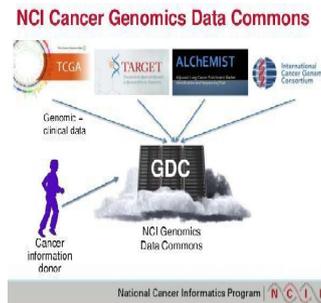


## University of Chicago to Establish Genomic Data Commons



The University of Chicago, in collaboration with the National Cancer Institute, will set up the country's most comprehensive computational facility that stores and harmonises cancer genomic data generated through NCI-funded research projects.

The establishment of the NCI Genomic Data Commons (GDC) will help scientists to speed up research, leading to faster discoveries for patients. The GDC will provide an interactive system for researchers, making the data easier to use. It also will provide resources to facilitate the identification of subtypes of cancer as well as potential therapeutic targets.

“The Genomic Data Commons has the potential to transform the study of cancer at all scales,” said Robert Grossman, PhD, director of the GDC project and professor in the Department of Medicine at the University of Chicago. “It supplies the data so that any researcher can test their ideas, from comprehensive ‘big-data’ studies to genetic comparisons of individual tumours to identify the best potential therapies for a single patient.”

A number of NCI-funded research programmes have collected genomic data on tumour types from more than 10,000 patients. However, the data for these studies are scattered across different locations and are in different formats, making it difficult for researchers to perform analyses. An Institute of Medicine report has cited the importance of having a system to store, harmonise and analyse existing cancer genomics data, which currently amount to roughly 20 petabytes of information (or 10 times as much as all of the publications currently housed in US academic research libraries).

The GDC will centralise existing NCI datasets through an approach to data storage and analysis similar to what is used by companies such as Google and Facebook. It will streamline access to data for researchers regardless of their institution's size or budget, thus effectively democratising access to the material. In addition, the GDC will enable previously unfeasible collaborative efforts between scientists.

The GDC will be built over a number of years to ensure individual projects can be combined to create broadly useful and accessible datasets and to inform guidelines for social, ethical, and legal issues that could arise as datasets become widely shared.

“With the GDC, the pace of discovery shifts from slow and sequential to fast and parallel,” said Conrad Gilliam, PhD, dean for basic science at the University of Chicago Biological Sciences Division. “Discovery processes that today would require many years, millions of dollars, and the coordination of multiple research teams could literally be performed in days, or even hours.”

The GDC is seen as a key step towards the development of precision medicine — targeted therapies that are tailored to individual patients. Once fully developed, the GDC will provide an interactive system for clinicians and researchers to upload their cancer genomics data and use it to identify the molecular subtype of cancer and potential therapeutic targets. Genetic data will be linked to extensive clinical information from patients and their responses to treatment.

“The availability of high-quality genomic data and associated clinical annotations is extremely important because this information can be combined and mined repeatedly to make new discoveries,” said Louis Staudt, PhD, MD, director of NCI's Center for Cancer Genomics.

The GDC also creates a foundation for future cloud-based technologies that one day will allow researchers to study large-scale datasets and perform experiments remotely. The open-source software being developed by the GDC could be a model for data-intensive research efforts for other diseases such as Alzheimer's and diabetes, which also need similar large-scale, data-driven approaches to develop cures.

The GDC will be constructed and operated with NCI funding through a subcontract from Leidos Biomedical Research, Inc. at the Frederick National Laboratory for Cancer Research. Some of the GDC's components are being developed by the Ontario Institute of Cancer Research through a subcontract with the University of Chicago.

Source: Newswise

Image Credit: National Cancer Institute

Published on : Wed, 10 Dec 2014