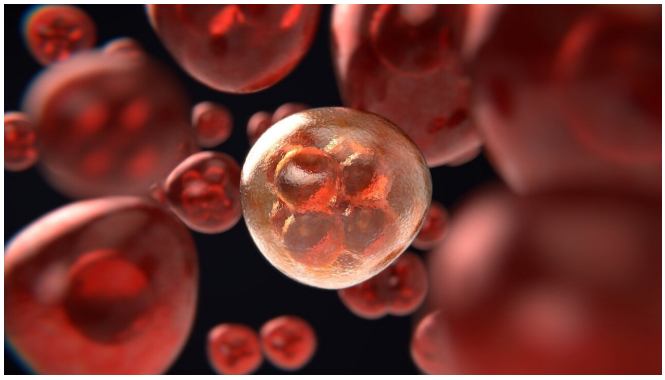
New model by CHOP researchers identifies noncoding mutations across five pediatric cancers

24 July 2020



Credit: Pixabay/CC0 Public Domain

Researchers at Children's Hospital of Philadelphia (CHOP) have developed a new computational algorithm that has, for the first time, identified a spectrum of mutations in the noncoding portion of the human genome across five major pediatric cancers. The study, which was published today in *Science Advances*, used the algorithm to analyze more than 500 pediatric cancer patients' mutations and gene expression profiles to develop a comprehensive list of potentially cancer-causing mutations.

"Noncoding mutations are very important because the noncoding portion of the genome typically regulates how genes are turned on and off, much like a control switch, which has implications for the uncontrolled growth that occurs in cancer," said Kai Tan, Ph.D., Professor of Pediatrics at CHOP and senior author of the study. "However, these regions are also challenging to study, and our knowledge about them not as developed as that of coding regions. Our computational model has identified a set of targets in pediatric cancers that we hope to study further and eventually move to clinical practice."

The researchers developed a computation tool called PANGEA (predictive analysis of noncoding genomic enhancer/promoter alterations) to analyze noncoding mutations and their impact on gene expression in more than 500 pediatric cancer patients who had five major types of pediatric cancer: B cell acute lymphoblastic leukemia (B-ALL), acute myeloid leukemia (AML), neuroblastoma, Wilms tumor, and osteosarcoma. PANGEA identified all types of mutations that are associated with gene expression changes, including single nucleotide variants, small indels, copy number variations, and structural variants.

Previous studies on noncoding mutations have focused on single nucleotide variants and small indels, which are insertions or deletions of bases in the genome that are relatively short in length. However, structural variants are regions of DNA much larger in size—1 kilobase or larger—a quality that makes them more difficult to identify but also more likely to contribute to changes in gene regulation that lead to cancer.

Using PANGEA, the researchers found that structural variants are indeed the most frequent cause of potentially cancer-causing mutations and identified 1,137 structural variants that affect the expression of more than 2,000 genes across the five pediatric cancer types.

In analyzing the data, the researchers found that coding and noncoding mutations affect distinct sets of genes and pathways, which is likely due to the different genomic locations of these two types of genes. The researchers found that genes involved in metabolism—the rewiring of which is a hallmark of cancer—are more frequently affected by noncoding mutations. However, it is unclear to what degree

noncoding mutations facilitate metabolism rewiring in the five cancer types the researchers studied.

"Our results highlight the need for comparative analysis of both coding and noncoding mutations, which might reveal novel cancer-related genes and pathways," said Tan. "Identifying putative mutations is a starting point that will facilitate experimental work to validates these predictions."

More information: Diverse noncoding mutations contribute to deregulation of cis-regulatory landscape in pediatric cancers," *Science Advances*, July 24, 2020. DOI: 10.1126/sciadv.aba3064

Provided by Children's Hospital of Philadelphia