

# Scientists generate largest data set of cancer-related genetic variations

July 15 2013

---

Scientists at the National Cancer Institute (NCI) have generated a data set of cancer-specific genetic variations and are making these data available to the research community, according to a study published in *Cancer Research*, a journal of the American Association for Cancer Research.

This will help cancer researchers better understand [drug response](#) and resistance to cancer treatments.

"To date, this is the largest database worldwide, containing 6 billion data points that connect drugs with genomic variants for the whole [human genome](#) across cell lines from nine tissues of origin, including breast, ovary, prostate, colon, lung, kidney, brain, blood and skin," said Yves Pommier, M.D., Ph.D., chief of the Laboratory of Molecular Pharmacology at the NCI in Bethesda, Md., in an interview. "We are making this data set public for the greater community to use and analyze.

"Opening this extensive data set to researchers will expand our knowledge and understanding of tumorigenesis [the process by which normal cells are transformed into cancer], as more and more cancer-related gene aberrations are discovered," Pommier added. "This comes at a great time, because genomic medicine is becoming a reality, and I am very hopeful this valuable information will change the way we use drugs for precision medicine."

Pommier and colleagues conducted whole-exome sequencing of the

NCI-60 [human cancer](#) cell line panel, which is a collection of 60 human cancer cell lines, and generated a comprehensive list of cancer-specific genetic variations. Preliminary studies conducted by the researchers indicate that the extensive data set has the potential to dramatically enhance understanding of the relationships between specific cancer-related genetic variations and drug response, which will accelerate the drug development process.

The NCI-60 human cancer cell line panel is used extensively by cancer researchers to discover novel anti-[cancer drugs](#). To conduct whole-exome sequencing, Pommier and his NCI team extracted DNA from the 60 different cell lines, which represent cancers of the lung, colon, brain, ovary, breast, prostate and kidney, as well as leukemia and melanoma, and cataloged the genetic coding variants for the entire human genome. The genetic variations identified were of two types: type I variants corresponding to variants found in the normal population, and type II variants, which are cancer-specific.

The researchers then used the Super Learner algorithm to predict the sensitivity of cells harboring type II variants to 103 anti-cancer drugs approved by the FDA and an additional 207 investigational new drugs. They were able to study the correlations between key cancer-related genes and clinically relevant anti-cancer drugs, and predict the outcome.

The data generated in this study provide means to identify new determinants of response and mechanisms of resistance to drugs, and offer opportunities to target genomic defects and overcome acquired resistance, according to Pommier. To enable this, the researchers are making these data available to all researchers via two database portals, called the CellMiner database and the Ingenuity systems database.

Provided by American Association for Cancer Research

Citation: Scientists generate largest data set of cancer-related genetic variations (2013, July 15)  
retrieved 4 May 2024 from

<https://medicalxpress.com/news/2013-07-scientists-largest-cancer-related-genetic-variations.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.