



Understanding the Integrated Gene Regulatory Networks for Hepatocellular Carcinoma

Long Gao*

Department of Genetics, University of Pennsylvania, USA

***Corresponding author:** Long Gao, Department of Genetics and Institute for Diabetes, Obesity, and Metabolism, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.

Received Date: July 15, 2018

Published Date: July 27, 2018

Editorial

Hepatic cancer is a malignant tumor that begins in the cells of the liver. The leading cause is a viral infection with hepatitis B and hepatitis C. Hepatocellular carcinoma (HCC) has become the most common form of hepatic cancer. It is the fastest-growing cause of cancer deaths in the United States. HCC is also found to be associated with obesity, type 2 diabetes and fatty liver disease. To understand the gene expression regulation during HCC development, scientists have identified a few related transcription factors (TFs) and characterize their roles such as E2F1 [1], Foxm1b [2] and hepatic nuclear factors [3]. However, these findings still cannot fully explain the underlying molecular mechanism during liver tumorigenesis. It is necessary to systematically study the global gene regulatory network (GRN) during HCC development.

The advancement of next-generation sequencing (NGS) technologies has enabled the rapid examination of the entire human genome. Many NGS applications have been developed to profile cells such as RNA-seq [4], ChIP-seq [5], ATAC-seq [6] etc. These allow scientists to study cells at different levels including genome, transcriptome, and epigenome. For instance, ChIP-seq and ATAC-seq data can be used to learn the chromatin states of certain biological processes by detecting regulatory elements in the genome and corresponding transcriptional regulators. Furthermore, recent developments in high-throughput single-cell technology provide the statistical power to study diverse population of tumor cells. This can greatly help scientists to understand intratumoral heterogeneity [7].

To take advantage of these high-throughput sequencing data, people from different fields have developed various sophisticated computational methods to reconstruct GRNs. A GRN is a collection of TFs that interact with their targets in the cell to govern gene expression. TFs are proteins that are involved in transcribing DNA to RNA by binding to specific DNA sequences such as enhancers and promoters. The interactions between TFs and their targets

can be detected via various technologies including electrophoretic mobility shift assay, DNase footprinting assay, chromatin immunoprecipitation and yeast one-hybrid system. GRNs can also be computationally reconstructed through reverse engineering methods using gene expression profiling data. GRN serves as a “blueprint” of TF-target interactions, which can be used to generate novel biological hypotheses. An important aspect of this application is that GRNs represent statistically significant predictions of TF-target interactions obtained from high throughput datasets. In this way, GRNs enable us to study regulatory relationship among genes and understand the underlying cellular processes in living cells. Gene expression data profiles generated from HCC patients can be used to construct GRN. Once corresponding GRNs are inferred, gene relationship can be understood using various network-based algorithms. For instance, hub genes in the GRNs are likely to play important roles in HCC development.

In addition to TF-target interactions, integrated GRNs can be also constructed with other types of interactions to obtain more biological insights. For instance, long noncoding RNAs (lncRNAs) and miRNAs can be integrated in GRNs using their expression profiles or genomic information. A few lncRNAs have been reported to be directly associated with tumorigenesis and cancer metastasis including H19, HOTAIR and HULC [8]. These lncRNAs have potential to contribute to activating WNT and TGF- β signaling pathways. For miRNAs, they are also likely to regulate key cancer-related pathways such as cell cycle control and DNA damage response [9]. Furthermore, mutational information and hyper-methylation of gene promoter regions identified in HCC patients can be incorporated into integrated networks to understand how upstream aberrant events cause downstream gene expression change [10]. By integrating various types of information, reconstructed GRNs will be more robust to reveal key transcriptional regulators for HCC study.

With constructed GRNs, network-based algorithms can be used to identify key TFs, mutations, hyper-methylated regions with unique topological features. There are a few commonly used topological features such as degree, centrality and cluster coefficient. All of them denote different types of importance of nodes (genes, mutations etc.) in given networks. Usually, nodes with large topological values tend to play a more important role in tumorigenesis. Besides traditional network algorithms, machine learning based approaches can be also used to predict key transcriptional regulators. The basic idea of this machine learning approach is to do prediction with prior known information. First collect a list of mutations, TFs, genes, lncRNAs that are known to be associated with HCC. Next, find their topological features in corresponding integrated GRNs and train a classifier. Lastly, predict novel regulators using this trained classifier [11].

References

1. Farra R, Grassi G, Tonon F, Abrami M, Grassi M, et al. (2017) The Role of the Transcription Factor E2F1 in Hepatocellular Carcinoma. *Curr Drug Deliv* 14(2): 272-281.
2. Kalinichenko VV, Major ML, Wang X, Petrovic V, Kuechle J, et al. (2004) Foxm1b transcription factor is essential for development of hepatocellular carcinomas and is negatively regulated by the p19ARF tumor suppressor. *Genes Dev* 18(7): 830-50.
3. Hayashi Y, Wang W, Ninomiya T, Nagano H, Ohta K, et al. (1999) Liver enriched transcription factors and differentiation of hepatocellular carcinoma. *Mol Pathol* 52(1): 19-24.
4. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10(1): 57-63.
5. Raha, D, Hong M, Snyder M (2010) ChIP-Seq: a method for global identification of regulatory elements in the genome. *Curr Protoc Mol Biol* Chapter 21: Unit 21.19.1-14.
6. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ (2015) ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* 109: 21.29.1-9.
7. Levitin HM, Yuan J, Sims PA (2018) Single-Cell Transcriptomic Analysis of Tumor Heterogeneity. *Trends in cancer* 4(4): 264-268.
8. Bhan A, Soleimani M, Mandal SS (2017) Long Noncoding RNA and Cancer: A New Paradigm. *Send to Cancer Res* 77: (15) 3965-3981.
9. Jansson MD, Lund AH (2012) MicroRNA and cancer. *Mol Oncol* 6(6): 590-610.
10. Wajed SA, Laird PW, DeMeester TR (2001) DNA methylation: an alternative pathway to cancer. *Ann Surg* 234(1): 10-20.
11. Long Gao, Yasin Uzun, Peng Gao, Kai Tan, Bing He, et al. (2018) Identifying noncoding risk variants using disease-relevant gene regulatory networks. *Nature Communications* 9: 702.