries. Seventy-six percent of the world's 7 billion people live in Asia and Africa, populations underexamined in genome studies. This lack of diversity limits the potential for discoveries and decreases the effectiveness of genetically targeted treatments. As the authors of this review conclude,

> *A central finding was that our results once again emphasized the potential for a cycle of disadvantage for underrepresented communities and despite continued efforts, infusing diversity into genomics remains challenging."*

So what are genome-wide association studies?

Where the Data Comes From

Why do genome studies lack diversity? This is due in large part to where the data is coming from. Many of these genome-wide studies rely on genetic material obtained from testing sites like 23andme. Those results are heavily skewed. Less than three years ago, researchers from Stanford estimated as many as 90 percent of the research into the genetics of

disease was based on people of European descent. Services like this also explain why Iceland, a country with less than 350,000 people, figures prominently into genetic research. The Icelandic company, deCODE Genetics, has been cataloging genetic information for more than 20 years, and nearly a third of the population has had at least part of their genome sequenced.

Other Influences

Funding also plays a role in genome study results. Of the more than 3,000 studies examined in this review, 85 percent of the funding acknowledgments referenced grants and other agencies in the U.S. Another 14 percent are based in the United Kingdom. This means the rest of the world is responsible for funding one percent of available genome-wide studies. Both the U.S. and the U.K. are predominantly composed of people with European ancestry (61 and 87 percent, respectively).

Another explanation for the overrepresentation of European ancestry in genome-wide studies is the people releasing them. Of the ten most connected and influential authors of genome studies, 9 are based in Europe. Three work for deCODE genetics in Iceland. Another three work for Erasmus Medical Center in the Netherlands, though that number would have been 4 prior to 2016. The only non-European author (based at Harvard University in the U.S.) also used to work at Erasmus. The work of these European men (and woman) is highly influential. Their articles and studies have been cumulatively cited over 200,000 times, and all of these scientists have strong European ties.

Echo Chambers

Genome-wide association studies are still a relatively new area of study (the first successful study was in 2002). These studies use a majority of data from those with European ancestry, receive funds from agencies in countries with a

majority of European ancestry, and are authored by white, mostly European scientists. It creates an echo chamber. That can limit, and in some cases, even hinder discovery.

The Oxford researchers responsible for this study call for more diversity, both ethnically and geographically.

> *GWAS that utilize data from diverse populations will provide more accurately targeted therapeutic treatments to more of the world's population, extend insights into the architecture of traits and uncover rare variants with significant effect sizes, which replicate across ancestries."*

Sources:

- *A scientometric review of genome-wide association studies* – Nature.com
- *23andMe has a problem when it comes to ancestry reports for people of color* – Quartz
- *Why Iceland is the World's Greatest Genetic Laboratory* – Wired