Cancer Letters xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Cancer Letters

journal homepage: www.elsevier.com/locate/canlet

Mini-review

Next generation sequencing reveals genetic landscape of hepatocellular carcinomas

Shuyu Li [a,*], Mao Mao [b,*]

[a] LRL IT China, Shanghai, China
[b] Pfizer Oncology, San Diego, CA, United States

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Hepatocellular carcinoma
Genome
Sequencing
Mutation

ABSTRACT

Liver cancer is one of most deadly cancers worldwide. Hepatocellualr carcinoma (HCC) represents a major histological subtype of liver cancers. As cancer is a genetic disease, genetic lesions play a major role in HCC tumorigenesis and progression. Although significant progress has been made to uncover genetic alterations in HCCs, our understanding of genetics involved in the initiation and progression of HCC is far from complete. Next generation sequencing (NGS) has provided a new paradigm in biomedical research to delineate the genetic basis of human diseases. While identification of cancer somatic mutations has been serendipitous, genome sequencing has provided an unbiased approach to systematically catalog somatic mutations and elucidate the mechanisms of tumourigenesis. A number of recently published NGS studies on HCCs have not only confirmed previously known mutations in CTNNB1 and TP53 in HCC, but also identified novel genetic alterations in HCC including mutations in genes involved in epigenetic regulation. WNT, cell cycle and chromatin remodeling pathways have emerged as key oncogenic drivers in HCCs. The frequently altered genes and pathways in HCC reflect classical cancer hallmarks. These findings have started to depict a genetic landscape in HCC and will facilitate development of novel therapeutics for the treatment of this deadly disease.

1. Introduction

Liver cancer is the sixth most commonly diagnosed cancer worldwide and the third leading cause of cancer death [1,2]. Hepatocellualr carcinoma (HCC), arising from hepatocytes, represents the most common histological subtype of liver cancers, accounting for approximately 70–85% of all cases [3]. Another subtype of liver cancer is cholangiocarcinoma originating from the biliary tract epithelium within either liver or the biliary tract [4,5]. Cholangiocarcinomas are relatively rare with high incidence rate in eastern Asia, particularly Thailand mainly due to high prevalence of liver fluke infection [6]. With respect to disease etiology, more than 80% of the HCCs are associated with chronic infections with hepatitis B virus (HBV) or hepatitis C virus (HCV). Other major risk factors for HCC include alcoholic liver diseases, non-alcoholic fatty liver diseases, dietary aflatoxin exposure, obesity, diabetes, smoking and iron overload. Less common causes include hereditary hemochromatosis, alpha1-antitrypsin deficiency, autoimmune hepatitis, some porphyrias, and Wilson's disease [7]. The distribution of these risk factors among HCC patients largely depends on geographic region and ethnic background. Development of liver cirrhosis has been recognized as a major step in HCC pathogenesis as it is found in 80–90% of HCCs [8]. Both HBV and HCV often cause liver inflammation, hepatic damage and subsequently cirrhosis. Cirrhosis is characterized by reduced hepatocyte proliferation, and associated with fibrosis, destruction of liver cells and occurrence of regenerative nodules [9]. It has been postulated that the mechanisms of HCC tumorigenesis in patients with cirrhosis involve accumulation of genetic alterations. Although significant progress has been made to uncover genetic aberrations in HCCs, including identification of point mutations in p53 (TP53) and β-catenin (CTNNB1), amplifications of MYC, FGF19 and cyclin D1 (CCND1), and HBV integrations into the TERT and MLL4 gene locus that encode telomerase reverse transcriptase and a histone lysine methyl transferase respectively [10], our understanding of genetic landscape in HCC is far from complete and the key drivers of HCC tumourigenesis remain poorly understood.

Next generation sequencing (NGS) has provided a new paradigm in biomedical research to uncover the genetic basis of human diseases [11,12]. In the last several years, specific emphasis has focused on applying NGS to further delineate cancer genomes and identify novel genetic mutations that drive tumorigenesis and cancer progression. These efforts have yielded significant results, including several seminal discoveries of recurrent mutations in genes such as IDH1 in glioblastoma multiforme (GBM) [13] and acute myeloid leukemia (AML) [14], CHD7 in small cell lung cancers [15], GRM3, TRRAP, MAP2K1/2, MAP3K5/9, PREX2 in

* Corresponding authors. Addresses: Lilly Research Laboratories IT China, Building 8, 338 Jia Li Lue Road, Zhangjiang Hi-Tech Park, Shanghai 201203, China (S. Li), Pfizer Oncology, 10724 Science Center Drive, San Diego, CA, United States (M. Mao).
E-mail addresses: li_shuyu_dan@lilly.com (S. Li), mao.mao@pfizer.com (M. Mao).

0304-3835/$ - see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.canlet.2012.09.027

Table 1
Published NGS studies on liver cancers.

Ref.	Study design	Summary of results
[32]	WGS and WES of a HCV positive HCC	Identified and validated 63 non-synonymous somatic substitutions including mutations of TP53 and AXIN1. Identified and validated 22 chromosome rearrangements generating four fusion transcripts
[55]	WGS or WES of nine tumor samples and seven adjacent normal liver samples from a HBV positive HCC	Identified and validated 24 somatic mutations that alter amino acid sequences. Evolution of tumors was inferred from mutations derived from multiple primary and recurrent tumor samples
[37]	WES of 10 HCV positive HCCs. Screening of five recurrently mutated genes in an additional cohort of 129 HCCs associated various disease etiologies	Identified frequent mutations in CTNNB1 and TP53. Identified novel and frequent inactivating mutations in ARID2 gene. Mutation frequencies of CTNNB1, TP53 and ARID2 are associated with disease etiology
[31]	WGS of four HCCs (3 HBV positive and 1 HBV negative)	Identified 255 HBV integration sites. HBV integration into MLL4, ANGPT1, and a novel transcript resulted in elevated gene expression
[30]	WES of 24 HCCs, among them one is HBV positive, four are HCV positive, 12 are alcohol related. Screening of 14 genes in an additional cohort of 125 HCCs	Identified frequent mutations in CTNNB1 and TP53. Identified novel and frequent inactivating mutations in ARID1A, ARID2 and other genes involved in chromatin remodeling. Mutation frequencies of these genes are associated with disease etiology
[29]	WGS of 27 HCCs (11 HBV positive, 15 HCV positive, two non-viral). Screening of significantly mutated genes in an additional cohort of 120 HCCs	Identified frequent and novel mutations in genes involved in chromatin remodeling, e.g. ARID1A, ARID1B, ARID2, MLL and MLL3. HBV integration in TERT gene in 4 of the 11 HBV positive HCCs
[49]	WGS of 88 HCCs (81 HBV positive, 7 HBV negative)	Recurrent HBV integration in TERT, MLL4 and CCNE1. HBV integrations led to elevated gene expression. HBV integration associated with chromosome instability, early onset and poor outcome
[33]	WES of 10 HBV positive HCCs. Screening 10 significantly mutated genes in an additional cohort of 100 HCCs	ARID1A mutations in 13% of HBV associated HCCs. Functional study suggested ARID1A mutation may be crucial in HCC invasion and metastasis
[40]	WES of eight cholangiocarcinomas (CCAs). Screening of 15 significantly mutated in 46 additional CCAs	Identified frequent mutations in TP53, KRAS, SMAD4 and MLL3

melanoma [16–20], NOTCH1 in chronic lymphocytic leukemia (CLL) [21], SF3B1 and other genes encoding splicing machinery proteins in myelodysplasia [22,23], and chromatin remodeling genes such as ARID1A in ovarian, kidney and gastric cancers [24–27]. While identification of cancer somatic mutations has historically been serendipitous because it is largely based on candidate gene approaches, whole genome sequencing (WGS) or whole exome sequencing (WES) has provided an unbiased platform to systematically discover somatic mutations and elucidate disease mechanisms of various types of cancer. During the last 18 months, a number of studies have been published on genome or exome sequencing of liver cancers (Table 1). In this review, we will summarize the findings from these studies that at least have begun to reveal the genetic landscape in HCC.

2. Mutation spectrum in HCC

Single point somatic mutations include transition and transversion. It has been shown that each tumor type often has a unique mutation spectrum with different rates of various types of transition and transversions [28]. For example, C:G > T:A transition is essentially the only nucleotide substitution occurring in gliomas and melanomas. While C:G > T:A is also a dominant mutation type in many other solid tumors, C:G > G:C transversion is a major alteration in breast, renal and ovarian cancers. In addition, C:G > A:T transversion and T:A > C:G transition occur frequently in lung and renal cancers respectively [28]. In HCC, multiple studies have indicated that the distribution of somatic mutations is also significantly deviated from the expected spectrum of various substitution patterns [29–32]. Similar to other solid tumors, C:G > T:A transition is one of the most prevalent substitutions. However, both T:A > C:G transition and C:G > A:T transversion are highly abundant, representing a unique mutation signature characteristic of HCC. It was indicated that C:G > A:T transversion is more common in HBV associated HCC than HCV associated HCC, implying aflatoxin B1 exposure may be a critical contributor in a subset of HBV associated HCCs [33]. It is noteworthy that all of the guanine substitutions (G > A, G > T, G > C) occur preferentially at the CpG sites in HCC.

Mutation frequency also varies significantly among different tumor types. Melanoma has the highest mutation rates with most metastatic melanomas harboring 200–600 non-synonymous single nucleotide mutations per genome [16], reflecting high rate of UV induced mutations. Lung cancers also have high mutation rates with up to 100–300 non-synonymous mutations in protein-coding regions [34]. In contrast, leukemia carries the least somatic mutations with less than 10 non-synonymous protein coding alterations in each genome [14,21,35]. Genome or exome sequencing of most other solid tumors including breast, colon, prostate and pancreatic cancers have revealed that on average approximately 20–80 non-synonymous mutations occur per genome in these tumor types. In HCC, the number of non-silent mutations in protein coding regions varies from study to study and also varies between different patients. However, the majority of the HCCs sequenced carry 20–100 non-synonymous mutations (Table 1), a frequency similar to most other solid tumor that have been studied to date. There is no apparent difference of mutation rates in HCCs related to different disease etiology.

3. Most commonly mutated genes in HCC

It has been well known that CTNNB1 is frequently mutated in several tumor types including colon cancer and HCC [36]. It is also well recognized that as a tumor suppressor, TP53 is inactivated through mutation or deletion in more than 50% of tumors. WGS or WES of HCCs have confirmed previous studies that CTNNB1 is the most frequently mutated oncogene and TP53 is the most frequently mutated tumor suppressors (Table 1). Overall, CTNNB1 is mutated in approximately 30% of HCCs, and mutation frequency varies with respect to different disease etiologies. It is mutated less frequently in HBV positive HCCs than HCCs related to other risk factors [30,37]. In a study where the prevalence of mutations in CTNNB1, TP53 and several previously unknown genes in HCCs was surveyed in a cohort of 50 HBV positive HCC, 43 HCV positive HCC, and 44 HCCs not associated with viral infections [37], it was shown that while CTNNB1 mutation is observed in only 10% of HBV-associated HCCs, it is detected in 30.2% and 20.5% of HCV-associated HCCs and non-viral related HCCs respectively,

Please cite this article in press as: S. Li, M. Mao, Next generation sequencing reveals genetic landscape of hepatocellular carcinomas, Cancer Lett. (2012), http://dx.doi.org/10.1016/j.canlet.2012.09.027