Medical Xpress
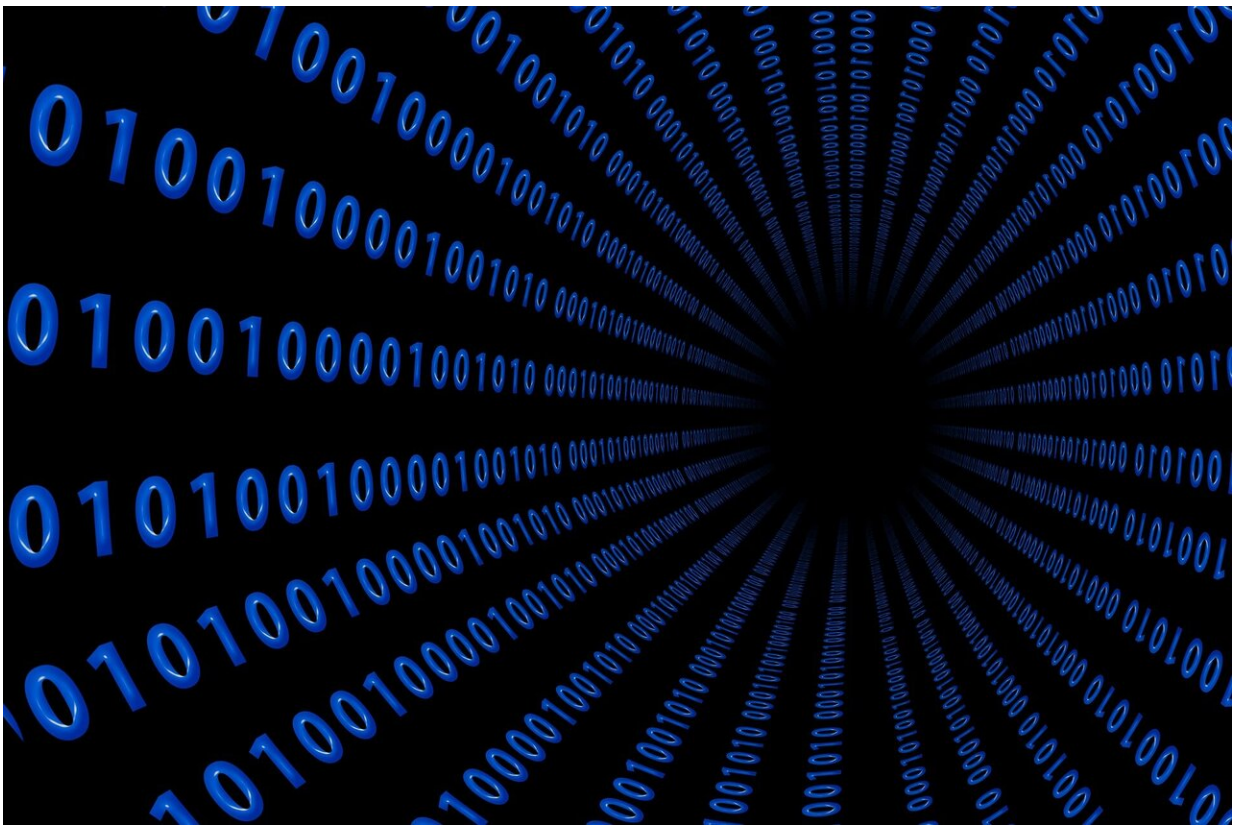
New method developed to detect and adjust population structure in genetic summary data

June 21 2021, by Meghan Azralon

Credit: Pixabay/CC0 Public Domain

In a new study published today in the *American Journal of Human Genetics*, researchers announced the development of a new method to

increase the utility and equity of large genetic databases. The research was conducted by Audrey Hendricks, an associate professor of statistics at the University of Colorado Denver (CU Denver).

Summix, the new method developed by Hendricks and her team of CU Denver undergraduate and graduate students, estimates the [genetic ancestry](#) in databases and adjusts the information to match the [ancestry](#) of a person or sample of people. This method leads large genetic databases to become more useful for people of various ancestries such as African American or Latinx, as they are underrepresented in genetic databases and studies. Hendricks compares this method to translating a book from English to another language.

"Think of DNA as the words of our body," says Hendricks. "All of the words of our body make the instruction book that makes each of us up. Right now, it's like the DNA books are only written in English so the information in the library is not as useful for people who don't speak English. We're working to create books in the library that are more universal."

According to Hendricks, individuals and samples from understudied populations, such as African American and Latnix, are the most likely to lack large public resources with precisely matched ancestry data. As a result, researchers working with those populations often resort to the closest, but still poorly matched ancestral group. This leads to biased results in the very populations where high-quality research is needed the most.

The team showed the effectiveness of Summix in over 5,000 simulation scenarios and in the widely used Genome Aggregation Database (gnomAD), a publicly available genetic resource. They found Summix's estimates of ancestry proportions to be highly accurate (within 0.001%) and the ancestry-adjusted genetic information to be less biased. The

Summix method is available in open access software increasing the utility of the method and its applications.

"Most people are a combination of multiple continental (e.g. African and European) or finer scale (e.g. Italian and German) ancestries," said Hendricks. "As healthcare moves forward with precision medicine, matching the unique ancestral makeup of each person will become increasingly important. The ability of Summix to update a genetic resource to match the ancestry of an individual is an important step in this direction and helps to increase the utility and equity of genetic summary data."

Provided by University of Colorado Denver