Many genes linked to alcohol and tobacco use are shared among diverse ancestries

December 7 2022, by Zach Sweger

Credit: Pixabay/CC0 Public Domain

Penn State College of Medicine researchers have co-led a large genetic study that identified more than 2,300 genes predicting alcohol and tobacco use after analyzing data from more than 3.4 million people.

They found that a majority of these genes were similar among people with European, African, American and Asian ancestries. The team's results were published in *Nature* on Dec 7.

Alcohol and tobacco use are associated with approximately 15% and 5% of deaths worldwide, respectively, and are linked with chronic conditions like cancer and heart disease. Although the environment and culture can affect a person's use and the likelihood of becoming addicted to these substances, genetics is also a contributing factor, according to the researchers. They helped identify around 400 genes that are associated with certain alcohol and tobacco use behaviors in people in a prior research study.

"We've now identified more than 1,900 additional genes that are associated with alcohol and tobacco use behaviors," said Dajiang Liu, professor and vice chair for research in the Department of Public Health Sciences. "A fifth of the samples used in our analysis were from non-European ancestries, which increases the relevance of these findings to a diverse population."

Collaborating with peers from the University of Minnesota and more than 100 other institutions, Liu and his team evaluated genetic datasets from more than 3.4 million people, at least 20% of whom were from non-European ancestries. According to Liu, their study is the largest genetic study on smoking and drinking behaviors to date, and is the most ancestrally diverse. He said his prior study in 2019 only included data from populations of European ancestry.

Liu and colleagues included genetic datasets from people of African, East Asian and American ancestries, and evaluated a variety of smoking and alcohol traits ranging from the initiation of drinking or smoking to the onset of regular use and the amount consumed. Using machine learning techniques, the researchers identified genes that were associated

with these behaviors.

Comparing the data between samples from different ancestries, the team found that there was a striking similarity in the genes related to alcohol and tobacco use behaviors between the different ancestries, with 80% of the variants showing consistent effects across the studied populations. While some genetic variants had different effects across ancestries or ancestry-specific effects, the genes associated with alcohol and tobacco use were largely consistent between samples from various ancestries.

The researchers used machine learning to develop a genetic risk score that could identify people at risk for certain alcohol and tobacco use behaviors. Despite the similarity of genetic effects, the model developed using data from individuals of European ancestry could only accurately predict alcohol and tobacco use behaviors for people of European ancestry. Since the model was not as accurate in predicting risk among people from other ancestries, Liu said there is a need to develop more sophisticated prediction methods by increasing sample sizes from non-European ancestries, which could improve risk prediction across diverse human populations.

"It is promising to see that the same genes are associated with addictive behaviors across ancestries," said Liu, a Penn State Cancer Institute and Penn State Huck Institutes of the Life Sciences researcher. "Having more robust and diverse data will help us develop predictive risk factor tools that can be applied to all populations."

Liu said that within two to three years, these genetic risk scores could be refined and become part of routine care for individuals already identified through basic screening to be at increased risk for alcohol and tobacco use. As interim director of the College of Medicine's second strategic plan goal, which seeks to develop and apply biomedical artificial intelligence, machine learning and informatics to make rapid

advances in biomedical research, he noted that this research is an example of how big data and sophisticated [machine learning](#) methods can help predict [health risks](#) so targeted interventions can be developed.

"This project leveraged large amounts of data to identify common genetic risk factors across diverse populations," said Kevin Black, MD, interim dean of the College of Medicine. "Using these findings to develop screening tools for diseases of despair is the kind of innovation that will help our college lead the way in using health informatics to contribute to health preservation and disease treatment in our communities."

According to Liu, future research will focus on diving deeper into their findings. A majority of the genes the team identified have unknown functions, so the researchers will try to understand their functions and how changes in those [genes](#), their function and interaction with the environment affect the risk for addictive behaviors. He also said increasing the diversity of genetic samples in the datasets will help the team develop predictive risk models for individuals from diverse ancestries.

More information: Gretchen R. B. Saunders et al, Genetic diversity fuels gene discovery for tobacco and alcohol use, *Nature* (2022). [DOI: 10.1038/s41586-022-05477-4](#)

Provided by Pennsylvania State University