Research
New Technology of Tumor Diagnosis and Treatment—Review

Novel Non-Protein Biomarkers for Early Detection of Hepatocellular Carcinoma

Ghassan K. Abou-Alfa [a,b,]*, Lin Wu [c], Augusto Villanueva [d]

[a] Memorial Sloan Kettering Cancer Center, New York City, NY 10065, USA
[b] Weill Medical College, Cornell University, New York City, NY 10065, USA
[c] Berry Oncology Corporation, Beijing 102200, China
[d] Division of Liver Diseases, Division of Hematology/Medical Oncology, Liver Cancer Program, Department of Medicine, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York City, NY 10029, USA

ABSTRACT

Early detection of hepatocellular carcinoma (HCC) while in its early stages is critical for reducing HCC mortality in high-risk patients. However, highly sensitive and specific surveillance biomarkers for early-stage HCC detection are still lacking. In recent years, great efforts have been made to research tumor-derived molecular features that are detectable in circulation, such as circulating tumor deoxyribonucleic acid and circulating tumor ribonucleic acid, in order to explore their potential as non-invasive biomarker candidates in many tumor types. In this review, we summarize current studies on these new approaches and their application in early HCC detection.

© 2021 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Hepatocellular carcinoma (HCC) accounts for 90% of liver cancers and is one of the most common and deadly cancers worldwide, with new cases still rapidly increasing in many countries. The major risk factors of HCC are chronic hepatitis B virus (HBV) infection, hepatitis C virus infection, alcohol use disorder, and non-alcoholic fatty liver disease (NAFLD). Cirrhosis of different causes predisposes patients to HCC at an annual incidence of 2%–4% [1]. The five year net survival rate of HCC patients was in the range of 5%–30% throughout 2000–2014, and has changed very little during the 20 year period from 1995 to 2014 in most countries [2]. If HCC was detected and treated in its early stages, the five year survival rate could increase to 70% [3]. Because many HCC patients are asymptotic in the early stages, almost half of them are diagnosed at an advanced stage [4], when the window for curative treatment is very narrow. Therefore, early HCC detection in the context of surveillance programs has been shown to decrease HCC mortality in high-risk patients [5].

Highly sensitive and specific surveillance biomarkers for early-stage HCC detection are still lacking. Currently, HCC surveillance depends on imaging examinations and serological tests. An abdominal ultrasound with or without serum alpha-fetoprotein (AFP) are the mainstream for HCC surveillance, as they are recommended by the American Association for the Study of Liver Diseases [6] and the European Association for the Study of the Liver guideline 2018 [7]. At a cutoff value of 20 $\mu g \cdot L^{-1}$, AFP has shown limited sensitivities ranging between 41% and 65% and specificities between 80% and 94% in cirrhotic patients, and the sensitivity of AFP for early-stage tumors is even lower, at only 32%–49% [8]. On the other hand, ultrasound is sub-effective for detecting early-stage HCC, with a sensitivity of 63% [9]. A recent comprehensive meta-analysis of more than 10 000 patients found that ultrasound and AFP have a pooled sensitivity of 63% for the detection of early-stage HCC [10]. Additional serum protein targets such as AFP-L3 and, des-γ-carboxy-prothrombin (DCP) (also known as protein induced by vitamin K absence or antagonist-II) have also been explored as biomarkers for early HCC detection [11], but their clinical utility has not been established in the setting of cohort studies. A diagnostic model named GALAD (namely, gender, age, AFP-L3, AFP, and DCP) involving the above three serum protein biomarkers as well as age and gender has been developed, with an area under the curve (AUC) of 0.95 and 0.98 for early and late TNM (where T describes the size of primary tumor, N describes whether regional lymph nodes are affected, and M describes whether distant metastasis is present) stages of HCC, respectively [12]. In early 2020, the

US Food and Drug Administration granted Breakthrough Device designation to the GALAD score to support earlier diagnosis of HCC [13].

In recent years, tumor-derived molecular features that are detectable in circulation, other than serum proteins, have been studied as potential biomarkers in many tumor types. In this review, we summarize current studies on these new approaches and their application in the early HCC detection space. As depicted in Table 1 [14–26], many new biomarkers have been identified and tested for HCC detection, and some have shown potential in early detection.

2. Circulating tumor deoxyribonucleic acid (DNA)

Circulating tumor DNA (ctDNA) refers to tumor-derived DNA fragments that are released into the bloodstream as a result of cellular death, through either apoptosis or necrosis. Such fragments carry tumor-specific alterations, including single nucleotide variants (SNV), insertion/deletion (In/Del), structural variations, and epigenetic alterations, and thus have potential as a biomarker. The biggest challenge in using ctDNA for the early detection of cancer is that it makes up only a minority of the total circulating cell-free DNA (cfDNA). It is estimated that the percentage of ctDNA in cfDNA of early cancer is below 1%, and could be as low as 0.01% [27]. Numerous technological advances have attempted to address this issue, such as the use of digital droplet polymerase chain reaction (PCR) or unique molecular identifiers in next-generation sequencing (NGS) [28]. In HCC, there is evidence that ultra-deep sequencing can detect tissue mutations in the blood of patients at early stages [29].

DNA methylation normally refers to 5-methylcytosine (5mC) modification, which is an epigenetic regulator of gene expression that usually results in gene silencing. Increased DNA methylation of tumor-suppressor genes is an early event in many tumors, making DNA methylation a potential biomarker for early detection. Unlike the limited number of DNA mutation events and sites available in each sample, DNA methylation occurs in multiple target regions and multiple altered cytosine-phosphate-guanine (CpG) sites within each targeted genomic region [30], and thus provides more potential targets than DNA mutations. Several groups have developed techniques for methylation detection and have explored their usefulness in early HCC detection. In 2015, Wen et al. [14] developed a methylated CpG tandem amplification and sequencing (MCTA-Seq) method that can detect hypermethylated CpG islands in cfDNA genome-wide, with a sensitivity of 0.25% allele frequency. Using this technique, they analyzed a small cohort of 27 HCC patients, 17 cirrhosis, and 28 normal individuals, and identified 19 high-performance markers in the blood for detecting small HCC (≤ 3 cm), with four (regulator of G-protein signaling 10 (RGS10), ST8 alpha-*N*-acetyl-neuraminide alpha-2,8-sialyltransferase 6 (ST8SIA6), RUNX family transcription factor 2 (RUNX2), and vimentin (VIM)) concordant with hypermethylation in tumor, and the other 15 already hypermethylated in normal liver tissues. A classifier model composed of these biomarkers achieved a sensitivity of 94% and a specificity of 89% for the plasma samples from 36 HCC patients and control subjects of 17 cirrhosis patients and 38 normal individuals. Notably, all 15 AFP-negative HCC patients were successfully identified, indicating that there is potential in combining these DNA methylation biomarkers with AFP in the future. In 2017, Xu et al. [15] conducted the methylation profiling of cfDNA samples from a much larger cohort consisting of 1098 HCC patients and 835 normal controls in order to identify and validate an HCC-specific methylation biomarker panel for early detection, with targeted bisulfite sequencing. Using so-called methylation-correlated blocks as the unit to quantify the CpG methylation level, the group constructed a diagnostic prediction model consisting of ten methylation markers (cg10428836, cg26668608, cg25754195, cg05205842, cg11606215, cg24067911, cg18196829, cg23211949, cg17213048, and cg25459300) with an AUC of 0.944 (95% confidence interval (CI), 0.928–0.961) in a validation cohort with 383 HCC patients and 275 normal individuals. The combined diagnosis score (cd-score) was highly correlated with tumor burden, treatment response, and stage. However, the majority of HCC cases in this study had tumors at advanced stages, which limits the extrapolation of these results to the early-detection clinical scenario. Another DNA methylation-based detection method also reported sensitivity and specificity higher than 90% [16], which could outperform the current recommended tools for HCC surveillance. A positive correlation of detection rate and tumor stages was also seen by the Circulating Cell-free Genome Atlas (CCGA) Consortium, which conducted bisulfite sequencing targeting a panel of more than 100 000 methylation regions in the plasma DNA of more than 50 types of cancers, including liver cancers [31], suggesting the suboptimal utility of DNA methylation biomarkers in early detection.

5-hydroxymethylcytosine (5hmC) is another type of epigenetic marker. It is a stable product of demethylation, generated through the oxidation of 5mCs by the 10–11 translocation family dioxygenases [32]. 5hmC modifications in enhancers, promotors, and gene bodies impact gene expression. The techniques for detecting 5hmC modification in cfDNA, hMe-Seal, and 5hmC-Seal, reported in 2017 from two different laboratories, involve the selective chemical labeling of 5hmC followed by enrichment and sequencing [33,34]. In 2019, Cai et al. [17] used the 5hmC-Seal technique to profile genome-wide 5hmCs in cfDNA samples from 1204 HCC patients, 392 chronic hepatitis B (CHB) infection/liver cirrhosis (LC) patients, and 958 healthy individuals/benign liver lesion patients. Focusing on the change of 5hmC in gene bodies, they developed a 32 gene classifier for distinguishing early HCC (stage 0/A, Barcelona Clinic Liver Cancer (BCLC)) from non-HCC at an AUC of 88.4% (95% CI, 85.8%–91.1%) and from a high-risk group at an AUC of 84.6% (95% CI, 80.6%–88.7%), both independent of potential confounders, such as smoking or alcohol intake history.

Structural variations are the hallmark of cancers. Several groups have developed methods to evaluate copy number variation (CNV) in ctDNA for the early detection of HCC. In 2015, Xu et al. [35] analyzed CNVs in a small cohort of plasma samples with 31 HCC and eight chronic hepatitis/cirrhosis patients based on low-depth whole-genome sequencing (WGS) of $0.1\times$–$0.2\times$. By CNV Z score analysis, they identified several differential variables (e.g., gain in 1q, 7q, and 19q in HCC) and some less differential variable (e.g., loss in 4q, 13q, gain in 17q, 22q) regions, based on which they proposed a CNV scoring method that generated a positive result in 26 of the 31 HCC patients (83.9%), or in 11 of the 16 HCC patients with a tumor dimension of up to 50 mm (68.8%), or in four of the seven HCC patients with a tumor dimension of up to 30 mm (57.1%). Notably, all eight samples with chronic hepatitis or cirrhosis scored negative. Although CNV analysis alone was not good enough for the early detection of HCC, it might serve as a parameter in model building. In 2020, Tao et al. [18] conducted a deeper low-depth WGS of $5\times$ to profile CNVs in a larger cohort with 384 plasma samples of HBV-related HCC and cancer-free HBV patients. They used machine learning to develop a model with a discovery cohort of 209 patients, achieving an AUC of 0.893, with 0.874 for early stages (BCLC stages 0–A) and 0.933 for more advanced stages (BCLC stages B–D). The performance of the model was validated in two cohorts (76 and 99 patients) that only consisted of patients with stages 0–A HCC and HBV infection, with an AUC of 0.920 and 0.812, respectively. In addition, the researchers found that, for early detection, lowering the sequencing depth decreased the

Table 1
Candidate biomarkers for HCC early detection.

Molecular candidate	Biomarkers	Technology	Performance	Validation cohort	EDRN phase	Reference
DNA methylation (5mC)	19 markers	MCTA-Seq	94% sensitivity and 89% specificity for HCC versus non-HCC	36 HCC/17 LC/38 normal	2	[14]
	Ten markers (cg10428836, cg26668608, cg25754195, cg05205842, cg11606215, cg24067911, cg18196829, cg23211949, cg17213048, and cg25459300)	Targeted bisulfite sequencing	AUC of 0.944 (95% CI, 0.928–0.961)	383 HCC/275 normal	2	[15]
	Six markers (HOXA1, EMX1, AK055957, ECE1, PFKP, and CLEC11A) normalized by B3GALT6	TELQAS	AUC of 0.96 (95% CI, 0.93–0.99) for HCC versus non-HCC HCC sensitivity of 95% (88%–98%) with 75% stage 0 and 93% stage A, at specificity of 92% (86%–96%)	95 HCC (4 stage 0, 42 stage A)/51 LC/98 healthy	2	[16]
5hmC modification	32 markers	5hmC-Seal	AUC of 88.4% (95% CI, 85.8%–91.1%) in cohort 1 for HCC versus non-HCC AUC of 84.6% (95% CI, 80.6%–88.7%) in cohort 2 for HCC versus controls	Cohort 1: 220 early HCC/129 CHB or LC/256 control Cohort 2: 24 early HCC/180 control	2	[17]
CNV	Global	Low-depth WGS	AUC of 0.920 in cohort 1 AUC of 0.812 in cohort 2	Cohort 1: 38 early HCC/38 HBV Cohort 2: 51 early HCC/48 HBV	2	[18]
Fragment size	Global	Low-depth WGS	AUC of 0.88 for HCC versus non-HCC	90 HCC/32 healthy/67 HBV/36 LC	2	[19]
miRNA	Seven markers (miR-122, miR-192, miR-21, miR-223, miR-26a, miR-27a, and miR-801)	qPCR	AUCs for BCLC stages 0, A, B, and C versus non-HCC were 0.888, 0.888, 0.901, and 0.881, respectively AUCs for HCC versus healthy, CHB, and cirrhosis were 0.941, 0.842, and 0.884, respectively	196 HCC/66 healthy/72 CHB/z 56 LC	2	[20]
	Eight markers (miR-206, miR-141-3p, miR-433-3p, miR-1228-5p, miR-199a-5p, miR-122-5p, miR-192-5p, and miR-26a-5p)	qRT-PCR	AUCs of 0.879 for HCC versus non-HCC, 0.893 for HCC versus healthy, and 0.892 for HCC versus cirrhosis	103 HCC/78 LC/60 healthy	2	[21]
	Seven markers (miR-29a, miR-29c, miR-133a, miR-143, miR-145, miR-192, and miR-505)	qPCR	AUC of 0.817 (0.769–0.865) in cohort 1 and 0.884 (0.818–0.951) in cohort 2 for HCC versus non-HCC 74.5% sensitivity and 89.9% specificity for HCC versus CHB/LC; 85.7% sensitivity and 83.3% specificity for HCC versus inactive HBsAg carrier Sensitivity of 29.6%, 48.1%, 48.1%, and 55.6% in 12, 9, 6, and 3 months for HCC before clinical diagnosis in cohort 3	Cohort 1: 153 HCC/60 healthy/68 CHB/71 LC Cohort 2: 49 HCC/48 healthy/42 IHC Cohort 3 (nested case-control study): 27 HCC/135 controls (CHB or LC)	2,3	[22]
lncRNA	Three markers (LINC00152, RP11-160H22.5, and XLOC014172)	qRT-PCR	AUC of lncRNA panel was unavailable AUC of 0.986 and 0.985 for HCC versus CH and HCC versus healthy, respectively, when combine with AFP	100 HCC/100 CHB/100 healthy	2	[23]
Viral exposure signature	Unique epitopes corresponding to 61 viral strains	Serological profiling with VirScan	AUC of 0.91 at baseline and 0.98 at diagnosis	Prospective at-risk cohort: 44 HCC/129 cancer-free	3	[24]
Multi-analyte	DNA mutation, HBV integrations, cfDNA concentration, protein markers, gender, and age	PCR-based sequencing and CLIA	17% PPV	Prospective AFP/ultrasound negative cohort: 24 HCC/307 cancer-free after 6–8 months	3	[25]
	5hmC, end motifs, fragmentation, and nucleosome footprints	5hmC sequencing and low-pass WGS	Sensitivity of 95.42% and specificity of 97.83% for HCC versus LC Sensitivity of 95.42% and specificity of 97.91% for HCC versus non-HCC	131 HCC/1800 LC/116 healthy	2	[26]

DNA: deoxyribonucleic acid; cfDNA: cell-free DNA; 5mC: 5-methylcytosine; 5hmC: 5-hydroxymethylcytosine; MCTA-Seq: methylated cytosine-phosphate-guanine (CpG) tandems amplification and sequencing; CNV: copy number variation; BCLC: Barcelona Clinic Liver Cancer; HBsAg: hepatitis B surface antigen; EDRN: Early Detection Research Network; TELQAS: target enrichment long-probe quantitative amplified signal; CI: confidence interval; WGS: whole-genome sequencing; LC: liver cirrhosis; CHB: chronic hepatitis B; IHC: inactive HBsAg carrier; CH: chronic hepatitis; qPCR: quantitative polymerase chain reaction; qRT-PCR: quantitative reverse-transcriptase polymerase chain reaction; PCR: polymerase chain reaction; RNA: ribonucleic acid; miRNA: microRNA; lncRNA: long non-coding RNA; CLIA: chemical luminescence immunity analyzer; PPV: positive predictive value; HOXA1: homeobox A1; EMX1: empty spiracles homeobox 1; ECE1: endothelin-converting enzyme 1; PFKP: phosphofructokinase; CLEC11A: C-type lectin domain containing 11A; B3GALT6: beta-1,3-galactosyltransferase 6.

sensitivity, which suggested that an adequate sequencing depth might be required for stable performance of the model.

3. Fragmentomics

cfDNA is highly fragmented due to the endonuclease digestion of nucleosome free regions. Fragmentation of cfDNA is not random, and may carry tissue or tumor-specific signatures. Fragmentomics refers to the analysis of the molecular characteristics of cfDNA fragmentation patterns, including plasma DNA sizes, end points, and nucleosome footprints [36]. These molecular characteristics of cfDNA can be readily analyzed from WGS data.

To understand the size distribution of cfDNA fragments for HCC, in 2015, Jiang et al. [37] performed a genome-wide analysis of cfDNA size profiles in 90 HCC patients, 67 CHB patients, 36 hepatitis B-associated cirrhosis patients, and 32 healthy controls. They found that the cfDNA of patients with HCC is more variable, with aberrantly short or long length. The short ones preferentially carried the tumor-associated copy number aberrations. The researchers also found that there were elevated amounts of mitochondrial DNA in the plasma of HCC patients. Such molecules were much shorter than the nuclear DNA in plasma. In 2019, Cristiano et al. [38] evaluated the fragmentation patterns of cfDNA across the genome and found that the profiles of healthy individuals reflected the nucleosomal patterns of white blood cells, whereas patients with cancer had altered fragmentation profiles. A machine learning model using genome-wide fragmentation features was found to have detection sensitivities ranging from 57% to more than 99% among seven cancer types at 98% specificity, with an overall AUC of 0.94. Unfortunately, this study did not include liver cancer samples.

To explore the utility of the end position of cfDNA fragments, in 2018, Jiang et al. [19] investigated whether there was a ctDNA signature in the form of preferred plasma DNA end coordinates associated with early HCC detection. Studying the DNA end characteristics in the plasma of patients with HCC and CHB, they identified millions of tumor-associated plasma DNA end coordinates in the genome. The ratios of tumor- to non-tumor-associated preferred ends were significantly increased in the plasma samples of the 90 HCC patients compared with those of non-HCC participants (32 healthy controls, 67 HBV carriers, and 36 LC), with an AUC of 0.88 to distinguish HCC patients from controls. Plasma DNA end coordinates were more readily detectable than somatic mutations as a specific cancer signature in plasma. To explore the utility of fragment end information, in 2020, the group further looked into the 5′ end motif of HCC and found a significant increase in the diversity of plasma DNA end motifs in HCC patients [39]. In particular, the abundance of the plasma DNA motif CCCA was much lower in patients with HCC than those without. Through a comparison of the aberrant end motifs with those of other cancer types, the researchers observed that the profile of plasma DNA end motifs originating from the same organ, such as the liver, placenta, and hematopoietic cells, generally clustered together, indicating that such markers carry tissue-of-origin information. Although a preferential pattern of 4-mer end motifs was identified for HCC, its role in distinguishing HCC from LC was not clear.

cfDNA reflects nucleosome footprints. In actively transcribed genes, the promoter region and downstream gene body are free of nucleosome, resulting in reduced frequencies of mapped reads. Nucleosome spacing inferred from cfDNA in healthy individuals correlates most strongly with the epigenetic features of lymphoid and myeloid cells [40]. Ulz et al. [41] demonstrated that nucleosome occupancy around the transcription start site in cfDNA could result in different read depth coverage patterns for expressed and silent genes. Most recently, Chen et al. [26] explored nucleosome footprints along with other genomic features in cfDNA for liver cancer detection, and found that nucleosome footprints alone could achieve an AUC of 0.973 in differentiating HCC from LC.

4. Circulating tumor ribonucleic acid (RNA)

Circulating microRNA (miRNA) and long non-coding RNA (lncRNA) are also potentially good biomarkers for cancers. miRNAs are a class of endogenous small non-coding RNA transcripts of about 22 nucleotides in length, while lncRNAs are longer non-protein coding transcripts of more than 200 nucleotides in length. Both miRNAs and lncRNAs are important regulatory molecules for gene expression as they are involved in multiple cellular processes, and their dysregulation is related to multiple diseases, including cancers. miRNAs and lncRNAs can be found in the circulation in healthy and diseased individuals, and studies of circulating RNA in HCC early detection have been carried out much earlier than those of ctDNA.

There are dozens of publications of studies of miRNA as HCC biomarkers; some have identified targets from microarray or NGS profiling, while others have tested miRNA candidates from a literature search. The method used to quantify miRNA targets is usually quantitative reverse-transcriptase polymerase chain reaction (qRT-PCR) assay. As early as 2011, Zhou et al. [20] conducted a study with three independent cohorts including 934 participants (healthy, CHB, cirrhosis, and HBV-related HCC). The researchers first profiled plasma miRNA expression with a microarray targeting 723 miRNAs in 137 samples and identified seven potential biomarkers (miR-122, miR-192, miR-21, miR-223, miR-26a, miR-27a, and miR-801) for distinguishing HCC from non-HCC. Then they evaluated the expression of the miRNA panel by means of quantitative polymerase chain reaction (qPCR). A logistic regression model built on a training cohort of 407 samples showed an AUC of 0.888 in a validation cohort of 390, which was independent of disease status, with AUCs for BCLC stages 0, A, B, and C of 0.888, 0.888, 0.901, and 0.881, respectively. The miRNA panel showed a better performance in differentiating HCC patients from healthy controls (AUC 0.941) than from CHB patients (AUC 0.842) and cirrhosis patients (AUC 0.884).

In 2014, Tan et al. [21] conducted a similar study with a total of 667 samples (261 HCC patients, 233 cirrhosis patients, and 173 healthy controls), in which the initial screening of miRNA expression was done by NGS using serum samples pooled from HCC patients and controls. The group identified eight miRNAs (hsa-miR-206, hsa-miR-141-3p, hsa-miR-433-3p, hsa-miR-1228-5p, hsa-miR-199a-5p, hsa-miR-122-5p, hsa-miR-192-5p, and hsa-miR-26a-5p), and the panel with a logistic regression model had an AUC of 0.887 and 0.879 for the training (357) and validation (241) sets, respectively, which was similar to the panel of Zhou et al. [20]. Unlike the panel of Zhou et al. [20], this miRNA panel had almost the same power in differentiating HCC patients from healthy controls (AUC = 0.893) as from cirrhosis patients (AUC = 0.892). However, it is not clear whether this panel was independent of HCC stage. In 2015, Lin et al. [22] reported a study testing their identified panel for predicting preclinical HCC. After discovery and validation phases using retrospective cohorts of HCC, cirrhosis patients related to HBV infection, inactive hepatitis B surface antigen (HBsAg) carriers, and healthy controls, they came up with a serum miRNA classifier (Cmi) of a seven miRNA panel (miR-29a, miR-29c, miR-133a, miR-143, miR-145, miR-192, and miR-505). The panel had a sensitivity of 74.5% and a specificity of 89.9% for distinguishing HCC from CHB plus LC patients, and a sensitivity of 85.7% and a specificity of 83.3% for HCC patients versus

an inactive HBsAg carrier group, respectively. The report included a nested case-control study of 27 cases, which found that the sensitivity of Cmi in detecting HCC was 29.6%, 48.1%, 48.1%, and 55.6% at 12, 9, 6, and 3 months before clinical diagnosis.

Compared with studies on circulating miRNA, studies on circulating lncRNA as an HCC biomarker are fewer and have much smaller cohorts. For example, in 2017, Yuan et al. [23] tested ten candidate circulating lncRNAs selected from the literature with qRT-PCR and identified four lncRNAs in a training set of 20 HCC patients and 20 controls, which were further narrowed down to three (LINC00152, RP11-160H22.5, and XLOC014172) in a validation set of 100 each of HCC patients and controls. The combination of three lncRNAs with AFP could distinguish the HCC patients from either chronic hepatitis patients or healthy controls with an AUC of 0.986 and 0.985, respectively.

5. Viral exposure signature

HCC is a virus-related malignancy, and virus infection may shape host immunity, thus defining the onset of the cancer. Therefore, a unique viral exposure signature resulting from virus–host interactions could reflect a cascade of events that may alter the risk of developing HCC. To test this hypothesis, Liu et al. [24] performed serological profiling of the viral infection history of 899 individuals from a National Cancer Institute–University of Maryland (NCI–UMD) case–control study using a synthetic human virome, VirScan. They developed a viral exposure signature and validated the results in a longitudinal cohort with 173 at-risk patients who had long-term follow-up for HCC development. The viral exposure signature was significantly associated with HCC status among the at-risk individuals in the validation cohort, with an AUC of 0.91 at baseline and 0.98 at diagnosis. The viral signature identified HCC patients prior to a clinical diagnosis and was superior to AFP.

6. Multiple analytes

Due to the inherent molecular heterogeneity of cancer, an early-detection biomarker may need to encompass multiple molecular dimensions in order to achieve a competitive performance. For example, Cohen et al. [42] developed a blood test called CancerSEEK to detect eight common cancer types through an assessment of the levels of mutations in cfDNA and 39 circulating proteins. The mutations were detected with a 61 amplicon panel, with each amplicon querying an average of 33 base pairs within one of 16 frequently mutated genes in common cancers. The sensitivity of CancerSEEK for liver cancer is as high as 98%, with an overall specificity of greater than 99%. However, the sensitivity of CancerSEEK is dependent on the stage of the cancer, and few HCC early-stage samples were included. In addition, a follow-up study from the same group reported relatively low positive predictive values using a blood-only test for the detection of different tumor types [43].

For the early detection of liver cancer, different approaches with multiple analytes have been reported. Qu et al. [25] developed an HCC screen assay, which includes DNA mutation, HBV integrations, cfDNA concentration, protein markers, gender, and age. The assay robustly separated HCC from non-HCC patients with a sensitivity of 85% and a specificity of 93% in the training set and with a 17% positive predictive value in the validation cohort. Chen et al. [26] developed a HIFI method by integrating four genomic features of cfDNA: 5hmC modification, end motifs, fragmentation, and nucleosome footprints. This method achieved high accuracy in differentiating HCCs from LC, with a sensitivity of 95.42% and a specificity of 97.83% in a test set, irrespective of demographics and clinical features including age, HBV status, Child–Pugh score, BCLC stage, tumor size, and AFP status.

7. Outlook

The need for better tools for early HCC detection cannot be overemphasized. In recent years, a number of new molecular approaches have been aimed at the detection of tumor components releases into the bloodstream, in the broader context of liquid biopsy applications in biomedicine. These new attempts have shed a bright light onto the early detection of HCC because, while direct comparison were available, these molecular biomarkers showed better AUC than AFP. Due to the heterogeneity of HCC and the relatively low ratio of tumor-specific genetic materials in circulation in the early stage, an early-detection model comprised of only one type of biomarker has limitations in terms of sensitivity and specificity. Although a combination of multi-dimensional parameters has barely been explored, it holds the promise of significantly increasing early-detection rates. Multi-dimensional parameters may also include traditional tools such as clinical pathological index, protein biomarkers, and molecular imaging.

Biomarker development for early detection generally requires five phases [44]. As listed by the Early Detection Research Network (EDRN) guideline, these are: a preclinical exploratory study, clinical assay development for clinical disease, a retrospective longitudinal repository study, a prospective screening study, and a cancer control study. Most of the early-detection methods summarized herein are still in phase 2, in which the ability to distinguish HCC from non-HCC is assessed using clinical samples. Several studies have progressed to phase 3, in which the capacity of the biomarker to detect preclinical disease is evaluated. All of them are retrospective, in the sense that no referral has been made based on these tests; thus, clinical usefulness needs to be further tested in prospective screening studies.

It should be noted that the new tools reviewed herein are for early detection, not for diagnosis, as patients with a positive early-detection test should undergo a definitive diagnostic procedure (e.g., magnetic resonance imaging and biopsy) according to the recall policy of the surveillance program. With this in mind, there are several challenging issues in the development and clinical use of cutting-edge techniques for the early detection of HCC:

(1) **The population targeted**. The targeted population comprises individuals at risk of HCC, including those with LC, chronical hepatitis virus infection, alcohol abuse, NAFLD, or a family history of HCC. Such individuals should take these tests every six months, as is currently done with AFP/ultrasound.

(2) **The selection and cost efficiency of the combination of multiple cutting-edge techniques and biomarkers.** A combination of multiple techniques and biomarkers can be selected based on the added detection value as well as the cost of each technique/biomarker. Cost reduction due to technological development should also be taken into consideration.

(3) **Acceptance of at-risk individuals for multiple biomarker examinations.** The acceptance of at-risk individuals for the practice of multiple biomarker examinations would mainly depend on the detection rates of the detection tools, the individual's current health situation and health awareness, the cost of the test, and governmental policy.

Compliance with ethics guidelines

Ghassan K. Abou-Alfa, Lin Wu, and Augusto Villanueva declare that they have no conflict of interest or financial conflicts to disclose.

References

[1] Villanueva A. Hepatocellular carcinoma. N Engl J Med 2019;380(15):1450–62.

[2] Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, et al. Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 2018;391 (10125):1023–75.

[3] Llovet JM, Zucman-Rossi J, Pikarsky E, Sangro B, Schwartz M, Sherman M, et al. Hepatocellular carcinoma. Nat Rev Dis Primers 2016;2(1):16018.

[4] Cancer stat facts: liver and and intrahepatic bile duct cancer [Internet]. Washington, DC: National Cancer Institute; c2020 [cited 2020 Jul 6]. Available from: https://seer.cancer.gov/statfacts/html/livibd.html.

[5] Van Meer S, de Man RA, Coenraad MJ, Sprengers D, van Nieuwkerk KMJ, Klümpen HJ, et al. Surveillance for hepatocellular carcinoma is associated with increased survival: results from a large cohort in the Netherlands. J Hepatol 2015;63(5):1156–63.

[6] Marrero JA, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM, et al. Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the American Association for the Study of Liver Diseases. Hepatology 2018;68(2):723–50.

[7] European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of hepatocellular carcinoma. J Hepatol 2018;69 (1):182–236.

[8] Kanwal F, Singal AG. Surveillance for hepatocellular carcinoma: current best practice and future direction. Gastroenterology 2019;157(1):54–64.

[9] Singal A, Volk ML, Waljee A, Salgia R, Higgins P, Rogers MAM, et al. Meta-analysis: surveillance with ultrasound for early-stage hepatocellular carcinoma in patients with cirrhosis. Aliment Pharmacol Ther 2009;30(1):37–47.

[10] Tzartzeva K, Obi J, Rich NE, Parikh ND, Marrero JA, Yopp A, et al. Surveillance imaging and alpha fetoprotein for early detection of hepatocellular carcinoma in patients with cirrhosis: a meta-analysis. Gastroenterology 2018;154 (6):1706–18.e1.

[11] Choi JY, Jung SW, Kim HY, Kim M, Kim Y, Kim DG, et al. Diagnostic value of AFP-L3 and PIVKA-II in hepatocellular carcinoma according to total-AFP. World J Gastroenterol 2013;19(3):339–46.

[12] Johnson PJ, Pirrie SJ, Cox TF, Berhane S, Teng M, Palmer D, et al. The detection of hepatocellular carcinoma using a prospectively developed and validated model based on serological biomarkers. Cancer Epidemiol Biomarkers Prev 2014;23(1):144–53.

[13] FDA grants Breakthrough Device Designation for Roche's Elecsys GALAD score to support earlier diagnosis of hepatocellular carcinoma [Internet]. Basel: Roche; c2021 [cited 2020 Jul 30]. Available from: https://www. roche.com/media/releases/med-cor-2020-03-04.htm.

[14] Wen L, Li J, Guo H, Liu X, Zheng S, Zhang D, et al. Genome-scale detection of hypermethylated CpG islands in circulating cell-free DNA of hepatocellular carcinoma patients. Cell Res 2015;25(11):1250–64.

[15] Xu RH, Wei W, Krawczyk M, Wang W, Luo H, Flagg K, et al. Circulating tumour DNA methylation markers for diagnosis and prognosis of hepatocellular carcinoma. Nat Mater 2017;16(11):1155–61.

[16] Kisiel JB, Dukek BA, VSR Kanipakam R, Ghoz HM, Yab TC, Berger CK, et al. Hepatocellular carcinoma detection by plasma methylated DNA: discovery, phase I pilot, and phase II clinical validation. Hepatology 2019;69(3):1180–92.

[17] Cai J, Chen L, Zhang Z, Zhang X, Lu X, Liu W, et al. Genome-wide mapping of 5-hydroxymethylcytosines in circulating cell-free DNA as a non-invasive approach for early detection of hepatocellular carcinoma. Gut 2019;68 (12):2195–205.

[18] Tao K, Bian Z, Zhang Q, Guo X, Yin C, Wang Y, et al. Machine learning-based genome-wide interrogation of somatic copy number aberrations in circulating tumor DNA for early detection of hepatocellular carcinoma. EBioMedicine 2020;56:102811.

[19] Jiang P, Sun K, Tong YK, Cheng SH, Cheng THT, Heung MMS, et al. Preferred end coordinates and somatic variants as signatures of circulating tumor DNA associated with hepatocellular carcinoma. Proc Natl Acad Sci USA 2018;115 (46):E10925–33.

[20] Zhou J, Yu L, Gao X, Hu J, Wang J, Dai Z, et al. Plasma microRNA panel to diagnose hepatitis B virus-related hepatocellular carcinoma. J Clin Oncol 2011;29(36):4781–8.

[21] Tan Y, Ge G, Pan T, Wen D, Chen L, Yu X, et al. A serum microRNA panel as potential biomarkers for hepatocellular carcinoma related with hepatitis B virus. PLoS ONE 2014;9(9):e107986.

[22] Lin XJ, Chong Y, Guo ZW, Xie C, Yang XJ, Zhang Q, et al. A serum microRNA classifier for early detection of hepatocellular carcinoma: a multicentre, retrospective, longitudinal biomarker identification study with a nested case-control study. Lancet Oncol 2015;16(7):804–15.

[23] Yuan W, Sun Y, Liu L, Zhou B, Wang S, Gu D. Circulating lncRNAs serve as diagnostic markers for hepatocellular carcinoma. Cell Physiol Biochem 2017;44(1):125–32.

[24] Liu J, Tang W, Budhu A, Forgues M, Hernandez MO, Candia J, et al. A viral exposure signature defines early onset of hepatocellular carcinoma. Cell 2020;182(2):317–28.e10.

[25] Qu C, Wang Y, Wang P, Chen K, Wang M, Zeng H, et al. Detection of early-stage hepatocellular carcinoma in asymptomatic HBsAg-seropositive individuals by liquid biopsy. Proc Natl Acad Sci USA 2019;116(13):6308–12.

[26] Chen L, Abou-Alfa GK, Zheng B, Liu JF, Bai J, Du LT, et al. Genome-scale profiling of circulating cell-free DNA signatures for early detection of hepatocellular carcinoma in cirrhotic patients. Cell Res 2021;31(5):589–92.

[27] Abbosh C, Birkbak NJ, Wilson GA, Jamal-Hanjani M, Constantin T, Salari R, et al. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. Nature 2017;545(7655):446–51.

[28] Von Felden J, Garcia-Lezana T, Schulze K, Losic B, Villanueva A. Liquid biopsy in the clinical management of hepatocellular carcinoma. Gut 2020;69 (11):2025–34.

[29] Labgaa I, Villacorta-Martin C, D'Avola D, Craig AJ, von Felden J, Martins-Filho SN, et al. A pilot study of ultra-deep targeted sequencing of plasma DNA identifies driver mutations in hepatocellular carcinoma. Oncogene 2018;37 (27):3740–52.

[30] Villanueva A, Portela A, Sayols S, Battiston C, Hoshida Y, Méndez-González J, et al. DNA methylation-based prognosis and epidrivers in hepatocellular carcinoma. Hepatology 2015;61(6):1945–56.

[31] Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV, Liu MC, et al. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Ann Oncol 2020;31(6):745–59.

[32] Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 2009;324(5929):930–5.

[33] Song CX, Yin S, Ma L, Wheeler A, Chen Y, Zhang Y, et al. 5-hydroxymethylcytosine signatures in cell-free DNA provide information about tumor types and stages. Cell Res 2017;27(10):1231–42.

[34] Li W, Zhang X, Lu X, You L, Song Y, Luo Z, et al. 5-hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for human cancers. Cell Res 2017;27(10):1243–57.

[35] Xu H, Zhu X, Xu Z, Hu Y, Bo S, Xing T, et al. Non-invasive analysis of genomic copy number variation in patients with hepatocellular carcinoma by next generation DNA sequencing. J Cancer 2015;6(3):247–53.

[36] Ivanov M, Baranova A, Butler T, Spellman P, Mileyko V. Non-random fragmentation patterns in circulating cell-free DNA reflect epigenetic regulation. BMC Genomics 2015;16(Suppl 13):S1.

[37] Jiang P, Chan CWM, Chan KCA, Cheng SH, Wong J, Wong VS, et al. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. Proc Natl Acad Sci USA 2015;112(11):E1317–25.

[38] Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. Nature 2019;570 (7761):385–9.

[39] Jiang P, Sun K, Peng W, Cheng SH, Ni M, Yeung PC, et al. Plasma DNA end-motif profiling as a fragmentomic marker in cancer, pregnancy, and transplantation. Cancer Discov 2020;10(5):664–73.

[40] Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA comprises an *in vivo* nucleosome footprint that informs its tissues-of-origin. Cell 2016;164(1–2):57–68.

[41] Ulz P, Thallinger GG, Auer M, Graf R, Kashofer K, Jahn SW, et al. Inferring expressed genes by whole-genome sequencing of plasma DNA. Nat Genet 2016;48(10):1273–8.

[42] Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science 2018;359(6378):926–30.

[43] Lennon AM, Buchanan AH, Kinde I, Warren A, Honushefsky A, Cohain AT, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. Science 2020;369(6499):eabb9601.

[44] Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, et al. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst 2001;93(14):1054–61.