Commentary

Profiling HBV integrations in hepatocellular carcinoma

Shaogui Wan[1], Jesse Civan[2], Simona Rossi[2], Hushan Yang[1]

[1]Division of Population Science, Department of Medical Oncology, Kimmel Cancer Center; [2]Division of Gastroenterology and Hepatology, Department of Medicine, Thomas Jefferson University, Philadelphia, PA 19107, USA

Corresponding to: Hushan Yang, PhD. Division of Population Science, Department of Medical Oncology, Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA 19107, USA. Email: hushan.yang@jefferson.edu.

Submitted Sep 27, 2012. Accepted for publication Oct 30, 2012.
doi: 10.3978/j.issn.2304-3881.2012.10.13
Scan to your mobile device or view this article at: http://www.thehbsn.org/article/view/1169/1856

Hepatocellular carcinoma (HCC) is the third leading cause of cancer death worldwide. In this context, chronic viral hepatitis B (HBV) infection represents the most common etiology of HCC. Notably, although other common causes of HCC including chronic viral hepatitis C and chronic alcoholic liver disease are mediated by progression through cirrhosis, the pathogenesis of HCC in HBV infection does not entirely depend on this mechanism. A major proposed pathway by which HCC may arise from chronic HBV infection is the integration of HBV DNA into the host genome, resulting in oncogene activation, tumor-suppressor gene inactivation, or other predisposition to chromosomal instability (1).

Since the 1980's, many studies have detected evidence of HBV DNA integrated not only into normal hepatocytes, but specifically into HCC cells, supporting the hypothesized role of HBV DNA integration in hepatocarcinogenesis. However, early studies observed viral DNA integration in a non-recurrent pattern, meaning that the integration was not found to occur at predictable and reproducible sites in the host genome. It is important to note that these initial studies were limited by technological constraints to the analysis of only certain portions of the host genome, and also constrained by small sample size.

Over the past several years, the new technology of next-generation sequencing (NGS) has resulted in an explosion of studies in which somatic aberrations in cancer cells have been explored with much greater efficiency and accuracy (2). Specifically, NGS has been increasingly used as a powerful means to help pinpoint the genomic events implicated in hepatocarcinogenesis, and to design preventive and therapeutic interventions tailored to target these events (3).

In a recent study, Sung and colleagues used NGS to construct a genome-wide HBV DNA integration map at single-base resolution by analyzing tumor and paired non-tumor liver tissues from 81 HBV-positive and 7 HBV-negative Chinese HCC patients (4). The authors identified 399 integration breakpoints in 75 HBV-positive and 1 HBV-negative patients. They found that viral integration occurred more frequently in tumors (344 events) than in paired non-tumor tissues (55 events). Of the 344 integrations into tumor tissue, 179 were found in known coding genes and were over-represented in the exons and promoter regions of these genes. Furthermore, they identified specific sites of recurrent integration, or "hotspots", of HBV DNA at breakpoints in the host genome within particular genes including *TERT*, *MLL4*, *CCNE1*, *SENP5*, and *ROCK1*. The findings are in line with previous studies using PCR- or hybridization-based techniques showing that HBV integrations commonly reside in *TERT* on chromosome 5, and *MLL4* on chromosome 19 (5-9).

Sung *et al*. went on to further characterize the consequences of HBV DNA integrations at these breakpoints, and found that mRNA expression levels of *TERT*, *MLL4* and *CCNE1* genes were significantly higher in tumor tissues carrying recurrent HBV DNA integrations at these breakpoints. In addition, the authors noted integrations at these sites were associated with dramatically increased copy number variations of the host genes, substantiating the link between HBV DNA integration and chromosomal instability. Finally, the authors reported an association between the total number of all detected integration sites and important clinical characteristics of the patient population including AFP levels, age at HCC diagnosis, and overall survival.

Collectively, these findings confirmed *TERT* and *MLL4*

genes as HBV integration hotspots in HCC patients. Moreover, they suggested that recurrent HBV DNA integration into the host genome is an important driving event in HBV-related hepatocarcinogenesis, as well as a potential target for individualized therapies.

Another key finding of the study by Sung *et al*. is that approximately 40% of the integration breakpoints on the HBV genome occurred within a 1,800-bp range, an insertion hotspot region harboring several essential viral genes. This is consistent with the results of another recent study by Jiang and colleagues who applied high-depth whole genome NGS to sequence tumor and adjacent normal liver tissues from 3 HBV-positive HCC patients and 1 HBV-negative HCC patient (10). Interestingly, although Jiang *et al*. also confirmed the link between HBV integration and elevated expression of host genes, there was only one shared recurrent integration site (*MLL4*) between that study and the one by Sung *et al*. This discrepancy may be partially explained by sample size limitations and individual variations. It may also be accounted for by potential intra-tumor heterogeneity, a major challenge in the design and analysis of current NGS studies, despite the commonly recognized theory of clonal expansion in cancer development (11,12). Neither study reported a detailed procedure of HCC sampling (4,10), and therefore it is possible that differences in sampling could have introduced bias into the findings.

Another difference between the two studies relates to sequencing depth. Sung *et al*. identified a total of 344 integration sites from 76 HCC tumors through a >30× depth of coverage, whereas Jiang *et al*. observed 148 integration sites from 3 HCC tumors by a >80× deeper sequencing of samples containing >80% tumor content. Jiang *et al*. further compared the results from the >80× depth to an even higher coverage of >240×, and found the number of detected integration sites were proportional to sequencing depth. Assuming the actual average numbers of integrations per sample were comparable between the tumors in these two studies, the fact that Jiang *et al*. identified a much higher number of integrations per sample might be due to the different stringencies in detection criteria used in the two studies, as well as differences in potential sequencing errors, analysis biases and artifacts. Furthermore, it could have reflected the presence of tumor heterogeneity that resulted from the inclusion of normal tissues in the tumor samples and thus reduced the sequencing sensitivity in the study of Sung *et al*. Further in-depth assessments of allelic bias and number of integration sites that were shared across samples in each study may provide additional clues to these observations.

In addition to constructing HBV integration maps, NGS has been widely applied to survey HCC genomes to identify other aberrations such as mutations, small insertions and deletions, copy number variations, and structural variations. As a result, many frequently mutated genes have been discovered and linked to HCC development, including *TP53*, *AXIN1*, *ARID1A*, *ARID2*, *CTNNB1*, as well as others (13-17). However, a causative link between HBV integration sites and specific gene mutations remains to be demonstrated. Moreover, although Sung *et al*. reported a significant association between the overall numbers of integration sites and patient survival, it is unknown whether this association was particularly prominent in patients with recurrent integrations in hotspot genes. Sung *et al*. (4) reported sequencing data on the largest HBV-positive HCC patient population to date, which may offer pivotal insights into the mechanism of malignant transformation from HBV infection to HCC. A thorough understanding of HBV DNA integration into the host genome, and the resulting aberrant gene expression and cancer development will likely necessitate even higher-depth sequencing coverage in larger patient populations, in order to tackle the heterogeneity issue and yield statistically robust findings. Due to the rapid decrease in NGS cost and increase in computational capacity, it will not be long before future studies likes Sung's and Jiang's pave new roads for researchers to establish a comprehensive understanding of HCC genomes, eventually facilitating the development of novel targeted and personalized therapeutic options to prevent and treat this devastating disease.

Acknowledgements

Disclosure: The authors declare no conflict of interest.

References

1. Farazi PA, DePinho RA. Hepatocellular carcinoma pathogenesis: from genes to environment. Nat Rev Cancer 2006;6:674-87.
2. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet 2010;11:685-96.
3. Tateishi R, Omata M. Hepatocellular carcinoma in 2011: Genomics in hepatocellular carcinoma--a big step forward. Nat Rev Gastroenterol Hepatol 2012;9:69-70.

4. Sung WK, Zheng H, Li S, et al. Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. Nat Genet 2012;44:765-9.
5. Paterlini-Bréchot P, Saigo K, Murakami Y, et al. Hepatitis B virus-related insertional mutagenesis occurs frequently in human liver cancers and recurrently targets human telomerase gene. Oncogene 2003;22:3911-6.
6. Murakami Y, Saigo K, Takashima H, et al. Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas. Gut 2005;54:1162-8.
7. Saigo K, Yoshida K, Ikeda R, et al. Integration of hepatitis B virus DNA into the myeloid/lymphoid or mixed-lineage leukemia (MLL4) gene and rearrangements of MLL4 in human hepatocellular carcinoma. Hum Mutat 2008;29:703-8.
8. Ferber MJ, Montoya DP, Yu C, et al. Integrations of the hepatitis B virus (HBV) and human papillomavirus (HPV) into the human telomerase reverse transcriptase (hTERT) gene in liver and cervical cancers. Oncogene 2003;22:3813-20.
9. Tamori A, Yamanishi Y, Kawashima S, et al. Alteration of gene expression in human hepatocellular carcinoma with integrated hepatitis B virus DNA. Clin Cancer Res 2005;11:5821-6.
10. Jiang Z, Jhunjhunwala S, Liu J, et al. The effects of hepatitis B virus integration into the genomes of hepatocellular carcinoma patients. Genome Res 2012;22:593-601.
11. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012;366:883-92.
12. Caldas C. Cancer sequencing unravels clonal evolution. Nat Biotechnol 2012;30:408-10.
13. Li M, Zhao H, Zhang X, et al. Inactivating mutations of the chromatin remodeling gene ARID2 in hepatocellular carcinoma. Nat Genet 2011;43:828-9.
14. Totoki Y, Tatsuno K, Yamamoto S, et al. High-resolution characterization of a hepatocellular carcinoma genome. Nat Genet 2011;43:464-9.
15. Guichard C, Amaddeo G, Imbeaud S, et al. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nat Genet 2012;44:694-8.
16. Fujimoto A, Totoki Y, Abe T, et al. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. Nat Genet 2012;44:760-4.
17. Huang J, Deng Q, Wang Q, et al. Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma. Nat Genet 2012;44:1117-21.

Cite this article as: Wan S, Civan J, Rossi S, Yang H. Profiling HBV integrations in hepatocellular carcinoma. Hepatobiliary Surg Nutr 2013;2(2):124-126. doi: 10.3978/j.issn.2304-3881.2012.10.13

