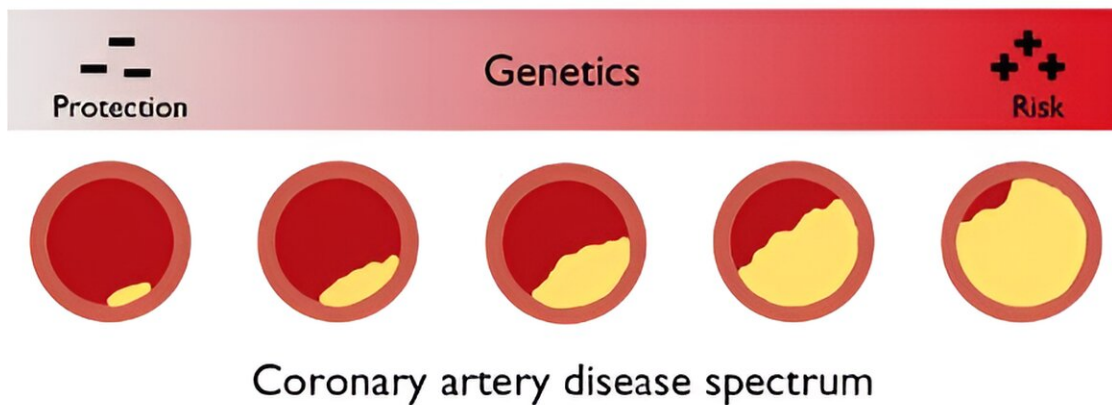


New AI tool finds rare variants linked to heart disease in 17 genes

June 11 2024



A rare variant association study using exome sequencing data was performed on a machine learning-based marker for coronary artery disease (CAD). This identified rare coding variants in 17 genes, revealing insights into the molecular basis of CAD. Credit: Ron Do, Ph.D., and Ben Omega Petrazzini, BS, at Icahn Mount Sinai.

Using an advanced artificial intelligence tool, researchers at the Icahn School of Medicine at Mount Sinai have identified rare coding variants in 17 genes that shed light on the molecular basis of coronary artery disease (CAD), the leading cause of morbidity and mortality worldwide.

The discoveries, detailed in *Nature Genetics*, reveal [genetic factors](#)

impacting heart disease that open new avenues for targeted treatments and personalized approaches to cardiovascular care.

The investigators used an in silico (computer-derived) score for [coronary artery disease](#) (ISCAD) that holistically represents CAD, as described in a [previous paper](#) by the team in *The Lancet*. The ISCAD score incorporates hundreds of different clinical features from the electronic health record, including [vital signs](#), laboratory test results, medications, symptoms, and diagnoses. To build the score, they trained machine learning models on the [electronic health records](#) of 604,914 individuals across the UK Biobank, All of Us Research Program, and BioMe Biobank in this comprehensive meta-analysis.

The score was then tested for association with rare and ultra-rare coding variants found in the exome sequences of these individuals. In addition, the research team conducted further investigation into the discovered [genes](#) to study their roles in causal CAD risk factors, clinical manifestations of CAD, and their connections with CAD status in traditional large-scale genome-wide association studies, among other factors.

"Our findings help us understand how these 17 genes are involved in coronary artery disease. Some of these genes are already known to influence heart disease development, while others have never been linked to it before," says Ron Do, Ph.D., senior study author and the Charles Bronfman Professor in Personalized Medicine at Icahn Mount Sinai. "Our study shows how machine learning tools can uncover genetic insights that traditional methods might miss when comparing cases and controls. This could lead to new ways to identify biological mechanisms of [heart disease](#) or gene targets for treatment."

Because they occur in only a small percentage of individuals, rare coding variants may have a significant impact on disease risk or susceptibility

when present. Therefore, studying these variants is essential to understanding the genetic basis of diseases and can inform therapeutic targets.

The study was driven by the challenges faced, over the last decade, in identifying rare coding variants associated with CAD using traditional methods relying on diagnosed cases and controls. Diagnostic codes' limitations in capturing the complexity of CAD prompted the researchers to explore new avenues of investigation.

"Our previous *Lancet* paper showed that a machine learning model trained with electronic health records can generate an in silico score for coronary artery disease, capturing disease across its spectrum," says lead author Ben Omega Petrazzini, Associate Bioinformatician in Dr. Do's lab at Icahn Mount Sinai. "Based on these findings, we hypothesized that the in-silico score for CAD could reveal novel rare coding variants related to CAD by offering a more holistic view of the disease."

Next, the investigators plan to further investigate the role of the identified genes in CAD biology and explore potential applications of machine learning in the genetic study of other complex diseases, as part of their ongoing efforts to advance understanding of disease mechanisms, discover new treatments, and improve patient outcomes.

The remaining authors of the paper, all with Icahn Mount Sinai except where indicated, are Iain S. Forrest, Ph.D. (MD candidate); Ghislain Rocheleau, Ph.D.; Ha My T. Vy, Ph.D.; Carla Márquez-Luna, Ph.D.; Áine Duffy, MS (Ph.D. candidate); Robert Chen, MS (MD/Ph.D. candidate); Joshua K. Park, BS (MD/Ph.D. candidate); Kyle Gibson, BA (MD candidate); Sascha N. Goonewardena, MD (University of Michigan); Waqas A. Malick, MD; Robert S. Rosenson, MD; and Daniel M. Jordan, Ph.D.

More information: Exome sequence analysis identifies rare coding variants associated with a machine learning-based marker for coronary artery disease, *Nature Genetics* (2024). [DOI: 10.1038/s41588-024-01791-x](https://doi.org/10.1038/s41588-024-01791-x)

Provided by The Mount Sinai Hospital

Citation: New AI tool finds rare variants linked to heart disease in 17 genes (2024, June 11) retrieved 18 June 2024 from <https://medicalxpress.com/news/2024-06-ai-tool-rare-variants-linked.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.