Researchers use 'blacklist' computing concept as novel way to streamline genetic analysis

December 24 2018

Credit: CC0 Public Domain

Researchers at the Icahn School of Medicine at Mount Sinai and The Rockefeller University have discovered a new use for a long-standing

computational concept known as "blacklisting," which is commonly employed as a form of access or spam control, blocking unwanted files and messages. Using blacklisting as a filter to single out genetic variations in patient genomes and exomes that do not cause illness, researchers have successfully streamlined the identification of genetic drivers of disease. This method is described in the December 2018 issue of *Proceedings of the National Academy of Science*.

In whole-exome sequencing—the process of identifying variations in protein-coding genes to determine the genetic underpinnings of any given illness—tens of thousands of genetic variants are identified, but only a few are deemed pathogenic, meaning disease-causing. Traditionally, in order to identify pathogenic mutations, scientists must sift through considerable amounts of data and remove genetic variants that are unlikely to cause disease, slowing down the process of genetic analysis and, subsequently, clinical treatment. To address this cumbersome process, researchers from the Icahn School of Medicine and The Rockefeller University investigated and subsequently identified a large portion of the non-pathogenic genetic variants, from which the "blacklist" was generated. Following this, they developed a program, known as ReFiNE, and a corresponding webserver that other researchers can use to automate the creation of their own blacklists.

"Until now, there has been no viable published method for filtering out non-pathogenic variants that are common in human genomes and absent from current genomic databases," said Yuval Itan, Ph.D., Assistant Professor of Genetics and Genomic Sciences at the Icahn School of Medicine and senior author of the publication. "Using the blacklist, researchers will now be able to remove genetic 'noise' and focus on true disease-causing mutations."

Noting the data-centric society we live in, Dr. Yuval says efficiency is key. His hope is that this contemporary tool can be used by clinicians,

researchers, and scientists across the globe to conduct genetic analysis more quickly and accurately, helping to accelerate the pace of genomic medicine.

More information: Patrick Maffucci el al., "Blacklisting variants common in private cohorts but not in public databases optimizes human exome analysis," *PNAS* (2018). www.pnas.org/cgi/doi/10.1073/pnas.1808403116

Provided by The Mount Sinai Hospital