

Big Data analysis identifies new cancer risk genes

July 10 2018



CRG researcher Solip Park developed a method to identify new cancer predisposition genes. Credit: CRG

Researchers at the Centre for Genomic Regulation (CRG) in Barcelona have developed a new method to identify genes contributing to heritable

cancer risk. Their work, which is published in *Nature Communications*, is a success story for data sharing and openness in science. The three researchers identified new cancer genes using only publicly available data.

There are many genetic causes of cancer. While some mutations are inherited, others are acquired throughout life due to external factors or mistakes in copying DNA. Large-scale genome sequencing has revolutionised the identification of cancers driven by such somatic mutations, but it has not been as effective in the identification of the inherited genetic variants that predispose to cancer. The main source for identifying these inherited mutations is still family studies.

Now, three researchers at the Centre for Genomic Regulation (CRG) in Barcelona, led by the ICREA Research Professor Ben Lehner, have developed a new statistical [method](#) to identify cancer predisposition genes from tumour sequencing data. "Our computational method uses an old idea that cancer genes often require 'two hits' before they cause cancer. We developed a method that allows us to systematically identify these genes from existing cancer genome datasets" explains Solip Park, first author of the study and Juan de la Cierva postdoctoral researcher at the CRG.

The method allows researchers to find risk variants without a control sample, meaning that they do not need to compare cancer patients to groups of healthy people, "Now we have a powerful tool to detect new cancer predisposition genes and, consequently, to contribute to improving cancer diagnosis and prevention in the future," adds Park.

The work, which is published in *Nature Communications*, presents their statistical method ALFRED and identifies 13 candidate cancer predisposition genes, of which 10 are new. "We applied our method to the genome sequences of more than 10,000 cancer patients with 30

different tumour types and identified known and new possible cancer predisposition genes that have the potential to contribute substantially to cancer risk," says Ben Lehner, principal investigator of the study.

"Our results show that the new cancer predisposition genes may have an important role in many types of cancer. For example, they were associated with 14 percent of ovarian tumours, 7 percent of breast tumours and to about one in 50 of all cancers. For example, inherited variants in one of the newly-proposed risk genes – NSD1 – may be implicated in at least three out of 1,000 [cancer patients](#)," explains Fran Supek, CRG alumnus and currently group leader of the Genome Data Science laboratory at the Institute for Research in Biomedicine (IRB Barcelona).

When sharing is key to advance knowledge

The researchers worked with genome data from several cancer studies from around the world, including The Cancer Genome Atlas (TCGA) project and also from several projects having nothing to do with cancer research. "We managed to develop and test a new method that hopefully will improve our understanding of cancer genomics and will contribute to cancer research, diagnostics and prevention just by using public data," says Solip Park.

Ben Lehner adds, "Our work highlights how important it is to share genomic data. It is a success story for how being open is far more efficient and has a multiplier effect. We combined data from many different projects and by applying a new computational method were able to identify important [cancer genes](#) that were not identified by the original studies. Many patient groups lobby for better sharing of genomic data because it is only by comparing data across hospitals, countries and diseases that we can obtain a deep understanding of many rare and common diseases. Unfortunately, many researchers still do not share

their data and this is something we need to actively change as a society."

More information: Solip Park et al. Systematic discovery of germline cancer predisposition genes through the identification of somatic second hits, *Nature Communications* (2018). [DOI: 10.1038/s41467-018-04900-7](https://doi.org/10.1038/s41467-018-04900-7)

Provided by Centre for Genomic Regulation

Citation: Big Data analysis identifies new cancer risk genes (2018, July 10) retrieved 20 April 2024 from <https://medicalxpress.com/news/2018-07-big-analysis-cancer-genes.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.