Clinical Annotation in High-throughput Data is Key to Progress of Translational Research in Hepatocellular Carcinoma

Mamatha Bhat1,2,3*, Elisa Pasini2, Chiara Pastrello3 and Igor Jurisica4,5

1Division of Gastroenterology, University Health Network, Canada
2Multi-Organ Transplant Program, University Health Network, Canada
3Krembil Research Institute, University Health Network, Canada
4Departments of Medical Biophysics and Computer Science, University of Toronto, Canada
5Institute of Neuroimmunology, Slovak Academy of Sciences, Canada

Letter to the Editor

Hepato Cellular Carcinoma (HCC) is a high fatality cancer that develops in the context of chronic liver disease, with increasing incidence worldwide. It is the second most common cancer in the world, with an estimated 788,000 deaths annually [1]. Five-year survival is the second worst worldwide among all cancers, at only 8.9% due to late diagnosis and inability to tolerate chemotherapy in patients with cirrhosis [2,3]. HCC is a multi-genic cancer occurring in the context of chronic liver disease, thereby rendering therapeutic targeting of this cancer difficult. The significant heterogeneity of HCC was highlighted by The Cancer Genome Atlas (TCGA) data on HCC [4].

In order to elucidate the key dys regulated genes and pathways in HCC, we recently undertook a comprehensive and systematic curation of publicly available gene expression data from patient HCC samples. We particularly aimed to identify genes characteristic of specific etiologies of liver disease (Hepatitis B, Hepatitis C, alcoholic liver disease and fatty liver disease). All available high-throughput microarray gene expression datasets related to HCC patient samples since 2002, which represents the start of high-throughput profiling, were considered for inclusion. Data were identified and downloaded using PubMed and Gene Expression Omnibus.

We identified 5,40 abstracts and created a curated database of the liver cancer gene signatures and deregulated genes from the 36 eligible datasets in a Cancer Data Integration Portal (CDIP), publicly available at http://ophid.utoronto.ca/CDIPLiver. However, there is still substantial lack of key demographic and clinical characteristics accompanying molecular profiles in HCC, as highlighted in (Figure 1A), which hinders translational research. Unfortunately, this limitation to translational research did not improve over the years (Figure 1B).

Awareness regarding the age, sex, underlying etiology of cirrhosis, Child-Pugh or MELD score as indices of liver function, grade and stage of tumor, and patient performance status are key pieces of information that influence clinical management [5]. Tumor-specific characteristics such as the
size, number, and presence of micro- or macro vascular invasion are important to consider alongside the interpretation of gene expression data. Conversely, gene expression data could influence the treatment approach adopted and overall survival. Most importantly, these various pieces of information are essential for progress to be made in treating this high-fatality cancer. We observed that most gene expression data has been obtained from HCC arising in the context of Hepatitis B and Hepatitis C. There was limited gene expression data from HCC arising in fatty liver disease, despite its dramatic increase as a dominant etiology worldwide with the rise in obesity and diabetes.

Although it is difficult to assign correlation, malignancies such as breast cancer have seen their mortality decrease with increased resource allocation and data generation over time [6,7]. This directly relates to genomic medicine – we need annotated datasets to build and validate prognostic and predictive signatures. This holds particularly true for a cancer that arises in the background of various chronic liver diseases. Therefore, there is a dire need to improve clinical annotation of the reams of high-throughput data being generated in HCC, in order to help research progress in better understanding the molecular pathogenesis of this highly fatal cancer.

References

1. Key facts cancer.