

Algorithm turns cancer gene discovery on its head

January 21 2020



An artificial intelligence system developed by KAUST scientists reveals genes that drive tumor growth. Here, Marwa Abdelhakim (left) and Sara Althubaiti are shown preparing the DNA library used for the study. Credit: 2020 KAUST

A method for finding genes that spur tumor growth takes advantage of

machine learning algorithms to sift through reams of molecular data collected from studies of cancer cell lines, mouse models and human patients.

By teaching the artificial intelligence system to link certain DNA mutations to altered functionality, a team led by Robert Hoehndorf from KAUST's Computational Bioscience Research Center showed that they could identify [genes](#) with a known causative role in [cancer](#) and pick out dozens of putative new ones for 20 different [tumor](#) types.

The [prediction method](#)—described in *Scientific Reports* and freely available online—could help clinicians tailor medicines to the molecular subtypes of patients. It could also be used by drug companies in the hunt for new therapeutic targets.

"Our method can be used as a framework to predict and validate cancer-driver genes in any database or real population sample," says Sara Althubaiti, a Ph.D. student in Hoehndorf's lab and the first author of the study.

Traditionally, scientists have approached the search for genes with a causal role in cancer by starting with DNA sequence data. By extensively cataloging tumor mutations shared among patients with a common type of cancer, the research community has documented hundreds of genes with a causal impact on tumor development. Experimental follow up is then used to functionally associate these genes with the hallmarks of cancer.



Robert Hoehndorf (left) and Sara Althubaiti discuss the performance of their algorithm. Credit: 2020 KAUST

"Our method turns this approach on its head," Althubaiti explains. "Essentially, our approach is knowledge-driven and we use tumor sequencing data as validation. This is unlike most approaches, which are data-driven combined with interpretation of the findings with respect to established knowledge."

The rate of discovery for new cancer-driving genes has been declining rapidly in recent years, leading the KAUST team to seek a new computational strategy. Instead of relying on sequence data, Althubaiti and Hoehndorf built a machine learning model that takes into account many biological features of genes and pathways involved in tumor

formation.

The researchers designed the algorithm to recognize functional and phenotypic patterns that predispose a gene toward playing a role in driving tumor development. They validated the model using a publicly available database of some 27,000 different tumor variants as well as functional and sequence data—showing that the algorithm could accurately categorize known cancer-driving genes and detect more than 100 other likely culprits, many with specific roles in particular tumor types.

The KAUST investigators then further tested the algorithm's performance on molecular data gathered from two cohorts of cancer patients. The first was from King Abdulaziz University Hospital in Saudi Arabia, comprising 26 tumor samples from individuals with a rare type of head and neck cancer called nasopharyngeal carcinoma. The other cohort comprised 114 colorectal cancer samples from patients treated at the University of Birmingham Hospital in the United Kingdom. In both patient groups, the model singled out candidate driver genes that were frequently mutated and shared pathogenic features of other cancer-causing genes.

Hoehndorf emphasizes the importance of the team effort involved. "This work is a good example for scientific collaboration within Saudi Arabia," he says, "but it also demonstrates the need for multidisciplinary collaborations between computer scientists, clinical researchers and biologists."

More information: Sara Althubaiti et al, Ontology-based prediction of cancer driver genes, *Scientific Reports* (2019). [DOI: 10.1038/s41598-019-53454-1](https://doi.org/10.1038/s41598-019-53454-1)

Provided by King Abdullah University of Science and Technology

Citation: Algorithm turns cancer gene discovery on its head (2020, January 21) retrieved 2 May 2024 from <https://medicalxpress.com/news/2020-01-algorithm-cancer-gene-discovery.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.