

RESEARCH

Open Access



HepaClear, a blood-based panel combining novel methylated CpG sites and protein markers, for the detection of early-stage hepatocellular carcinoma

Yi Bai^{1†}, Juan Xu^{3†}, Deqiang Li^{4†}, Xiaoyu Zhang^{5†}, Dapeng Chen⁵, Fucun Xie^{2,6,7,8}, Longmei Huang⁴, Xiaotian Yu⁴, Haitao Zhao^{2*} and Yamin Zhang^{1*}

Abstract

Background Early screening and detection of hepatocellular carcinoma (HCC) can efficiently improve patient prognosis. We aimed to identify a series of hypermethylated DNA markers and develop a blood-based HCC diagnosis panel containing DNA methylation sites and protein markers with improved sensitivity for early-stage HCC detection.

Results Overall, 850K methylation arrays were performed using paired tissue DNA samples from 60 HCC patients. Ten candidate hypermethylated CpG sites were selected for further evaluation by quantitative methylation-specific PCR with 60 pairs of tissue samples. Six methylated CpG sites, along with α -fetoprotein (AFP) and des-gamma-carboxyprothrombin (DCP), were assayed in 150 plasma samples. Finally, an HCC diagnosis panel, named HepaClear, was developed in a cohort consisting of 296 plasma samples and validated in an independent cohort consisting of 198 plasma samples. The HepaClear panel, containing 3 hypermethylated CpG sites (cg14263942, cg12701184, and cg14570307) and 2 protein markers (AFP and DCP), yielded a sensitivity of 82.6% and a specificity of 96.2% in the training set and a sensitivity of 84.7% and a specificity of 92.0% in the validation set. The HepaClear panel had higher sensitivity (72.0%) for early-stage HCC than AFP (≥ 20 ng/mL, 48.0%) and DCP (≥ 40 mAU/mL, 62.0%) and detected 67.5% of AFP-negative HCC patients (AFP ≤ 20 ng/mL).

Conclusions We developed a multimarker HCC detection panel (HepaClear) that shows high sensitivity for early-stage HCC. The HepaClear panel exhibits high potential for HCC screening and diagnosis from an at-risk population.

Keywords DNA methylation, Hepatocellular carcinoma, 850K methylation array, Multi-target panel, Early detection

[†]Yi Bai, Juan Xu, Deqiang Li and Xiaoyu Zhang have contributed equally to this work.

*Correspondence:

Haitao Zhao
zhaoh@pumch.cn
Yamin Zhang
5020200824@nankai.edu.cn

¹ Department of Hepatobiliary Surgery, Tianjin First Central Hospital, School of Medicine, Nankai University, Tianjin, China

² Department of Liver Surgery, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

³ Department of Infectious Diseases, Central Hospital of Shengli Oilfield, Dongying, China

⁴ Hangzhou New Horizon Health Technology Co., Ltd, Hangzhou, China

⁵ The First Central Clinical School, Tianjin Medical University, Tianjin, China

⁶ Department of Thoracic Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

⁷ State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

⁸ Central Laboratory, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital and Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen, China



Background

Liver cancer is the third leading cause of cancer death worldwide [1] and creates a significant public health burden. Hepatocellular carcinoma (HCC) accounts for up to 85–90% of liver cancers [2]. In China, HCC is the second most deadly cancer [3], and the 5-year survival rate of HCC patients is only approximately 14% [4], while over 50% of early-stage HCC (BCLC 0/A) patients can live more than 5 years after standard treatment [5]. A poor early detection rate (less than 50%) and a lack of effective therapies for advanced-stage HCC cause poor prognosis in HCC patients [6]. Therefore, the screening and diagnosis of early-stage HCC in high-risk populations are essential for improving the overall survival rate and reducing treatment costs [7]. According to The Asian Pacific Association for the Study of the Liver (APASL) HCC guidelines, the major risk factors for HCC in China include chronic hepatitis B virus (HBV) infection and liver cirrhosis (LC). Hence, patients with the abovementioned risk factors are recommended for HCC surveillance [8].

Currently, early detection of HCC or monitoring of HCC recurrence mainly relies on ultrasonography (US), serum alpha-fetoprotein (AFP) levels and tissue biopsy [9]. However, these methods show limitations in diagnostic accuracy and sensitivity for early-stage HCC (BCLC 0/A), including approximately 50% sensitivity for AFP and 45% for US alone, even though the sensitivity of combining AFP and US is only 63% [10–12]. For protein markers, although a combination of AFP, des-gamma-carboxyprothrombin (DCP) and lectin-bound AFP (AFP-L3) has higher sensitivity and accuracy than AFP alone [13, 14], neither DCP nor AFP-L3 can improve the performance in distinguishing HBV-associated early HCC and LC with chronic hepatitis B (CHB) [15]. Therefore, it is critical to find novel biomarkers to detect early-stage HCC in high-risk populations.

Cell-free DNA (cfDNA), released from apoptotic, necrotic, and living cells, has been widely used as a plasma-based biomarker in cancer screening and diagnosis [16]. Circulating tumor DNA (ctDNA), a small portion of cfDNA derived from tumor cells, contains genetic defects identical to the tumor cells from which they originated. It has been reported that cfDNA plays an important role in the detection of early-stage HCC [17, 18]. Among the different mechanisms underlying genetic/epigenetic alterations, DNA methylation changes have been reported to contribute to tumorigenesis [19]. Aberrant DNA methylation can occur at the precancerous lesion or early stage and has a strong correlation with metastasis and recurrence [20]. Moreover, abnormal DNA methylation patterns usually cause up/downregulated gene expression, resulting in silencing of tumor suppressor genes involved in hepatocarcinogenesis [21]. Therefore,

aberrant methylation signals can be useful targets for cancer surveillance and treatment.

Currently, cfDNA methylation biomarkers have been applied in the evaluation of HCC diagnosis and prognosis, and different blood-based methylation panels have shown high sensitivity and specificity in clinical practice [22–24]. However, the limited tissue sample number and traditional 450K Methylation BeadChip array used in previous studies may have led to bias and missed some potential markers. In recent studies, almost all HCC patients and normal individuals included were from the United States, and these patients had different genetic backgrounds and major HCC etiologies from the Chinese population [25]. The leading cause of HCC in China is chronic hepatitis B virus (HBV) infection [26], while in the United States is chronic hepatitis C virus (HCV) infection. The carcinogenic mechanism of HBV and HCV is different. HBV can integrate into the host genome and cause the aberrant activation or suppression of the nearby gene [27]. In contrast, HCV disturbs cell growth and apoptosis through some specific proteins, such as HCV core protein and HCV non-structural protein NS3 [28]. As it reported by Lambert et al. [29], the different mechanisms lead to distinct methylation profiles of HBV-related HCC and HCV-related HCC. Therefore, it is essential to identify more HCC methylation biomarkers applicable to Chinese HCC patients. Here, we conducted an HCC-specific DNA methylation biomarker screening and verification study and then selected appropriate hypermethylated DNA markers and protein markers to build and validate a multitarget panel named HepaClear. These results show the potential of the HepaClear panel for clinical HCC screening and diagnosis in the Chinese population with high HCC risk.

Results

Aberrant methylation signatures in 60 paired HCC tissue samples

To screen for genes that contain hypermethylated CpGs in HCC samples, we evaluated genome-wide methylation profiles between 60 pairs of tumor and adjacent tissues by 850K Methylation BeadChip array. The clinical and pathological features of 60 HCC patients are summarized in Additional file 1: Table S1. Etiologically, 41 patients (68.3%) were HBV-positive, and 46 patients (76.7%) had cirrhosis, partly due to the high rate of HBV infection and liver cirrhosis among Chinese HCC patients. For screening early-stage HCC methylation biomarkers, we enrolled 73.4% of HCC patients had early-stage tumors, and 58.3% had low AFP levels (< 20 ng/mL).

We found that ~38.7% of targets were significantly altered ($p < 0.05$, $|\Delta\beta| > 0.1$), while only 3.7% were hypermethylated (Fig. 1A). Principal component analysis

(PCA) revealed that normal tissues clustered with one another, whereas HCC tumor tissues were more dispersed, indicating greater heterogeneity in methylation signatures in HCC (Fig. 1B). The top 1000 ranked targets based on $|\Delta\beta|$ could efficiently distinguish HCC and normal tissues (Fig. 1C). Gene ontology (GO) analyses showed enrichment in terms related to keratinization and positive chemotaxis (Additional file 1: Fig. S1A). We also found enrichment in pathways involved in tight junction, cardiomyopathy, and inflammatory mediator regulation (Additional file 1: Fig. S1B).

From 31,937 significantly hypermethylated target CpG sites, we selected the top 1000 sites with the highest $|\Delta\beta|$ (Additional file 1: Table S2) and found that 16% of them were not included in the 450K Methylation BeadChip array (Fig. 1D). These methylated sites may have the potential to become biological markers, which the traditional 450K Methylation BeadChip array cannot detect because of technical limitations. Among the other 840 target sites included in the 450K Methylation BeadChip array, 300 overlapped with both the top 1000 hypermethylated sites in the TCGA-LIHC dataset and the top 1000 sites in the GSE56588 dataset, indicating good consistency of our study with previous studies (Fig. 1E).

Selection of the markers used in the HCC diagnostic model

To reduce false-positive cases caused by partially methylated CpG sites from normal tissue-derived cDNA in plasma, we sought to identify potential hypermethylation HCC markers, which show little or no apparent methylation in adjacent normal tissues. From the top 1000 sites, we identified 132 sites with $\Delta\beta$ greater than 0.3 and an average β less than 0.1 in adjacent tissues (Additional file 1: Table S3). We further integrated these sites with ROC curve analysis data from 60 pairs of tissues and selected 32 sites with AUC values higher than 0.85 and YIs no less than 0.80 (Additional file 1: Table S4). All 32 sites showed good performance in distinguishing HCC and normal tissues (Additional file 1: Fig. S2). The top 10 methylated CpG sites with the highest $\Delta\beta$ were chosen for further tissue validation assays. The UCSC RefGenes of the 10 methylated CpG sites are cyclin-dependent kinase-like 2 (CDC2-related kinase) (*CDKL2*), ubiquitin specific peptidase 44 (*USP44*), zinc finger family member 783 (*ZNF783*), forkhead box E3 (*FOXE3*), methylene-tetrahydrofolate dehydrogenase 2 (*MTHFD2*),

cyclin-dependent kinase inhibitor 2A (*CDKN2A*), lysyl oxidase-like 3 (*LOXL3*), TLR4 interactor with leucine-rich repeats (*TRIL*), and chromosome 5 open reading frame 49 (*C5orf49*). In addition, cg20172627 is located within the intergenic region (Chr2.25439110) (Additional file 1: Table S4).

The 10 hypermethylated CpG sites were tested by TaqMan qMSP in 60 pairs of HCC and adjacent normal tissues, along with leukocytes and two HCC cell lines. All ten sites showed significantly higher methylation in HCC tissues than in normal tissues (Fig. 2A), and nine of them showed high methylation in both HepG2 and Huh7 cells (Additional file 1: Fig. S3). As a result of ROC curve analysis, AUC values of the ten marker genes ranged from 0.750 to 0.915 among 60 pairs of tissue samples, while YI varied from 0.617 to 0.850 (Table 1). Six of the ten sites (cg14263942, cg12701184, cg14570307, cg15457058, cg07689503, and cg20172627) showed a sensitivity of 78.3–85.0% and a specificity of 96.7–100%.

Determining the marker combination for the HCC diagnostic model

The abovementioned 6 hypermethylated CpG sites were selected for further verification in a pilot study of 150 plasma samples. Two protein markers, AFP and DCP, which have relatively high clinical utility for the diagnosis of HCC in Chinese patients [30] were also included in the biomarker assay. Clinical characteristics from 150 participants are listed in Additional file 1: Table S5. All 8 markers exhibited significantly increased methylation or protein levels in the HCC group (Additional file 1: Fig. S4A). Three methylated CpG sites (cg14263942, cg12701184, and cg14570307) and two protein markers had AUC values of over 0.80 (Fig. 2B), and each of the three methylated CpG sites detected $\geq 65\%$ HCCs with a specificity of $> 93\%$ (Additional file 1: Fig. S4B).

Furthermore, we applied a stepwise logistic regression analysis with backward marker elimination to determine the optimized marker combination. The performance of diagnostic models built from different marker groups is shown in Additional file 1: Table S6. Since the model performance dropped little after eliminating cg15457058, cg07689503 and/or cg20172627 (Fig. 2C, Additional file 1: Table S6), the remaining 3 hypermethylated CpG sites, combined with AFP and DCP, were finally included for multitarget HCC diagnostic model training and

(See figure on next page.)

Fig. 1 Differential methylation between HCC and normal liver tissues detected by 850K Methylation BeadChip array. **A** The volcano plot shows 31,937 hypermethylated CpG sites (red) and 216,246 hypomethylated sites (blue). The x-axis represents the DNA methylation differences, while the y-axis represents the *p* value. **B** Principle component plot of HCC and normal methylation based on beta value. **C** Unsupervised hierarchical clustering of the top 1000 methylated CpG sites. The rows represent different CpG sites, and the columns represent different tissue samples. The color in the heatmap represents the methylation beta values. **D** Hypermethylation locus distribution, whether included in 450K or not. **E** Venn diagram of the hypermethylation markers identified in this study ($n = 840$, 450K included), in TCGA-LIHC ($n = 1000$) and in GSE56588 ($n = 1000$)

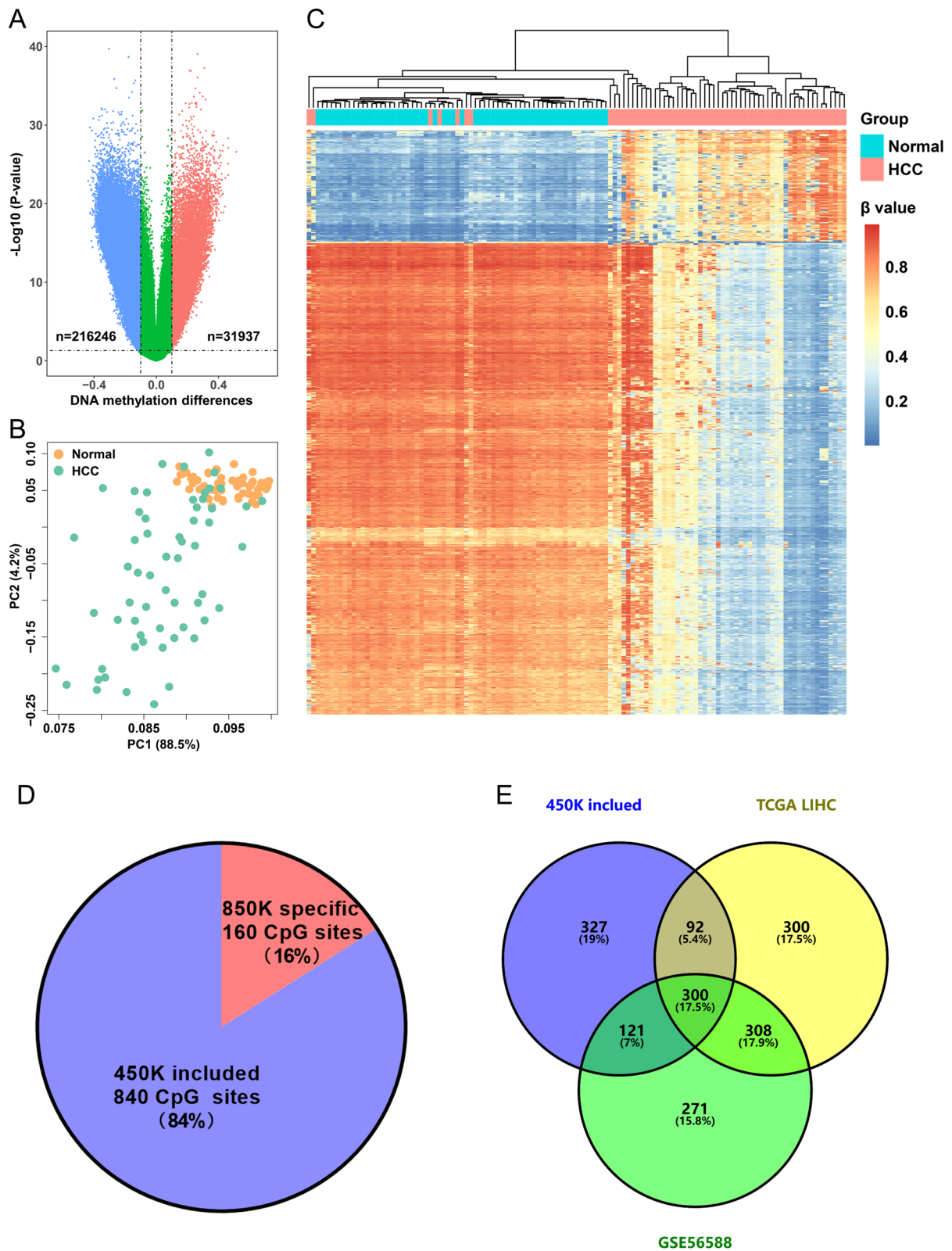


Fig. 1 (See legend on previous page.)

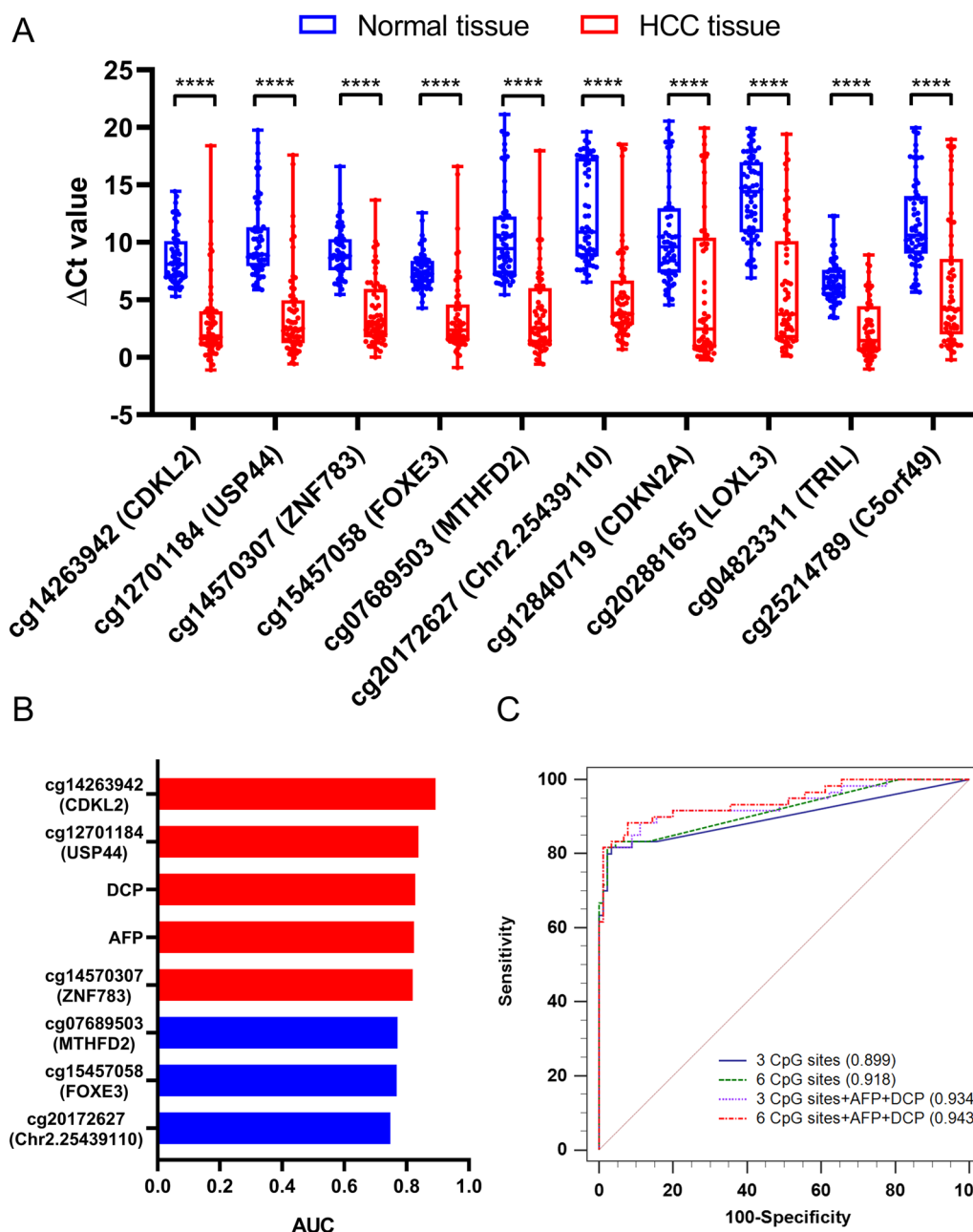


Fig. 2 Identification of biomarkers for the HCC diagnostic model development by qMSP. **A** Methylation levels of 10 candidate methylated CpG sites as quantified by qMSP in 60 pairs of HCC and normal tissues. The y-axis represents methylation levels ($\Delta Ct = Ct_{reference} - Ct_{target}$), in which a lower value represents a relatively higher methylation level. $****p < 0.0001$. **B** Area under the ROC curve (AUC) of eight biomarkers in diagnosing HCC from normal plasma samples. Markers that have AUCs of > 0.8 and < 0.8 are indicated in red and blue, respectively. **C** ROC curves for 4 different biomarker panels in the plasma pilot study. The 4 panels were as follows: 3 methylated CpG sites (cg14263942, cg12701184, and cg14570307), 3 methylated CpG sites with AFP and DCP, 6 methylated CpG sites (cg14263942, cg12701184, cg14570307, cg15457058, cg07689503, and cg20172627), and 6 methylated CpG sites with AFP and DCP. The AUC values for each curve are included in parentheses

verification. These markers were included to develop the diagnostic model named HepaClear.

To evaluate the limit of detection (LOD) of qMSP reaction system for the three methylation biomarkers, we

used positive control (PC) and negative control (NC) composed of different concentrations of Huh7 cell DNA containing methylated target sequences. The quadruplex qMSP assay could fully detect three hypermethylated

Table 1 Distinguishing performance of 10 hypermethylation markers in tissue validation set

Marker region (Gene)	AUC (95% CI)	Cut-off Δ Ct value ($Ct_{tar}-Ct_{ref}$)	Sensitivity (%)	Specificity (%)	Youden index
cg14263942(CDKL2)	0.914 (0.848–0.957)	4.99	85.0	100.0	0.850
cg12701184(USP44)	0.907 (0.841–0.953)	5.35	80.0	100.0	0.800
cg14570307(ZNF783)	0.914 (0.849–0.958)	6.22	80.0	96.7	0.767
cg15457058(FOXE3)	0.879 (0.807–0.932)	5.38	81.7	96.7	0.783
cg07689503(MTHFD2)	0.915 (0.850–0.958)	6.03	78.3	98.3	0.767
cg20172627(Chr2.25439110)	0.891 (0.822–0.941)	6.99	78.3	96.7	0.750
cg12840719(CDKN2A)	0.750 (0.663–0.824)	3.8	61.7	100.0	0.617
cg20288165(LOXL3)	0.870 (0.796–0.924)	6.62	68.3	100.0	0.683
cg04823311(TRIL)	0.887 (0.816–0.937)	3.24	73.3	100.0	0.733
cg25214789(C5orf49)	0.809 (0.728–0.875)	5.27	61.7	100.0	0.617

CpG sites with a minimum of 50 pg Huh7 cell DNA in 0.5% PC (Additional file 1: Figure S5, Table S7).

Construction and validation of the HepaClear panel

In total, 494 participants were recruited to participate further in the plasma biomarker study. Plasma samples from 296 participants were used for diagnostic model construction, and samples from 198 participants were used for model validation. The general clinical characteristics of the study population are shown in Table 2. Among HCC cases, 55% in the training set and 51% in the validation set were diagnosed at BCLC stage 0 and A, while patients with relatively low AFP levels (< 20 ng/mL) accounted for 47.8% and 40.8% in the two sets.

The Ct value of methylated CpG sites and log-transformed AFP/DCP levels from the training set were used to develop a logistic regression algorithm. As shown in Fig. 3A and Table 3, among 138 all-stage HCC patients in the training set, the HepaClear panel yielded an AUC value of 0.915 using a cutoff score of 0.481, with 82.6% (95% CI 75.2–88.5%) sensitivity at 96.2% (95% CI 91.9–98.6%) specificity. The novel HCC biomarker panel showed better performance than AFP at the clinical cutoff value of 20 ng/mL [31], which had a sensitivity of 52.2% (95% CI 43.5–60.7%) with 93.7% (95% CI 87.9–96.5%) specificity (Table 3). The DCP marker also showed lower performance than the HepaClear panel, with a sensitivity of 72.9% (95% CI 64.8–81.0%) and specificity of 92.0% (95% CI 86.9–95.5%) at a cutoff value of 40 mAU/mL (Table 3) [32]. Among 76 early-stage HCC patients in the training set, the HepaClear panel achieved an AUC value of 0.848 and sensitivity of 68.4% (95% CI 56.7–78.6%), which were higher than those of AFP (AUC: 0.705, 34.2% sensitivity, 95% CI 23.7–46.0%) and DCP (AUC: 0.762, 64.7% sensitivity, 95% CI 54.0–75.2%) (Fig. 3B, Table 3). The median value

of the predicted score in the HCC groups was significantly higher than that in the normal groups ($p < 0.01$, Fig. 3C). Meanwhile, there was a gradually increasing trend of the predicted score from healthy, CHB, LC to HCC individuals, consistent with the development of HBV-related HCC.

To further validate the performance of the HCC diagnosis model based on the training set, we used the HepaClear assay on a validation set consisting of an independent cohort. At a cutoff value of 0.481, the HepaClear panel showed 84.7% (95% CI 76.0–91.2%) sensitivity and 92.0% (95% CI 84.8–96.5%) specificity. In contrast, the AFP single biomarker yielded a sensitivity of only 60.2% (95% CI 49.8–70%) and 94.0% (95% CI 87.4–97.8%) specificity using a 20 ng/mL cutoff (Table 3), while the DCP marker yielded a sensitivity of 74.5% (95% CI 63.6–81.9%) and 92.0% (95% CI 84.8–96.5%) specificity at a cutoff value of 40 mAU/mL. Among 50 early-stage HCC patients, the panel showed 72.0% (95% CI 57.7–83.4%) sensitivity, while AFP alone showed a sensitivity of only 48.0% (95% CI 33.7–62.6%), and DCP alone showed a sensitivity of 62.0% (95% CI 46.2–74.6%) (Table 3). Similar to the training set, the predicted scores in the HCC groups were much higher than those in the normal groups ($p < 0.01$, Fig. 3D) and showed a gradual increasing trend.

In addition, we evaluated the performance of the HepaClear panel within different subgroups of the validation set (Table 4). The sensitivity was 83.0% (95% CI 73.8–90.0%) in CHB patients and 85.6% (95% CI 76.7–91.7%) in cirrhosis patients at 90% specificity. For AFP-negative HCC patients, 67.5% (95% CI 50.9–81.4%) obtained positive results using the HepaClear assay. Collectively, these results indicated the advantage of HepaClear over AFP and DCP in differentiating HCC and non-HCC individuals from the high-risk population.

Table 2 Characteristics of study participants in the training and validation sets

	Training set			Validation set			P value*	Healthy (n = 30)	CHB (n = 30)	LC (n = 40)	HCC (n = 98)	LC (n = 40)	CHB (n = 30)	Healthy (n = 30)	P value*	
	HCC (n = 138)	LC (n = 59)	CHB (n = 59)	Healthy (n = 40)	HCC (n = 98)	LC (n = 40)										CHB (n = 30)
<i>Age (year)</i>																
Median (IQR)	56 (50–64)	50 (43–56)	45 (37–54)	42 (34–54)	57 (51–64)	51 (43–57)	< 0.001	48 (35–55)	50 (35–56)	51 (43–57)	57 (51–64)	51 (43–57)	50 (35–56)	48 (35–55)	< 0.001	
<i>Gender</i>																
Male (%)	113 (79.7)	45 (76.3)	30 (50.8)	25 (62.5)	82 (83.7)	30 (75.0)	< 0.001	17 (56.7)	21 (70.0)	30 (75.0)	82 (83.7)	30 (75.0)	21 (70.0)	17 (56.7)	0.02	
Female (%)	25 (20.3)	14 (23.7)	29 (49.2)	15 (37.5)	16 (16.3)	10 (25.0)		13 (43.3)	9 (30.0)	10 (25.0)	16 (16.3)	10 (25.0)	9 (30.0)	13 (43.3)		
<i>Etiology</i>							0.016	NA						NA	0.507	
HBV (%)	129 (93.5)	48 (81.4)	59 (100)	NA	88 (89.8)	35 (87.5)		NA			88 (89.8)	35 (87.5)	30 (100)	NA		
Alcohol (%)	2 (1.4)	4 (6.8)	0		3 (3.1)	2 (5.0)					3 (3.1)	2 (5.0)	0			
NAFLD (%)	2 (1.4)	2 (3.4)	0		2 (2)	0					2 (2)	0	0			
Other (%)	5 (3.7)	5 (8.4)	0		5 (5.1)	3 (7.5)					5 (5.1)	3 (7.5)	0			
<i>Cirrhosis (%)</i>	123 (89.1)	59 (100)	NA	NA	90 (91.8)	40 (100)		NA	NA	40 (100)	90 (91.8)	40 (100)	NA	NA		
<i>AFP (ng/mL)</i>																
Median (IQR)	28.9 (4.8–822.7)	4.2 (2.7–7.6)	4.1 (2.4–6.1)	2.0 (1.3–3.2)	54.9 (3.5–992.9)	4.4 (2.9–9.4)	< 0.001	3.6 (3.0–4.3)	4.3 (3.1–7.1)	4.4 (2.9–9.4)	54.9 (3.5–992.9)	4.4 (2.9–9.4)	4.3 (3.1–7.1)	3.6 (3.0–4.3)	< 0.001	
< 20 (%)	66 (47.8)	53 (89.8)	56 (94.9)	40 (100)	40 (40.8)	36 (90.0)	< 0.001	30 (100)	28 (93.3)	36 (90.0)	40 (40.8)	36 (90.0)	28 (93.3)	30 (100)	< 0.001	
20–400 (%)	31 (22.5)	6 (10.2)	3 (5.1)	0	30 (30.6)	4 (10.0)		0	2 (6.7)	4 (10.0)	30 (30.6)	4 (10.0)	2 (6.7)	0		
> 400 (%)	41 (29.7)	0	0	0	28 (28.6)	0		0	0	0	28 (28.6)	0	0	0		
<i>BCLC stage</i>																
0 (%)	6 (4.3)	NA	NA	NA	6 (6.1)	NA		NA	NA	NA	6 (6.1)	NA	NA	NA		
A (%)	70 (50.7)				44 (44.9)						44 (44.9)					
B (%)	27 (19.6)				14 (14.3)						14 (14.3)					
C (%)	27 (19.6)				31 (31.6)						31 (31.6)					
D (%)	8 (5.8)				3 (3.1)						3 (3.1)					

AFP a-fetoprotein, BCLC Barcelona Clinic Liver Cancer, CHB chronic hepatitis B, LC liver cirrhosis, HBV hepatitis B virus, HCC hepatocellular carcinoma, IQR interquartile range, NAFLD nonalcoholic fatty liver disease

*Statistical tests performed: Chi-square test of independence, Kruskal–Wallis test, and Wilcoxon rank-sum test

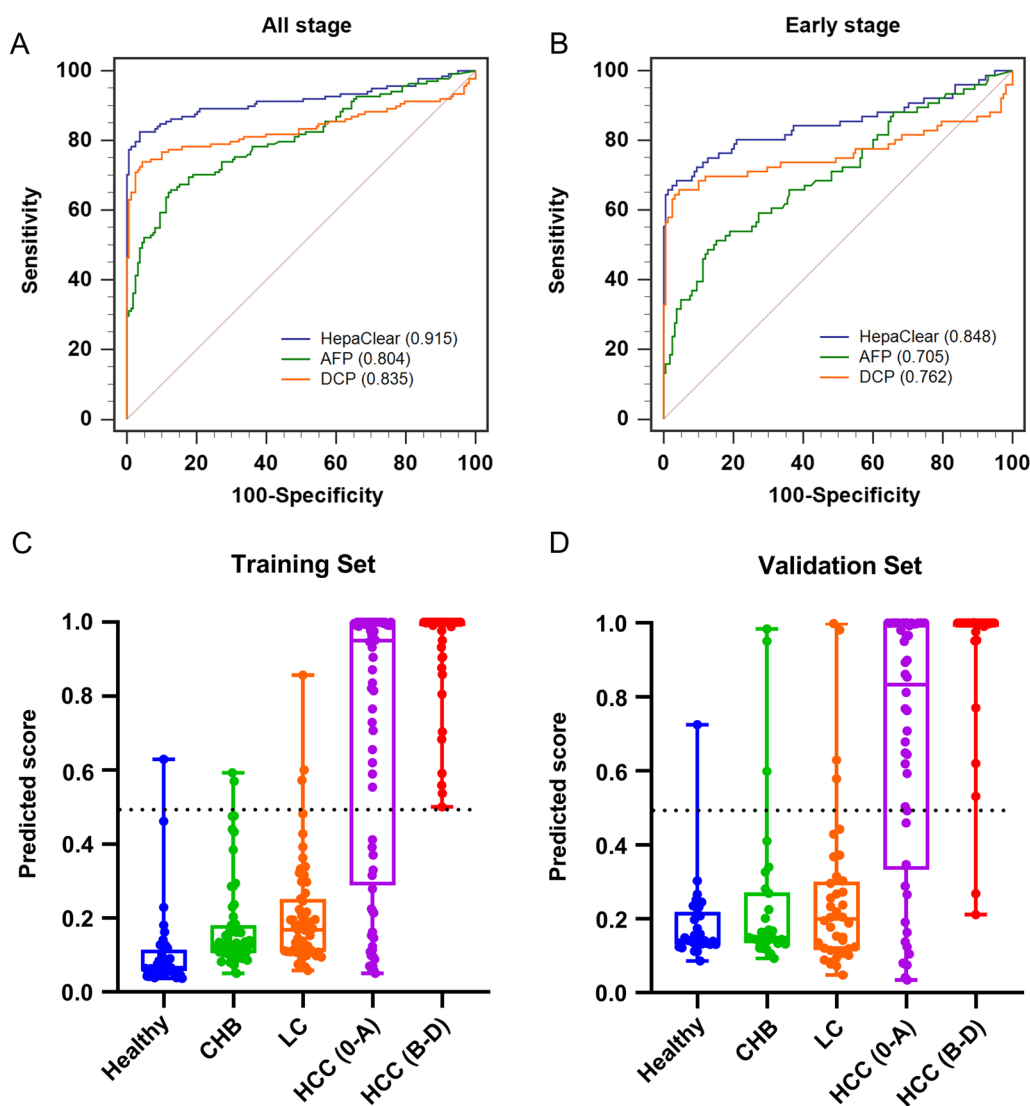


Fig. 3 The performance of HepaClear in the training and validation sets. **A, B** ROC curves of the HepaClear panel compared with AFP and DCP alone in the training group among all-stage patients (**A**) and early-stage HCC patients (**B**). **C, D** Predicted score distributions among healthy controls, CHB, LC and HCC patients in the training group (**C**) and the validation group (**D**)

Table 3 Performance of the HepaClear and comparison with AFP/DCP tests in the training set and the validation set

Marker(s) (cutoff)	Training set				Validation set			
	Specificity (95% CI) (%)	Sensitivity (95% CI) (%)			Specificity (95% CI) (%)	Sensitivity (95% CI) (%)		
		Stage 0-A (n = 76)	Stage B-D (n = 62)	All Stages (n = 138)		Stage 0-A (n = 50)	Stage B-D (n = 48)	All Stages (n = 98)
HepaClear (> 0.481)	96.2 (91.9–98.6)	68.4 (56.7–78.6)	100.0 (94.2–100.0)	82.6 (75.2–88.5)	92.0 (84.8–96.5)	72.0 (57.7–83.4)	95.8 (85.7–99.5)	84.7 (76.0–91.2)
AFP (> 20 ng/mL)	93.7 (87.9–96.5)	34.2 (23.7–46.0)	74.2 (61.5–84.5)	52.2 (43.5–60.7)	94.0 (87.4–97.8)	48.0 (33.7–62.6)	72.9 (58.2–84.7)	60.2 (49.8–70.0)
DCP (> 40 mAU/mL)	92.0 (86.9–95.5)	64.7 (54.0–75.2)	83.9 (72.3–92.0)	72.9 (64.8–81.0)	92.0 (84.8–96.5)	62.0 (46.2–74.6)	87.5 (74.8–95.3)	74.5 (63.6–81.9)

AFP a-fetoprotein, CI confidence interval, DCP des-gamma-carboxyprothrombin

Table 4 Sensitivity and specificity of HepaClear in different subgroups from the validation set

Subgroup	Sensitivity (95% CI) (%)	Specificity (95% CI) (%)
Male (n = 149)	85.4 (75.8–92.2)	91.0 (81.5–96.6)
CHB (n = 153)	83.0 (73.8–90.0)	89.2 (79.1–95.6)
Cirrhosis (n = 130)	85.6 (76.7–91.7)	90.0 (76.3–97.2)
AFP < 20 ng/mL (n = 134)	67.5 (50.9–81.4)	92.6 (85.3–97.0)

AFP α -fetoprotein, CHB chronic hepatitis B, CI confidence interval

Discussion

We developed a cost-effective liquid biopsy HCC panel that demonstrated higher sensitivity and diagnostic accuracy than current biomarker-based tests, including AFP and US. The HepaClear panel showed higher sensitivity than AFP and DCP for early-stage and late-stage HCC. Additionally, the panel performed consistently in both the training set and validation set and yielded high accuracy in different subgroups, including viral etiology and cirrhosis, indicating a valuable opportunity for HCC detection during disease surveillance. In conclusion, our study shows that the HepaClear assay may have superior clinical performance than AFP and has potential for HCC surveillance of high-risk subjects.

Three methylated CpG sites (cg14263942, cg12701184 and cg14570307), for which the RefGenes are *CDKL2*, *USP44* and *ZNF783*, respectively, were finally chosen to build the HCC diagnosis model. *CDKL2* is a member of the cdc2-related serine/threonine kinase subfamily and negatively regulates the cell cycle [33]. *CDKL2* has been reported to be hypermethylated in HCC tissues, and the hypermethylation of *CDKL2* causes decreased *CDKL2* mRNA expression [34]. Low expression of *CDKL2* is positively correlated with tumor cell proliferation and invasion [35]. *USP44* is a member of the deubiquitinating enzyme (DUB) family and is regarded as a key factor involved in DNA double-strand break repair as well as the regulation of mitotic spindle formation and centrosome positioning [36, 37]. It has been reported that *USP44* hypermethylation promotes tumorigenesis and metastasis in multiple cancers [38–40]. In HCC, a lower level of *USP44* expression in HCC samples predicted poor prognosis. *ZNF783* belongs to the zinc finger protein (ZFP) family, which plays a significant role in HCC oncogenesis and progression as a transcription factor [41]. Although the mechanism and biological function of *ZNF783* hypermethylation in HCC tumorigenesis have not been reported, our data demonstrated that *ZNF783* could still be a potential DNA marker in HCC.

In the methylation biomarker discovery and validation phase, several criteria were established to minimize the potential false-positive and false-negative cases in the following plasma cfDNA assay. First, based on a high differential methylation level (≥ 0.3) between HCC and paired adjacent tissues, we selected candidate markers with an average methylation level of ≤ 0.1 to control for methylation signatures contributed from other liver diseases. Normal individuals with liver damage may display much higher levels of liver-derived cfDNA, which could reduce the specificity of liquid biopsy if candidate markers were not sufficiently hypomethylated in normal tissues. Second, we evaluated the performance of each candidate marker in discriminating HCC and control samples using AUC, sensitivity, specificity, and YI analysis. Identifying methylation markers in such criteria can avoid choosing inappropriate biomarkers, which show extremely high $\Delta\beta$ values in a small portion of paired tissues and low $\Delta\beta$ values in relatively more tissues. Third, in addition to adjacent normal tissues, we also used buffy coat tissue (Additional file 1: Fig. S3) as a control sample during biomarker tissue validation. Since leukocyte-derived DNA constitutes most of the plasma circulating DNA pool, even trace amounts of methylated target sites may cause false-positive results.

Although the HepaClear assay showed 82–85% sensitivity for all-stage HCC at $\geq 92\%$ specificity, $\sim 30\%$ of early-stage HCC cases were missed in both the training and validation cohorts. One possible reason can be attributed to the bottleneck of the ctDNA methylation detection method. For some early-stage HCC patients, ctDNA comprises a very small fraction ($< 1\%$) of cfDNA [42] and the actual fraction of methylated DNA fragments would easily fall below the limit of detection of qMSP if cfDNA loss during preanalytical sample processing is considered. On the other hand, HepaClear yielded 14 false-positive cases among the two cohorts. Among these cases, 6 had abnormally elevated AFP or DCP levels, and 10 showed elevated methylation signals from at least one marker gene. In addition to technical errors, these abnormal marker signals may arise as a driver or a consequence of physiological disorder. We tracked the fourteen individuals with false-positive results for six months, seven cases were lost and no new clinical information was obtained. Of the remaining seven individuals who completed follow-up, one developed HCC and the others were consistent with the original case information (hepatitis B or cirrhosis).

Over recent years, many HCC diagnosis models based on cfDNA methylation and protein biomarkers have been reported [43]. Chalasani et al. reported a multitarget HCC panel integrating 4 DNA methylation

markers and 2 protein markers, with a sensitivity of 80% for any-stage HCC and 71% for early-stage HCC at 90% specificity. Our HepaClear panel yielded similar sensitivity and specificity as the multitarget HCC panel, and the model performance was validated in an independent cohort. Compared to the study from Chalasani et al., in which hepatitis C virus (HCV)- or nonalcoholic fatty liver disease (NAFLD)-related cases were enrolled, our study indicates that the HepaClear panel is more applicable to Chinese high-risk populations because over 80% of HCC patients in China were diagnosed as HBV-positive [44]. Luo et al. [45] presented an HCC screening model based on 2321 methylation markers, and the screening model achieved a sensitivity of 84% and specificity of 96% in the validation set. However, cancer screening assays based on next-generation sequencing (NGS) are time-consuming, while the HepaClear assay has the advantages of low financial cost (< 10 dollars per test) and low time cost (1 day from sample processing to data analysis). Furthermore, the DNA methylation detection technology used in our study was the 850K Methylation BeadChip array, which provides more comprehensive coverage of genome-wide methylation sites (> 850,000 methylation sites) than the traditional 450K Methylation BeadChip array. cg12701184, which is included in HepaClear, is a novel methylated site detected by the 850K Methylation BeadChip array. Therefore, as a time-efficient assay with novel methylated CpG sites showing high sensitivity and specificity, HepaClear could have comparable clinical applicability in the Chinese high-risk population. Our results are consistent with a previous study that showed the superiority of DCP to AFP in HCC surveillance and strengthen the viewpoint that DCP has higher sensitivity of detecting HCC with an HBV-positive background.

Despite the encouraging results, there are still limitations to our study. First, the sensitivity of HepaClear in the early stage needs to be improved by optimizing sample processing and adding new biomarkers. Second, the CHB/LC group and healthy group had shorter median ages than the HCC group in both the training and validation cohorts, which may have influenced the biomarker levels. Third, the novel methylated site detected by the 850K Methylation BeadChip array in our study lacks verification from public data. Few data from the 850K Methylation BeadChip array are currently accessible in online databases, such as TCGA or GEO. Therefore, additional studies, including algorithm optimization and prospective, multicenter clinical validation, will further demonstrate the robustness, efficiency, and cost-effectiveness of the HepaClear assay as an HCC screening tool.

Conclusions

In summary, we have developed HepaClear, a novel non-invasive HCC screening assay that integrates DNA methylation markers and protein markers. HepaClear can be effectively implemented on HCC high-risk (CHB/LC) individuals with favorable performance and low cost. Additional studies should be performed to demonstrate the better clinical sensitivity and utility of the HepaClear assay than US plus AFP in future clinical applications.

Methods

Participants and sample collection

In this retrospective study, 644 participants were enrolled from September 2019 to March 2022, including 100 healthy individuals, 119 CHB patients, 129 LC patients and 296 HCC patients. Participants were recruited from three institutions: the Department of Hepatic Surgery, Tianjin First Central Hospital ($n=246$), the Department of Hepatic Surgery, Peking Union Medical College Hospital ($n=248$), and Department of Infection, Central Hospital of Shengli Oilfield ($n=150$).

Tissues used in biomarker discovery and validation were obtained at the time of surgical operation from HCC patients without any previous treatment, including 80 paired HCC tumor and adjacent nontumor liver tissue samples. All tissue samples within this study were stored at -80°C before use, and tumor tissues underwent histopathological evaluation before DNA extraction. Sixty paired tissues were selected for 850K Methylation BeadChip array. Forty of 60 paired tissues, with an additional 20 pairs, were further used for the TaqMan qMSP assay to validate the biomarkers.

Overall, 644 blood samples from the abovementioned participants were obtained, and those from HCC patients were collected after definite diagnosis and before treatment. Peripheral blood (5 mL, collected in EDTA tubes) was centrifuged twice at $1500\times g$ for 15 min within 4 h of collection to isolate plasma, which was then stored at -80°C until experimental analysis. In most cases, 2 mL of plasma was used for cfDNA extraction, and 0.5 mL plasma was used to measure the concentrations of AFP and DCP.

Biomarker discovery and validation

This study was performed in four sequential procedures, according to the flow diagram shown in Fig. 4. First, we used a 850K Methylation BeadChip array (Illumina, San Diego, CA) to detect altered methylation regions in tissue DNA and select candidate hypermethylated CpG sites in HCC tissue DNA. Methylation 450K BeadChip array datasets collected from The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO)

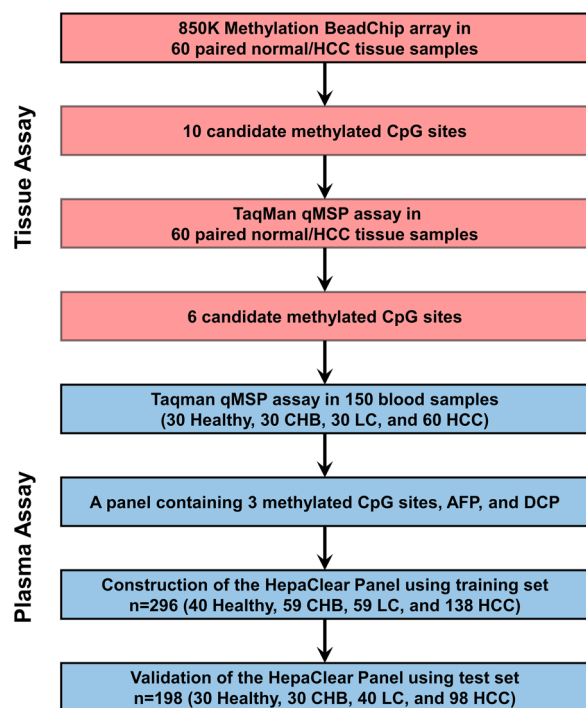


Fig. 4 Flow diagram of the study

database were used for comparison with the results of our study. Second, quantitative methylation-specific PCR (qMSP) was used for tissue validation, and HCC tissue-specific differential methylation sites were further confirmed. Third, selected methylated CpG sites, with HCC protein markers, were tested on 150 plasma samples for the pilot study. Area under the ROC curve (AUC) values and Youden index (YI) at the corresponding biomarker combinations were evaluated. Finally, the selected methylated CpG sites were tested by TaqMan qMSP assay on blood samples in the two groups. A blood-based panel of DNA methylated CpG sites and protein markers (AFP, DCP) was constructed based on the data in the training set and further evaluated based on the data in the validation set.

Tissue/cell/plasma DNA extraction and bisulfite conversion

For tissue and cell samples, genomic DNA was extracted using a QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Circulating cfDNA from ~2 mL plasma samples was extracted using the Magnetic Serum/Plasma Circulating DNA Maxi Kit (TianGen Biotech, Beijing). Tissue and cell DNA was quantified with a NanoDrop 2000 Spectrophotometer (ThermoFisher Scientific, Waltham, MA), and plasma cfDNA was quantified with a Qubit 4.0 fluorometer (ThermoFisher Scientific, Waltham, MA). For bisulfite conversion, 500 ng of tissue DNA or 40 μ L of

plasma cfDNA was converted according to the manufacturer's instructions using an EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA).

Methylation BeadChip array

Converted DNA was used for hybridization on an Illumina Infinium 850K Methylation BeadChip array (Illumina, San Diego, CA), and an Illumina HD methylation assay kit (Shanghai Biotechnology Corporation) was used to assay DNA methylation levels. After quality filtering and data normalization, differentially methylated CpGs between HCC and normal liver tissues were screened using the R package IMA. We used the mean difference in β -value ($\Delta\beta$) between the two abovementioned groups to evaluate the methylation level for each CpG site. CpG sites that showed $|\Delta\beta| > 0.1$ and p value < 0.05 were defined as significantly differentially methylated sites.

TaqMan qMSP for biological tissue validation

For validation of candidate methylated sites, tissue DNA samples were used for qMSP based on a TaqMan probe. The total volume of the qMSP reaction was 20 μ L, containing 2 μ L bisulfite-converted gDNA and 1 unit TaKaRa TaqTM. DNA samples were assayed on an ABI7500 (Applied Biosystems, Foster City, CA) under the following conditions: 94 $^{\circ}$ C for 3 min, 45 cycles of 94 $^{\circ}$ C for 30 s and 60 $^{\circ}$ C for 35 s. Positive control (PC) and negative control (NC) were included in each plate. PC was composed of HepG2 cell DNA, while NC was composed of HIE-2 cell DNA. PCR results were normalized to β -2-microglobulin (*B2M*) amplified from the same sample, and the Δ Ct of each sample was calculated for statistical analysis. Primers and probes for qMSP are summarized in Additional file 1: Table S8.

Plasma biomarker detection

For the methylation marker assay, 10 μ L bisulfite-converted cfDNA was amplified in a volume of 30 μ L containing 2 units PlatinumTM Taq DNA Polymerase, and qPCR was performed with an ABI 7500 system (Applied Biosystems, Foster City, CA) as above. qMSP was performed in triplex assays (2 methylation markers plus internal reference gene *B2M*) for biomarker verification and in quadruplex assays (3 methylation markers plus *B2M*) for model training and validation. PC (Huh7 cell DNA) and NC (HIE-2 cell DNA) were also included, in qMSP. The experiments with plasma samples were performed in a blinded fashion.

Limit of detection (LOD) of quadruplex methylation biomarkers assay

LOD assay of three methylation biomarkers were performed using NC and PC DNA. NC was composed of

HIE-2 cell DNA, while PC was composed of Huh7 cell DNA spiked in HIE-2 cell DNA at ratios of 5%, 1%, 0.5% or 0.25%. All the three methylation biomarkers were found methylated in Huh7 cell DNA, and unmethylated in HIE-2 cell DNA using sanger sequencing. Both PC and NC DNA were diluted to 2 ng/ μ L before use. For LOD assays, NC and different PCs were converted using an EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA), and qMSP with a 30 μ L reaction system containing 5 μ L DNA template was performed with an ABI 7500 system (Applied Biosystems, Foster City, CA). Each bisulfite-treated DNA sample were assayed in twenty replicate wells.

Statistical analysis

Differences between groups were analyzed using the two-tailed Student's *t* test, the Kruskal–Wallis test, or the chi-square test, where appropriate. A *p* value < 0.05 was regarded as statistically significant. To evaluate the clinical performance of each biomarker, receiver operating characteristic (ROC) curves were constructed, and the cutoff for each candidate marker was determined based on the YI. In the plasma pilot study, DNA methylation marker levels, protein marker levels and clinical information were used for HCC diagnosis model construction and validation. The Ct value of methylated CpG sites and the log₂-transformed protein values were collected before data input. All Ct values of undetected methylation markers from qMSP were rounded up to 45, and log₂-transformed AFP and DCP values (calculated as < 0) were rounded up to 0. Logistic regression was used for constructing models of marker combinations, and the logistic regression algorithm developed from HepaClear panel was as follows: score = 30.9062 – 0.3385 × cg14263942 – 0.2114 × cg12701184 – 0.2202 × cg14570307 + 0.3362 × log₂AFP + 0.2656 × log₂DCP. All statistical tests were performed with the Statistical Program for Social Sciences (SPSS 20.0 for Windows, USA) or GraphPad Prism 7 Software (GraphPad, USA). GraphPad Prism 7 and MedCalc Statistical Software version 20.0.4 (MedCalc Software Ltd, Ostend, Belgium) were used for the creation and analysis of graphs.

Abbreviations

HCC	Hepatocellular carcinoma
qMSP	Quantitative methylation-specific PCR
AFP	α -Fetoprotein
DCP	Des-gamma-carboxyprothrombin
APASL	Asian Pacific Association for the Study of the Liver
HBV	Hepatitis B virus
HCV	Hepatitis C virus
LC	Liver cirrhosis
US	Ultrasonography
BCLC	Barcelona Clinic Liver Cancer
CHB	Chronic hepatitis B

cfDNA	Cell-free DNA
ctDNA	Circulating tumor DNA
TCGA	The Cancer Genome Atlas
GEO	Gene Expression Omnibus
ROC	Receiver operating characteristic
AUC	Area under ROC curve
YI	Youden index
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
B2M	β -2-Microglobulin
CDKL2	Cyclin-dependent kinase-like 2
USP44	Ubiquitin specific peptidase 44
ZNF783	Zinc finger family member 783
NAFLD	Non-alcoholic fatty liver disease
NGS	Next generation sequencing
PC	Principal component

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-023-01508-7>.

Additional file 1. Figure S1. GO and KEGG pathway enrichment analysis of differentially methylated CpG sites. **Figure S2.** Performance of 32 candidate CpG sites identified from methylome profiling data. **Figure S3.** Methylation levels of 10 candidate CpG sites in two HCC cell lines and leukocytes using qMSP. **Figure S4.** Methylation/Protein levels and diagnostic performance of eight candidate markers in plasma pilot test set. **Figure S5.** Amplification curves of cg14263942, cg12701184 and cg14570307 within quadruplex assays. **Table S1.** Demographic and clinical characteristics of HCC patients for biomarker screening and tissue validation. **Table S2.** Top-1000 hypermethylation markers screened from genome-wide methylation profiling. **Table S3.** List of 132 hypermethylation markers with *p* < 0.05, $\Delta\beta$ > 0.3 and β_{normal} < 0.1. **Table S4.** List of 32 markers further screened from Table S3, with AUC > 0.85 and Youden Index (YI) \geq 0.8. **Table S5.** Characteristics of study participants in plasma pilot study. **Table S6.** Performance of different biomarker combinations in 150 plasma samples. **Table S7.** Limit of detection (LOD) of three methylation markers in HepaClear panel. **Table S8.** List of primers and probes for Taqman qMSP.

Acknowledgements

This work was supported by the Tianjin Natural Science Foundation (20JCYBJC01310 and 21JCYBJC00320), the Tianjin Science and technology committee (19ZXDBSY00010), and the Tianjin Health Science and technology project (TJWJ2021ZD002 and ZC20218).

Author contributions

Conception and design: YB, HTZ, and YMZ; Acquisition of data: YB, JX, DQL, and XYZ; Analysis and interpretation of data: YB, JX, DQL, XYZ, DPC, FCX, LMH, XTY; Drafting the article: YB, JX, DQL, XYZ, HTZ, and YMZ. All authors revised the manuscript and approved the final version.

Funding

This work was supported by the Tianjin Natural Science Foundation (20JCYBJC01310 and 21JCYBJC00320), the Tianjin Science and technology committee (19ZXDBSY00010), and the Tianjin Health Science and technology project (TJWJ2021ZD002 and ZC20218).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The study was approved by the ethic committee of Tianjin First Central Hospital (2021N066KY), Peking Union Medical College Hospital (HS-2108) and Shengli Oilfield Central Hospital (Q/ZXY-YZYWB-LL202160). All collection

and processing of clinical samples and date were in accordance with the principles of the Declaration of Helsinki. Written informed consents were provided by all participants.

Consent for publication

All authors give consent for the publication of the manuscript.

Competing interests

Deqiang Li, Longmei Huang and Xiaotian Yu are employees of Hangzhou New Horizon Health Technology Co., Ltd, Hangzhou, China. The remaining authors have nothing to declare.

Received: 1 December 2022 Accepted: 19 May 2023

Published online: 12 June 2023

References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–49.
- Wong MCS, Huang JLW, George J, Huang J, Leung C, Eslam M, et al. The changing epidemiology of liver diseases in the Asia-Pacific region. *Nat Rev Gastroenterol Hepatol*. 2019;16(1):57–73.
- Cao W, Chen HD, Yu YW, Li N, Chen WQ. Changing profiles of cancer burden worldwide and in China: a secondary analysis of the global cancer statistics 2020. *Chin Med J (Engl)*. 2021;134(7):783–91.
- Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Niksic 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.
- Huang J, Yan L, Cheng Z, Wu H, Du L, Wang J, et al. A randomized trial comparing radiofrequency ablation and surgical resection for HCC conforming to the Milan criteria. *Ann Surg*. 2010;252(6):903–12.
- Lin J, Zhang H, Yu H, Bi X, Zhang W, Yin J, et al. Epidemiological characteristics of primary liver cancer in Mainland China From 2003 to 2020: a representative multicenter study. *Front Oncol*. 2022;12:906778.
- 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.
- Xie DY, Ren ZG, Zhou J, Fan J, Gao Q. 2019 Chinese clinical guidelines for the management of hepatocellular carcinoma: updates and insights. *Hepatobiliary Surg Nutr*. 2020;9(4):452–63.
- Tayob N, Lok AS, Do KA, Feng Z. Improved detection of hepatocellular carcinoma by using a longitudinal alpha-fetoprotein screening algorithm. *Clin Gastroenterol Hepatol*. 2016;14(3):469–475 e2.
- Marrero JA, Feng Z, Wang Y, Nguyen MH, Befeler AS, Roberts LR, et al. Alpha-fetoprotein, des-gamma carboxyprothrombin, and lectin-bound alpha-fetoprotein in early hepatocellular carcinoma. *Gastroenterology*. 2009;137(1):110–8.
- Singal AG, Conjeevaram HS, Volk ML, Fu S, Fontana RJ, Askari F, et al. Effectiveness of hepatocellular carcinoma surveillance in patients with cirrhosis. *Cancer Epidemiol Biomark Prev*. 2012;21(5):793–9.
- 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):1706–1718 e1.
- Berhane S, Toyoda H, Tada T, Kumada T, Kagebayashi C, Satomura S, et al. Role of the GALAD and BALAD-2 serologic models in diagnosis of hepatocellular carcinoma and prediction of survival in patients. *Clin Gastroenterol Hepatol*. 2016;14(6):875–886 e6.
- Caviglia GP, Abate ML, Petrini E, Gaia S, Rizzetto M, Smedile A. Highly sensitive alpha-fetoprotein, Lens culinaris agglutinin-reactive fraction of alpha-fetoprotein and des-gamma-carboxyprothrombin for hepatocellular carcinoma detection. *Hepatol Res*. 2016;46(3):E130–5.
- Song T, Wang L, Su B, Zeng W, Jiang T, Zhang T, et al. Diagnostic value of alpha-fetoprotein, Lens culinaris agglutinin-reactive alpha-fetoprotein, and des-gamma-carboxyprothrombin in hepatitis B virus-related hepatocellular carcinoma. *J Int Med Res*. 2020;48(3):300060519889270.
- Ng CKY, Di Costanzo GG, Terracciano LM, Piscuoglio S. Circulating cell-free DNA in hepatocellular carcinoma: current insights and outlook. *Front Med (Lausanne)*. 2018;5:78.
- Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med*. 2014;6(224):224ra24.
- Liu X, Ren J, Luo N, Guo H, Zheng Y, Li J, et al. Comprehensive DNA methylation analysis of tissue of origin of plasma cell-free DNA by methylated CpG tandem amplification and sequencing (MCTA-Seq). *Clin Epigenet*. 2019;11(1):93.
- Jiang P, Chan KCA, Lo YMD. Liver-derived cell-free nucleic acids in plasma: biology and applications in liquid biopsies. *J Hepatol*. 2019;71(2):409–21.
- Baylin SB, Jones PA. A decade of exploring the cancer epigenome—biological and translational implications. *Nat Rev Cancer*. 2011;11(10):726–34.
- Yang B, Guo M, Herman JG, Clark DP. Aberrant promoter methylation profiles of tumor suppressor genes in hepatocellular carcinoma. *Am J Pathol*. 2003;163(3):1101–7.
- Chalasanani NP, Ramasubramanian TS, Bhattacharya A, Olson MC, Edwards VD, Roberts LR, et al. A novel blood-based panel of methylated DNA and protein markers for detection of early-stage hepatocellular carcinoma. *Clin Gastroenterol Hepatol*. 2021;19(12):2597–2605 e4.
- Kisiel JB, Dukek BA, Kanipakam VSR, Ghosh 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.
- 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.
- Makarova-Rusher OV, Altekruse SF, McNeel TS, Ulahannan S, Duffy AG, Graubard BI, et al. Population attributable fractions of risk factors for hepatocellular carcinoma in the United States. *Cancer*. 2016;122(11):1757–65.
- Villanueva A. Hepatocellular carcinoma. *N Engl J Med*. 2019;380(15):1450–62.
- D'Souza S, Lau KC, Coffin CS, Patel TR. Molecular mechanisms of viral hepatitis induced hepatocellular carcinoma. *World J Gastroenterol*. 2020;26(38):5759–83.
- Chidambaramathan-Reghupaty S, Fisher PB, Sarkar D. Hepatocellular carcinoma (HCC): epidemiology, etiology and molecular classification. *Adv Cancer Res*. 2021;149:1–61.
- Lambert MP, Paliwal A, Vaissiere T, Chemin I, Zoulim F, Tommasino M, et al. Aberrant DNA methylation distinguishes hepatocellular carcinoma associated with HBV and HCV infection and alcohol intake. *J Hepatol*. 2011;54(4):705–15.
- Feng X, Song P, Bie P, Jiang P, Ma K, Li X, et al. Des-gamma-carboxyprothrombin plasma level in diagnosis of hepatocellular carcinoma in a Chinese Population Undergoing Surgery. *Med Sci Monit*. 2016;22:1663–72.
- Gupta S, Bent S, Kohlwees J. Test characteristics of alpha-fetoprotein for detecting hepatocellular carcinoma in patients with hepatitis C. A systematic review and critical analysis. *Ann Intern Med*. 2003;139(1):46–50.
- Ji J, Wang H, Li Y, Zheng L, Yin Y, Zou Z, et al. Diagnostic evaluation of des-gamma-carboxy prothrombin versus alpha-fetoprotein for hepatitis B virus-related hepatocellular carcinoma in China: a large-scale, multicentre study. *PLoS ONE*. 2016;11(4):e0153227.
- Malumbres M, Harlow E, Hunt T, Hunter T, Lahti JM, Manning G, et al. Cyclin-dependent kinases: a family portrait. *Nat Cell Biol*. 2009;11(11):1275–6.
- Zhou Y, Qiu XP, Li ZH, Zhang S, Rong Y, Yang GH, et al. Clinical significance of aberrant cyclin-dependent kinase-like 2 methylation in hepatocellular carcinoma. *Gene*. 2019;683:35–40.
- Fang CL, Uen YH, Chen HK, Hseu YC, Lin CC, Hung ST, et al. Loss of cyclin-dependent kinase-like 2 predicts poor prognosis in gastric cancer, and its overexpression suppresses cells growth and invasion. *Cancer Med*. 2018;7(7):2993–3002.
- Chen Y, Zhao Y, Yang X, Ren X, Huang S, Gong S, et al. USP44 regulates irradiation-induced DNA double-strand break repair and suppresses tumorigenesis in nasopharyngeal carcinoma. *Nat Commun*. 2022;13(1):501.
- Nijman SM, Luna-Vargas MP, Velds A, Brummelkamp TR, Dirac AM, Sixma TK, et al. A genomic and functional inventory of deubiquitinating enzymes. *Cell*. 2005;123(5):773–86.

38. Chen X, Wu X, Lei W. USP44 hypermethylation promotes cell proliferation and metastasis in breast cancer. *Future Oncol.* 2021;17(3):279–89.
39. Londra D, Mastoraki S, Bournakis E, Zavidou M, Thanos A, Rampias T, et al. USP44 promoter methylation in plasma cell-free DNA in prostate cancer. *Cancers Basel.* 2021;13(18):4607.
40. Zhang Y, Foreman O, Wigle DA, Kosari F, Vasmatzis G, Salisbury JL, et al. USP44 regulates centrosome positioning to prevent aneuploidy and suppress tumorigenesis. *J Clin Invest.* 2012;122(12):4362–74.
41. Li X, Han M, Zhang H, Liu F, Pan Y, Zhu J, et al. Structures and biological functions of zinc finger proteins and their roles in hepatocellular carcinoma. *Biomark Res.* 2022;10(1):2.
42. Campos-Carrillo A, Weitzel J, Sahoo P, Rockne R, Mokhnatkin J, Murtaza M, et al. Circulating tumor DNA as an early cancer detection tool. *Pharmacol Ther.* 2020;207:107458.
43. Tran NH, Kisiel J, Roberts LR. Using cell-free DNA for HCC surveillance and prognosis. *JHEP Rep.* 2021;3(4):100304.
44. Wang M, Wang Y, Feng X, Wang R, Wang Y, Zeng H, et al. Contribution of hepatitis B virus and hepatitis C virus to liver cancer in China north areas: experience of the Chinese National Cancer Center. *Int J Infect Dis.* 2017;65:15–21.
45. Luo B, Ma F, Liu H, Hu J, Rao L, Liu C, et al. Cell-free DNA methylation markers for differential diagnosis of hepatocellular carcinoma. *BMC Med.* 2022;20(1):8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

