

# Team develops big data analysis platform to unveil gene interactions in cancer

June 23 2015

---



The breakthrough has been made by a research team consists of Professor Benjamin Yung (centre), Dr. Lawrence Chan (left) and Dr. Cesar Wong (right) from HTI at PolyU and biostatisticians from Harvard University. Credit: Hong Kong Polytechnic University

The Hong Kong Polytechnic University (PolyU) has achieved a

breakthrough in the cancer genomics by developing a novel big data analysis platform for analyzing the interactions among genes. The analysis platform unveils the unregulated patterns of gene network in cancer and discovers potential diagnostic and therapeutic target genes, Nucleophosmin (NPM1) and its associated genes, in Chronic Myelogenous Leukemia (CML).

These discoveries help the future development of NPM1-oriented treatment strategy in CML and enable targeted therapy to be more targeted. The analysis platform can also be readily applied to other diseases. The breakthrough has been made by a research team led by chair professor Benjamin Yung, which consists of biomedical scientists, Dr. Lawrence Chan and Dr. Cesar Wong from Department of Health Technology and Informatics (HTI) at PolyU and biostatisticians from Harvard University.

Traditional way to identify cancer gene markers is to examine genes in the whole human genome one-by-one with respect to their individual expressional differences between cancer patients and healthy people. Such deep-rooted method cannot provide any information about the interactions among genes and their systemic changes that compose the complicated cancer mechanism.

The research team hypothesized that genes participating in the same mechanism are co-expressed (i.e. the expression levels of gene pairs are correlated), and explored pair-to-pair correlations between genes so as to decipher the mechanism underlying the cancer. Combinations of gene pairs increase geometrically with the number of genes of interest. As the human genome has more than 20,000 protein-coding [genes](#), the co-expression analysis could involve about 0.2 billion possible gene pairs. Moreover, the statistical test for co-expression analysis has to be refined to maintain the reliability of the detection results. Biostatistics experts of Harvard University helped the team perform a large-scale computer

simulation to establish the statistical method for analyzing cancer-specific difference in co-expression. The team overcame challenges in processing massive data, and accomplished a performance of analyzing 0.2 billion gene pairs in two days, a mission impossible by the traditional method. The team has built an unprecedented scientific foundation of co-expression analysis, paving the way to the future of translational medicine.

This study has assured a novel structural co-expression network analysis platform, which unveils cancer pathogenesis and its potential NPM1-oriented treatment strategy. Co-expression analysis discovers novel unregulated patterns of [gene network](#) for understanding [cancer](#) biology, identifying new targets for treatment and all these innovations contribute to great science. This platform can be readily applied to other diseases for diagnostic, prognostic and therapeutic investigation. This finding will be published shortly in *Scientific Reports* from the publishers of *Nature*, showcasing the success of scientific research as a result of cross disciplinary collaboration.

Provided by Hong Kong Polytechnic University

Citation: Team develops big data analysis platform to unveil gene interactions in cancer (2015, June 23) retrieved 24 May 2024 from <https://medicalxpress.com/news/2015-06-team-big-analysis-platform-unveil.html>

<p>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.</p>
--