

# New tool facilitates inclusion of people of diverse ancestry in large genetics studies

February 2 2021

---



Credit: CC0 Public Domain

Genome-wide association studies (GWAS) have typically excluded diverse and minority individuals in the search for gene variants that confer risk of disease. Researchers at Massachusetts General Hospital (MGH), the Broad Institute of MIT and Harvard, and other institutions around the world have now developed a free-access software package called Tractor that increases the discovery power of genomics in

understudied populations. A study of Tractor's performance and accuracy was published in *Nature Genetics*.

Researchers perform GWAS to identify where genetic variants responsible for causing disease are located in the genome. Recently, geneticists have begun creating models from published GWAS data to predict risks of disease in individuals. But the clinical utility of these models is currently limited, since most are based on genomic studies of people with European [ancestry](#).

"If you build disease-risk models on available data and attempt to extrapolate them to diverse populations, the accuracy of predicting who will get sick is reduced," says Elizabeth Atkinson, Ph.D., lead author of the paper and an investigator in the Analytic and Translational Genetics Unit (ATGU) at MGH. "These errors exacerbate existing health disparities, in part because we aren't finding specific gene variants that may contribute to higher risk of a particular disease in diverse populations."

Another significant shortcoming of current GWAS is that "they leave many opportunities for genetic discovery on the table for all populations," says Atkinson. People of African descent, for example, have a million more genetic variations on average than someone who doesn't have African ancestry due to human migration patterns over the ages. Conducting a GWAS with diverse populations allows geneticists to pinpoint genetic associations to disease at many more spots across the genome, says Atkinson.

"Within these genomic regions identified in a GWAS, the genetic mutation that actually causes disease is shared across ancestries most of the time," she adds. By studying admixed populations—people with recent ancestry from two or more previously isolated [population](#) groups, such as Africa and Europe—"we can get more powerful and precise

genetic association signals and do a better job at pinpointing where the causal mutation is, which improves our understanding of disease for everyone."

Until now, there was no fine-scale way to control for ancestry composition in mixed groups being studied in a GWAS. "Different ancestry groups have gene variants that occur at different frequencies due to the populations' demographic history," explains Atkinson. "Not taking ancestry into account in a GWAS can lead to false-positive hits or to gene variants canceling themselves out and dismissed as not important. So, until now, it's been easier to exclude people with multiple ancestries from GWAS to avoid being confounded by different patterns of gene variants."

Tractor, however, allows researchers to account for ancestry in a precise manner so admixed individuals can be included in large-scale gene discovery efforts. The software colors pieces of each person's chromosomes according to its ancestry origin, which researchers can infer from reference genome sequences, and uses this information in a new GWAS model. "Tractor takes into account the ancestry backbone of each genetic variant so we can correctly calibrate the GWAS results to find causal variants in specific [population groups](#)," says Atkinson.

Tractor also provides estimates of ancestry-specific effect sizes, which isn't possible in a standard GWAS. "Instead of getting a weighted average of the disease-risk effect size for a particular [gene variant](#), Tractor can determine how large or small the effect of a variant is in various ancestry groups," says Atkinson. "This will be informative for building genetic risk scores in diverse populations." Another advantage of Tractor is its ability to improve the power of GWAS by detecting risk gene variants across multiple ancestries. "With Tractor, we can get stronger disease-association signals by leveraging ancestral genomic differences," says Atkinson.

"Tractor advances the existing methodologies for studying the genetics of complex disorders in diverse and minority populations," she adds.

"We hope that this method increases the inclusion of admixed participants in large-scale association studies going forward."

**More information:** Elizabeth G. Atkinson et al, Tractor uses local ancestry to enable the inclusion of admixed individuals in GWAS and to boost power, *Nature Genetics* (2021). [DOI: 10.1038/s41588-020-00766-y](https://doi.org/10.1038/s41588-020-00766-y)

Provided by Massachusetts General Hospital

Citation: New tool facilitates inclusion of people of diverse ancestry in large genetics studies (2021, February 2) retrieved 26 April 2023 from <https://medicalxpress.com/news/2021-02-tool-inclusion-people-diverse-ancestry.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.