Large genomic databases hold clues linking genetic mutations to future disease risk

September 12 2017, by Greta Friar

Credit: CC0 Public Domain

Physicians are increasingly using genetic testing to unravel mysterious or rare diseases in patients with baffling symptoms. But untangling mystifying symptoms is only half the battle. The next frontier in genomic analysis is predicting whether a genetic mutation is a harbinger of future disease.

Now scientists from Harvard Medical School have developed a way to

increase the predictive accuracy of genetic mutations as forerunners of diseases that may develop down the road. Using large-scale genomic analysis, the team has succeeded in illuminating patterns that can help predict symptoms likely to be caused by poorly understood disease-causing mutations and can help gauge the likelihood that a patient might develop clinical disease later on.

The researchers said their findings, published Aug. 16 in the journal *Cell Systems*, could help physicians better monitor future disease risk in patients with obvious symptoms affecting one organ but whose mutation may put them at risk for future disease affecting other organs and systems.

The analysis focused on rare genetic diseases because these tend to be easier to match to root-cause mutations, the researchers said. To pinpoint mutations common across multiple patients and link them to a specific disease, researchers typically must look at large numbers of sequenced genomes from people with the same diagnoses and then compare them against the genomes of people without such diagnoses. In the case of rare diseases—classified in the United States as those that affect fewer than 200,000 Americans—the pool for sequencing is limited to the genomes of only a handful of people.

On first blush, this may seem like a niche-value proposition, researchers note, but, in truth, the cumulative impact of characterizing the symptoms and future disease risk of many rare conditions could be profound. The National Institutes of Health estimate that as many as 30 million Americans have a rare disease.

"It can be hard to know where to look to diagnose and characterize a rare genetic disease," said study first author Ariel Feiglin, research fellow in biomedical informatics at HMS. "Our goal is to give physicians a better starting point for connecting genetic mutations to symptoms in order to

plan for their patient's future needs."

In the current study, researchers tested the hypothesis that genetic diseases that affect an organ likely stem from mutations in the genes most expressed in the tissue that makes up that very organ. For example, if a certain gene produces a protein that is most prevalent in the heart, a mutation to that gene would be the most likely culprit behind a patient's genetic heart disorder. The notion is, by itself not new, the researchers said, but previous studies have yielded conflicting findings. The HMS study results bring some much needed clarity, the researchers said.

In their analysis, the team compared relative expression of the same gene across different tissues as well as expression of different genes relative to each other within the same tissue.

The hypothesis panned out in some tissues but not in others. Mutations in genes highly expressed in the heart, brain, skin and muscle caused disease in these organs more often, the study showed. However, in breast, thyroid, and stomach tissue, the levels of gene expression were not linked to disease risk, a finding that suggests the existence of a more complex interplay in certain tissues.

"Our findings underscore the importance of using big data to tease out important patterns about gene expression—the activity of each gene in each tissue—which can improve our understanding of the meaning of a genetic mutation," said senior author Isaac Kohane, chair of the Department of biomedical informatics at Harvard Medical School.

The researchers said the new findings could give physicians clues about which organs to monitor in patients who come in with symptoms in one organ but turn out to have a mutation linked to multiple organs. For example, the mutation behind a well-known genetic muscle disease was recently found to cause neurodevelopmental disorders in a small

percentage of affected individuals. In this case, researchers said, awareness of the tell-tale high level of the relevant gene's expression in the brain could have clued in a physician to check for signs of neurodevelopmental issues in patients with the muscle disease.

"The beauty of these huge databases is that we can leverage existing data to make new discoveries," Feiglin said.

More information: Ariel Feiglin et al. Comprehensive Analysis of Tissue-wide Gene Expression and Phenotype Data Reveals Tissues Affected in Rare Genetic Disorders, *Cell Systems* (2017). DOI: 10.1016/j.cels.2017.06.016

Provided by Harvard Medical School