Abstracts

A.45

Proteomic Data Commons (PDC): A Next-Generation Proteomics Data Repository

Xu Zhang1, Ratna Thangudu2, Paul Rudnick3, Michael Holck2, Michael MacCoss4, Nathan Edwards5, Peter McHarvey5, Karen Ketchum2, Sudha Venkatachalan6, John Otridge6, Emi Casas-Silva1, Erika Kim1, Henry Rodriguez1

1National Cancer Institute, NIH, Rockville, MD, 2ICF International, INC., Rockville, MD, 3Spectragen Informatics LLC, Bainbridge Island, WA, 4University of Washington, Seattle, WA, 5Georgetown University, Washington D.C., 6Frederick National Laboratory, Frederick, MD

Proteomics has become an important field in cancer research, as it provides valuable information on the identity, expression levels, and modifications of proteins. Cancer proteomics unraveled key information in molecular mechanisms on tumor growth and metastasis, which has contributed to the identification of diagnostic protein biomarkers as well as therapeutic targets. Several cancer proteome databases have been established and shared worldwide. However, the data were extracted from multiple publications, the process could be very different across the studies and the results may not be comparable. The National Cancer Institute’s (NCI’s) Clinical Proteomic Tumor Analysis Consortium (CPTAC) is a national effort to accelerate the understanding of the molecular basis of cancer through the application of large-scale proteogenomics. The Proteomic Data Commons (PDC) serves as the central repository for the proteomics data and distributes it to physicians, clinicians, and scientists in the cancer research community. It is the largest cancer proteomic data warehouse in the world.

The PDC provides a unified data repository that enables data sharing across cancer proteomic studies and enables multi-omic integration in support of precision medicine. As a domain-specific repository within the Cancer Research Data Commons (CRDC), the vision for the PDC is to provide researchers the ability to find and analyze proteomic data across a wide variety of tumor types. Currently, the PDC houses data, supported by a large collection of metadata attributes, for nearly 100 studies from over 12 cancer types produced by several large-scale cancer research programs, each with cohort sizes greater than 100 patients. The data in the PDC are structured and queryable using the PDC data model and data dictionary. All the data in PDC are harmonized through a Common Data Analysis Pipeline (CDAP), removing the data analysis variable, and enabling comparisons across datasets. The PDC provides an intuitive interface for researchers to search, visualize, and analyze protein expression across diverse studies. Other features include building pan-cancer cohorts using highly curated, clinical metadata, and comprehensively viewing a study without downloading the data. Through an application programming interface (API), PDC facilitates interoperability with NCI Cloud Resources for efficient access to cloud computation.

A.46

Profiling the cellular signalling state by targeted phosphoproteomics

Mirjam van Bentum1, Henrik Zauber1, Bertram Klinger2, Mohamed Haji3, Philipp Mertins1, Nils Blüthgen5, Matthias Selbach1

1Max Delbrück Center (MDC), Berlin, Germany, 2Charité – Universitätsmedizin Berlin, Berlin, Germany

Normal cell functioning is regulated by a complex signalling network that responds to various extra- and intracellular signals like growth factors, cytokines and nutrients. Aberrations in cellular signalling often lead to disease, for example cancer. Understanding the functioning of the cellular signalling network, as well as changes in the network between healthy and diseased cells is a key challenge in biology research.

We have developed a targeted phospho-proteomic assay, that aims to provide a high-throughput, quantitative, reproducible profile of the cellular signalling state. The assay enables sensitive measurement of the phosphoproteins that give a comprehensive overview of the cellular signalling state. The panel includes 443 phospho-peptides mapping to 249 proteins in all major signalling pathways. The phospho-peptides are measured in a targeted fashion, using internal standard triggered PRM (Surequant).

The data is analysed with a newly developed in-house tool: Vali. This tool enables semi automatic analysis of PRM data and is optimised to work directly with Picky generated spectral libraries. We confirmed sensitivity and accuracy of quantification of the PRM-Vali pipeline using a SILAC dilution series. We observed linearity in quantitation down to 63-127 fold ratios using the described setup in a pure automated fashion.

Performance of the assay was determined by measuring perturbations of HCT-116 colon cancer cells with growth factors and cytokines. We were able to validate previously identified signatures of these perturbations (for example, increased phosphorylation of Met Y1234 when treated with HGF, EGFR Y1172 when treated with EGF and Jun S63 and S73 when treated with TNFalpha). The assay is furthermore applied to a colon cancer cell line panel to uncover previously unknown synergistic feedback mechanisms.

100321, https://doi.org/10.1016/j.mcpro.2022.100464

100321, https://doi.org/10.1016/j.mcpro.2022.100463