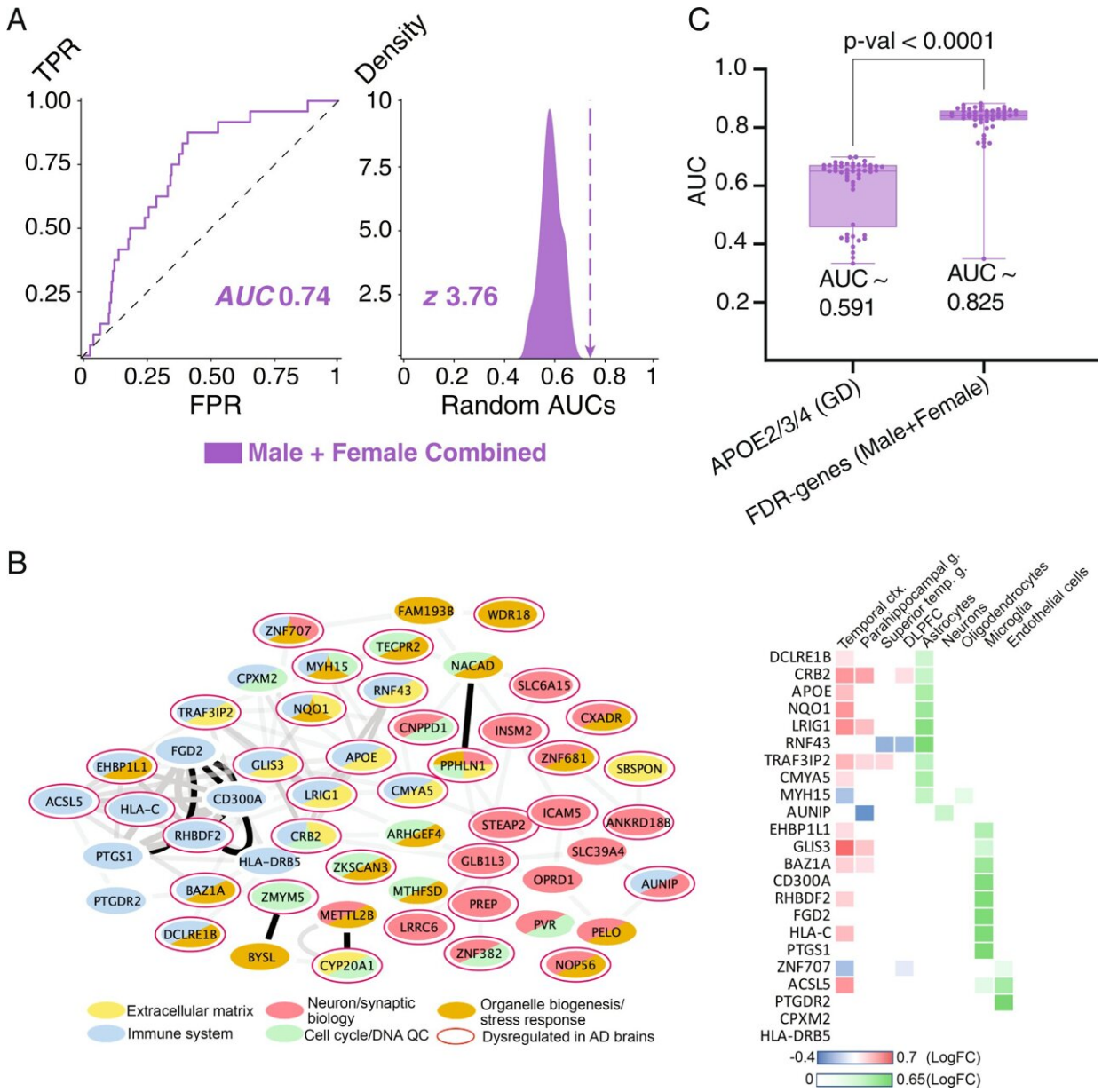
Machine-learning program reveals genes responsible for sex-specific differences in Alzheimer's disease progression

May 20 2023



The top 98 EAML genes are connected to GWAS genes, dysregulated in AD patients, and capable of separating AD and healthy control samples. Credit: *Nature Communications* (2023). DOI: 10.1038/s41467-023-38374-z

Alzheimer's disease (AD) is a complex neurodegenerative illness with genetic and environmental origins. Females experience faster cognitive decline and cerebral atrophy than males, while males have greater mortality rates. Using a new machine-learning method they developed called "Evolutionary Action Machine Learning (EAML)," researchers at Baylor College of Medicine and the Jan and Dan Duncan Neurological Research Institute (Duncan NRI) at Texas Children's Hospital have discovered sex-specific genes and molecular pathways that contribute to the development and progression of this condition. The study was published in *Nature Communications*.

"We have developed a unique machine-learning software that uses an advanced computational predictive metric called the evolutionary action (EA) score as a feature to identify genetic factors that influence AD risk separately in males and females," Dr. Olivier Lichtarge, MD, Ph.D., professor of biochemistry and molecular biology at Baylor College of Medicine, said. "This approach lets us exploit a massive amount of evolutionary data efficiently, so we can now probe with greater accuracy smaller cohorts and identify genes involved in sex-specific differences in AD."

EAML is an ensemble computational approach that includes nine machine learning algorithms to analyze the functional impact of non-synonymous coding variants, defined as DNA mutations that affect the structure and function of the resulting protein, and estimates their

deleterious effect on biological processes using the evolutionary action (EA) score.

Lichtarge and team used EAML to analyze coding variants in 2,729 AD patients and 2,441 control subjects to identify 98 genes that are associated with AD. These included several genes known to play a major role in AD biology which supported the general value of combining machine-learning approach with the phylogenetic evolutionary information embodied in EA to identify genes and pathways linked to a complex disease such as AD. They also showed that these genes made functional connections and discovered they were expressed abnormally in AD brains. Specific pathways involved mediated pathways for neuroinflammation, and microglial and astrocytic biology, consistent with their potential involvement in AD pathophysiology.

Next, they collaborated with Dr. Ismael Al-Ramahi, Dr. Juan Botas, and their teams at the Center for Alzheimer's and Neurodegenerative Diseases and Duncan NRI, to test the homologs of the 98 EAML candidate genes using two fruit fly models of AD. For this, they used a robot-assisted state-of-the-art behavioral testing platform, which allows for high-throughput screens *in vivo*. They found 36 genes modulated tau-induced degeneration and 29 genes modulated Aβ42-induced neurodegeneration. These included 9 genes able to ameliorate the neurodegeneration caused by both Tau and Aβ42, the two proteins known to accumulate in AD patients. This strongly validated the functional involvement of the identified candidates in mediating neurodegeneration *in vivo* and highlighted potential therapeutic avenues that could be gained by targeting these genes.

Since the goal of this study was to understand how AD manifests and progresses differently in males and females, they next applied EAML analysis separately to males and females within this cohort. They found 157 AD-associated genes in males and 127 in females. The genes

identified in this sex-separated study were found to be more closely connected to known AD GWAS genes than those identified in the combined sex studies. These findings suggest that sex-separated analysis increases the sensitivity of identifying AD-associated genes and improves risk prediction ability.

Moreover, they discovered that certain biological pathways may have a more significant impact on AD development for one sex than the other. For instance, female-specific EAML candidates were found to be involved in a module related to cell cycle control and DNA quality control. "We were excited to find a group of genes that were neuroprotective in females and that were linked to BRCA1, a gene known for its association with breast cancer. These findings suggest potential biological connections between AD and breast cancer, two diseases that are more frequent in females than males." Dr. Ismael Al-Ramahi said. These findings could have important implications for developing therapeutic strategies and in designing sex-stratified clinical trials for AD.

In addition, EAML retained its predictive capability with consistent and robust targets, even when the team tested it with smaller sample sizes. Even with just 700 samples, EAML could recover over 50% of the candidates found in the entire data set, which is significantly better than the predictive algorithms in use currently. The authors think this remarkably improved capability will enable researchers to use smaller data sets to arrive at accurate and reliable predictions, paving the way for incorporating sex-specific analyses to disease-gene association studies that may have not yielded reliable results using known methods.

"Our success in using EAML to find new targets for AD not only provides a fresh perspective on the genetic factors influencing this disorder but also underscores the importance of systematically applying sex-specific analyses when studying disease-gene associations," Dr. Juan

Botas, professor in the department of Molecular and Human Genetics at Baylor, added. "This innovative approach has the potential to revolutionize our understanding of complex diseases like AD and drive the development of personalized treatments tailored to each individual's genetic makeup."

Others involved in the study include Thomas Bourquard, Kwanghyuk Lee, Minh Pham, Dillon Shapiro,Yashwanth Lagisetty,Shirin Soleimani, Samantha Mota, Kevin Wilhelm,Maryam Samieinasab,Young Won Kim,Eunna Huh, Jennifer Asmussen,and Panagiotis Katsonis. They are affiliated with one or more of the following institutions: Baylor College of Medicine, Jan and Dan Duncan Neurological Research Institute at Texas Children's Hospital, UTHealth McGovern Medical School.

More information: Thomas Bourquard et al, Functional variants identify sex-specific genes and pathways in Alzheimer's Disease, *Nature Communications* (2023). DOI: 10.1038/s41467-023-38374-z

Provided by Texas Children's Hospital