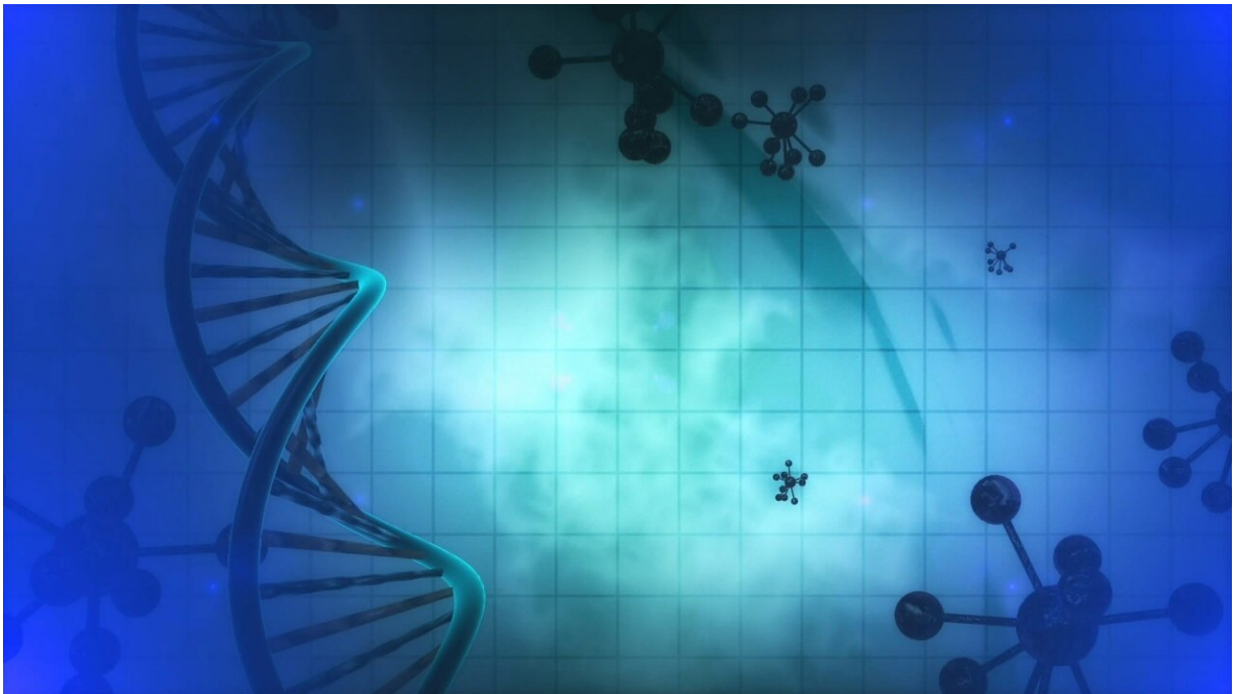# Research team identifies differences between benign and pathogenic gene variants

October 27 2020

Credit: CC0 Public Domain

An international team of researchers led by Cleveland Clinic's Lerner Research Institute has performed for the first time a wide-scale characterization of missense variants from 1,330 disease-associated genes. Published in *Proceedings of the National Academy of Sciences*, the study identifies features associated with pathogenic and benign variants that reveal the effects of the mutations at a molecular level.

"Our study serves as a powerful resource for the translation of personal genomics to personal diagnostics and precision medicine, and can aid variant interpretation, inform experiments and help accelerate personalized drug discovery," said Dennis Lal, Ph.D., assistant staff, Genomic Medicine, and the study's lead author.Recent large-scale DNA sequencing efforts have detected millions of missense variants, where mistakes in the DNA code change the amino acid (molecular building block of a protein) makeup of proteins. Some of these variants are pathogenic, meaning they alter the structure and function of a protein in a way that leads to disease, while others are benign with no impact on health. The vast majority, however, are considered variants of uncertain significance because their effects remain unknown.

While methods to predict variant pathogenicity exist, they do not elucidate why some variants are more or less likely to cause disease than others or establish their functional impact. Additionally, pathogenic and benign variants can co-exist in almost every disease-associated gene. As such, gaining a better understanding of the mechanistic differences between benign and pathogenic variants will be a critical next step in the development of novel therapies for genetic disorders.

Considering that a protein's function is closely linked to its three-dimensional structure, in this study the research team identified and compared the protein features of amino acids affected by pathogenic versus benign missense variants. Features that are more frequently mutated in pathogenic variants compared to benign variants (3-D mutational hotspots) are likely crucial to protein fitness and thus could help explain the molecular determinants of pathogenicity.

Looking at 1,330 disease-associated genes, the researchers analyzed a set of 40 features and found that 18 were significantly associated with pathogenic variants, 14 were significantly associated with benign variants and the remaining eight had no significant association with any

variant type.

"By considering genetic variation in the context of proteins' three-dimensional organization, we present for the first time an atlas of molecular properties of pathogenic mutations that addresses the differences between benign and disease-causing mutations," said Lal. "This study focused on 1,330 genes associated with rare types of genetic disorders, so we are currently extending our project to look at more genes and milder disorders."Data from this study (including precomputed P3DFiDAGS1330 and P3DFiProteinclass values for every possible amino acid exchange in proteins encoded by 1,330 disease-associated genes, along with the explicit listing of the 3-D features of the altered site as the rationale for the index) is available through the dedicated web server MISCAST.

More information: Sumaiya Iqbal et al, Comprehensive characterization of amino acid positions in protein structures reveals molecular effect of missense variants, *Proceedings of the National Academy of Sciences* (2020). DOI: 10.1073/pnas.2002660117

Provided by Cleveland Clinic