# Epigenetic algorithm accurately predicts male sexual orientation

October 8 2015

An algorithm using epigenetic information from just nine regions of the human genome can predict the sexual orientation of males with up to 70 percent accuracy, according to research presented at the American Society of Human Genetics (ASHG) 2015 Annual Meeting in Baltimore.

"To our knowledge, this is the first example of a predictive model for sexual orientation based on molecular markers," said Tuck C. Ngun, PhD, first author on the study and a postdoctoral researcher at the David Geffen School of Medicine of the University of California, Los Angeles.

Beyond the genetic information contained in DNA, the researchers examined patterns of DNA methylation - a molecular modification to DNA that affects when and how strongly a gene is expressed - across the genome in pairs of identical male twins. While identical twins have exactly the same genetic sequence, environmental factors lead to differences in how their DNA is methylated. Thus, by studying twins, the researchers could control for genetic differences and tease out the effect of methylation. In all, the study involved 37 pairs of twins in which one twin was homosexual and the other was heterosexual, and 10 pairs in which both twins were homosexual.

"A challenge was that because we studied twins, their DNA methylation patterns were highly correlated," Dr. Ngun explained. Even after some initial analysis, the researchers were left with over 400,000 data points to sort through. "The high correlation and large data set made it difficult to identify differences between twins, determine which ones were relevant

to sexual orientation, and determine which of those could be used predictively," he added.

To sort through this data set, Dr. Ngun and his colleagues devised a machine learning algorithm called FuzzyForest. They found that methylation patterns in nine small regions, scattered across the genome, could be used to predict study participants' sexual orientation with 70 percent accuracy.

"Previous studies had identified broader regions of chromosomes that were involved in sexual orientation, but we were able to define these areas down to the base pair level with our approach," Dr. Ngun said. He noted that it will take additional research to explain how DNA methylation in those regions may be related to sexual orientation. The researchers are currently testing the algorithm's accuracy in a more general population of men.

"Sexual attraction is such a fundamental part of life, but it's not something we know a lot about at the genetic and molecular level. I hope that this research helps us understand ourselves better and why we are the way we are," Dr. Ngun said.

Provided by American Society of Human Genetics