LETTER TO THE EDITOR

Genome-scale profiling of circulating cell-free DNA signatures for early detection of hepatocellular carcinoma in cirrhotic patients

Cell Research (2021) 31:589–592; <https://doi.org/10.1038/s41422-020-00457-7>

Dear Editor,

Hepatocellular carcinoma (HCC) is the second most deadly cancer worldwide.¹ Cirrhosis of different causes predisposes patients to HCC, increasing the annual HCC incidence by 2%–4%.¹ The development of cirrhosis facilitates a series of genetic or epigenetic changes, resulting in the formation of dysplastic nodules, a premalignant stage in HCC.¹ HCC diagnosis at an early stage contributes to an improved prognosis with the possibility of curative treatment. Due to the low accuracy of current diagnostic methods, it is urgently needed to explore new non-invasive strategies for early HCC diagnosis in cirrhotic patients. Hence, we collected cell-free DNA (cfDNA) samples from a total of 2250 patients with liver cirrhosis (LC), 508 with HCC, and 476 healthy controls (CTRL), from 13 hospitals in 11 provinces of China, and randomly assigned them to training, validation and test cohorts for development and evaluation of diagnostic model. We employed a state-of-the-art next-generation sequencing (NGS) technology to acquire genome-wide 5-hydroxymethylcytosine (5-hmc),² nucleosome footprint (NF),³ 5' end motif⁴ and fragmentation^{5–7} profiles of cfDNAs from all enrolled patients. Using a logistic regression method, we constructed a weighted diagnostic model based on the performance of these four features.

A total of 3204 individuals were finally included in our study according to our exclusion criteria and randomly assigned to training/validation/test cohorts (Supplementary information, Fig. S1a). For the 2247 LC patients from 13 hospitals, the median age was 56, 76% were males, 83.3% were diagnosed with HBV infection, and Child-Pugh scores A, B, and C were 80.9%, 15.8%, and 0.2%, respectively (Supplementary information, Table S1). For the 481 HCC patients, the median age was 51, 87.5% were males, 79% were diagnosed with HBV infection, 64% were reported with LC, and BCLC stage 0/A, B/C patients accounted for 41.8% and 58.2%, respectively (Supplementary information, Table S1). For the 476 CTRL, the median age was 54, 46.4% were males, and no one had history of chronic hepatitis virus infection, LC or HCC.

Individual features were profiled to see their respective relevance in distinguishing HCC from LC and CTRL. We found distinct 5-hmc modification landscapes in HCC and CTRL (Supplementary information, Fig. S1b). NF analysis showed that the read coverage of differential genes held power to distinguish HCC from CTRL (Supplementary information, Fig. S1c). Meanwhile, 256 end motif patterns also exhibited as a potential classification parameter for HCC against CTRL (Supplementary information, Fig. S1d). Additionally, our data showed that the fragment size of HCC cfDNA was more variable and much shorter (median size < 150 bp) in HCC (Supplementary information, Fig. S1e). Collectively, all four features showed promising diagnostic potential for HCC.

However, LC patients were much different from CTRL, which significantly increased the diagnostic difficulty.

By integrating the above four genomic features (Supplementary information, Fig. S1f), HIFI (5-Hydroxymethylcytosine/motif/Fragmentation/nucleosome footprint) method held a strong diagnostic value in differentiating HCC from LC, yielding a sensitivity of 95.79% and a specificity of 95.00% for HCC in validation set (95 HCC and 100 LC) and a sensitivity of 95.42% and a specificity of 97.83% in test set (148 HCC and 1800 LC). The differentiation power of the HIFI method for HCC vs LC (validation: AUC = 0.995 [0.990–1.000], test: AUC = 0.996 [0.992–0.999]) was superior to AFP (validation: AUC = 0.845 [0.788–0.897], test: AUC = 0.826 [0.785–0.864]) (Fig. 1a; Supplementary information, Fig. S2a, b). In addition, HIFI method achieved promising diagnostic performances in differentiating HCC from non-HCC (LC + CTRL) and CTRL (Supplementary information, Fig. S2c, f).

In differentiating HCC from LC, NF was the best individual diagnostic feature in both validation (AUC = 0.971 [0.948–0.993]) and test sets (AUC = 0.973 [0.953–0.993]), but still inferior to HIFI method (validation, AUC = 0.995 [0.990–1.000]; test, AUC = 0.996 [0.992–0.999]) (Supplementary information, Fig. S2g, h). Furthermore, the three remaining genomic features performed much better than AFP (validation, AUC = 0.845 [0.788–0.897]; test, AUC = 0.826 [0.785–0.864]). The superior diagnostic power of HIFI method over individual genomic features was also seen when differentiating HCC from non-HCC or CTRL (Supplementary information, Fig. S2i–l). To obtain the best model, we tried different combinations of genome features, and evaluated their diagnostic powers. The results showed that HIFI outperformed any combinations of genome features (Supplementary information, Fig. S2m). Furthermore, HIFI method also achieved a high accuracy in 5-hmc misdiagnosed cases (Supplementary information, Fig. S2n). In our previously published study,⁸ the weighted diagnosis score (wd-score) calculated with 32 5-hmc markers derived mainly from gene bodies showed superior diagnostic power in HCC. To this end, we evaluated its diagnostic potential in our validation and test sets (HCC vs non-HCC), respectively, which showed inferior diagnostic accuracy than HIFI score (Supplementary information, Fig. S2o).

cfDNA concentration varied among individuals, with the highest in HCC group (Supplementary information, Fig. S3a). To test the feasibility in clinical application, we then evaluated the consistency of 5hmc-seq/whole-genome sequencing (WGS) data between duplicated HCC/LC samples across a series of input DNA amount (1, 2, 5, and 10 ng) and different batches (different technicians and reagent lots) (Supplementary information, Fig. S3b), and observed high correlations (mean Pearson's $r >$

0.90) among duplicate samples and different batches, indicating the robustness and feasibility of our 5hm-seq/low-pass WGS and the feasibility for clinical usage. In addition, the diagnostic performance of HIFI method was not affected by cfDNA

concentration (Supplementary information, Fig. S3c), although the fragment size was significantly longer in LC patients with low cfDNA concentration (<10 ng/mL) (Supplementary information, Fig. S3d).

Fig. 1 The diagnostic value of HIFI method. **a** ROC curves of HIFI method and AFP for HCC patients vs LC patients in the test set. **b** ROC curves of HIFI method for HCC patients vs LC patients in the 11 hospital-based validation + test sets. Hospitals 4, 9 and 10 were integrated into one due to their small size. **c** HIFI scores of all enrolled patients were shown. The cutoff value of HIFI score was 0.38. Upper: information of HCC patients. Lower: information of LC patients. **d** Proportions of positive and negative calling by HIFI method in LC patients with AFP < 20 µg/L, AFP 20–400 µg/L, and AFP > 400 µg/L, in the validation + test sets. **e** Proportions of positive and negative calling by HIFI method and AFP in LC patients with different ages, Child-Pugh stages, and HBV infection status in validation + test sets. **f, g** Proportions of positive and negative calling by HIFI method in HCC patients with different AFP and PIVKA-II levels in the validation + test sets. **h–j** Proportions of positive and negative calling by HIFI/AFP/PIVKA-II in HCC patients with different BCLC stages, AJCC stages, and tumor sizes in validation + test sets.

The 2247 LC patients were recruited from 11 different provinces of China (Supplementary information, Fig. S3e). To examine the possible bias due to geographic distribution, we rebuilt 11 independent test sets based on sample origin (hospitals 4, 9, and 10 were integrated into one set) and assessed the performance of each genome feature. NF exhibited a stable and the best performance in all 11 test sets with AUC > 0.97. Motif also showed a comparatively high diagnostic capability in different hospitals. The diagnostic capability of 5-hmc varied slightly with AUC values ranging from 0.893 [0.857–0.929] to 0.923 [0.900–0.947]. Meanwhile, fragmentation also achieved a good classification performance with AUC values surpassing 0.900 in all hospitals (Supplementary information, Figs. S3f–i). HIFI method exhibited a higher diagnostic accuracy than any individual features and a consistently good performance in all 11 test sets (AUC > 0.99, accuracy > 99%) (Fig. 1b), indicating that the HIFI method was stable and robust for clinical application regardless of demographic distribution.

Overall, HIFI score showed a gradually increasing trend from CTRL to HCC patients (Fig. 1c). In the LC cohort (Supplementary information, Fig. S4a), the difference of HIFI scores between patients with or without ascites reached statistical significance ($P = 0.0413$). However, it displayed negative correlation with Child-Pugh stage (A vs B + C, $P = 3.665e-05$), total bilirubin level (< 34 µmol/L vs ≥ 34 µmol/L, $P = 0.003665$), albumin level (> 35 vs ≤ 35, $P = 0.0002762$), but was positively correlated with platelet level (< 125 K/mL vs 125–350 K/mL, $P = 0.03885$). In HCC cohort (Supplementary information, Fig. S4b), patients with larger tumor size (> 5 cm, $P = 2.784e-05$) and higher PIVKA-II level (> 40, $P = 0.001591$) had a significantly higher HIFI score. HCC patients with cirrhosis history had much higher HIFI score than those without ($P = 0.03194$).

In the assessment of diagnostic value of HIFI method in LC patient subgroups (Supplementary information, Table S2), HIFI method exhibited a high accuracy in AFP-positive patients who might be misclassified as HCC (157 of 163 patients for 20 < AFP < 400 µg/L, 96.3%; 16 of 16 patients for AFP > 400 µg/L, 100%) (Fig. 1d). Not surprisingly, HIFI method also accurately distinguished between AFP-negative HCC and LC (Supplementary information, Fig. S4c). Furthermore, HIFI method showed a consistent and higher accuracy than AFP in LC patients regardless of age, Child-Pugh stage and HBV infection status (Fig. 1e). However, the accuracy of HIFI method was slightly decreased in LC patients with diabetes and hypertension (χ^2 test, $P_{\text{diabetes}} = 0.063$, $P_{\text{hypertension}} = 0.003$), which were predisposing factors for HCC^{9,10} (Supplementary information, Fig. S4d). In HCCs (Supplementary information, Table S3), HIFI method achieved high diagnostic accuracies in both AFP-negative (92 of 98 patients, 93.9%) and PIVKA-II-negative patients (30 of 33 patients, 90.9%) (Fig. 1f, g). The accuracies of HIFI method were irrespective of age, cirrhosis history and HBV infection status, outperforming AFP and PIVKA-II in any HCC patient subgroups (Supplementary information, Fig. S4e–j). Most importantly, HIFI method was powerful for detecting early-stage HCC (BCLC: A, 94.4%/AJCC: I, 94.7%) and even very early-stage HCC (BCLC: 0, 88.9%). As comparison, the diagnostic accuracies of AFP/PIVKA-II and individual features were much lower than HIFI method for early HCC (Fig. 1h, i; Supplementary information, Fig. S4k). The

combination of AFP/PIVKA-II with HIFI could not further improve the diagnostic accuracy (Supplementary information, Fig. S4l). HIFI method showed great performance in diagnosing small HCC (diameter < 3 cm, 96.7%), whereas the performance of AFP and PIVKA-II was positively correlated with tumor size (Fig. 1j). In AFP/PIVKA-II misdiagnosed HCC cases, HIFI achieved a high accuracy (93.9%/90.9%) (Supplementary information, Fig. S4m). 5-hmc modifications of human genome affect gene expression activity,⁸ while NF, which belongs to cfDNA fragmentomics, reflects transcription initiation and elongation.³ In our study, we found that multiple 5-hmc and NF markers were strongly enriched in HCC. The other two fragmentomic markers, end motif and fragment size carrying tissue-of-origin information,⁴ were also included in HIFI method. However, their underlying mechanisms in hepatocarcinogenesis need to be further studied.

In summary, based on the unique genome features of cfDNA, we performed a large-scale, multi-center study and developed a novel integrated method for HCC diagnosis. HIFI method showed a high accuracy for differentiating HCC from LC and CTRL groups, including AFP/PIVKA-II-negative HCC, reinforcing the idea that HIFI method had great potentials as a new strategy for HCC early diagnosis and surveillance. HIFI method fills an urgent need to identify early-stage HCC which is qualified for curative treatments. However, there are a few limitations of our study. The patients in our cohort were all from China, therefore, the capability of HIFI method to diagnose HCC in other ethnicities will need more investigations. Second, the number of CTRL in the test set is rather small as our main aim is to distinguish HCC from LC. Hence, the conclusion that HIFI also performed excellently in healthy control group needs to be further proved in a larger cohort. Collectively, considering its high accuracy, non-invasiveness and universality, HIFI method has a great potential for public health screening in high-risk population, especially when both equipment and infrastructure required for HCC imaging screening are not available.

ACKNOWLEDGEMENTS

This work was supported by the State Key Project for Infectious Diseases (2018ZX10732202-001, 2018ZX10302207-004, 2018ZX10301202-006), the National Research Program of China (2017YFA0505803, 2017YFC0908100), the National Natural Science Foundation of China (81790633, 81672860, 81702298, 61922047, 81830054 and 91859205), the Natural Science Foundation of Shanghai (17ZR143800).

AUTHOR CONTRIBUTIONS

H.-Y.W., J.-L.H. and L.W. had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Concept and experimental design: L.C., G.K.A., B.Z., J.-F.L., J.B., L.-T.D., Y.-S.Q., R.F., X.-L.L. Acquisition, analysis, or interpretation of data: all authors. Drafting of the manuscript: L.C., G.K.A., B.Z. Critical revision of the manuscript for important intellectual content: all authors. Statistical analysis: Q.-Z.Z. and F.-M.S. Obtaining funding: L.W., J.-L.H. and H.-Y.W. Administrative, technical, or material support: Y.-C.W., X.-T.F., G.-Q.J., G.-H.D., C.-Y.W., Q.G., F.S., H.-P.H., Y.-N.Y., J.Z., Y.-H.G., J.X., H.-D.Y., M.-F.L., Y.-L.Y., Y.-J.G., J.S., C.-X.Z., Y.W., H. Wang, F.K., J.-M.C., H. Wen, D.Z., Y.Y. and C.-X.W. Supervision: L.W., J.-L.H. and H.-Y.W.

ADDITIONAL INFORMATION

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41422-020-00457-7>.

Competing interests: The authors declare no competing interests.

Lei Chen^{1,2}, Ghassan K. Abou-Alfa^{3,4}, Bo Zheng^{1,2}, Jing-Feng Liu⁵, Jian Bai⁶, Lu-Tao Du^{7,8}, Yun-Song Qian⁹, Rong Fan¹⁰, Xiao-Long Liu⁵, Lin Wu⁶, Jin-Lin Hou¹⁰, Hong-Yang Wang^{1,2,11,12} and The PreCar Team
¹National Center for Liver Cancer, Shanghai 210822, China; ²International Cooperation Laboratory on Signal Transduction, Eastern Hepatobiliary Surgery Institute, Second Military Medical University, Shanghai 200438, China; ³Memorial Sloan Kettering Cancer Center, New York, NY, USA; ⁴Weill Medical College at Cornell University, New York, NY, USA; ⁵The United Innovation of Mengchao Hepatobiliary Technology Key Laboratory of Fujian Province, Mengchao Hepatobiliary Hospital of Fujian Medical University, Fuzhou, Fujian 350025, China; ⁶Berry Oncology Corporation, Beijing 100102, China; ⁷Department of Clinical Laboratory, The Second Hospital of Shandong University, 247 Beiyuan Street, Jinan, Shandong 250033, China; ⁸The Clinical Research Center of Shandong Province for Clinical Laboratory, 247 Beiyuan Street, Jinan, Shandong 250033, China; ⁹Hepatology Department, Ningbo Hwamei Hospital, University of Chinese Academy of Sciences, Ningbo, Zhejiang 315010, China; ¹⁰Department of Infectious Diseases, State Key Laboratory of Organ Failure Research, Guangdong Provincial Key Laboratory of

Viral Hepatitis Research, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong 510515, China; ¹¹Key Laboratory of Signaling Regulation and Targeting Therapy of Liver Cancer (SMMU), Ministry of Education, Shanghai 200438, China and ¹²Shanghai Key Laboratory of Hepatobiliary Tumor Biology (EHBH), Shanghai 200438, China

A list of the PreCar Team members and their affiliations appears at the end of the paper.

These authors contributed equally: Lei Chen, Ghassan K. Abou-Alfa, Bo Zheng, Jing-Feng Liu, Jian Bai, Lu-Tao Du, Yun-Song Qian, Rong Fan, Xiao-Long Liu

Correspondence: Lin Wu (wulin@berryoncology.com) or Jin-Lin Hou (jlhou@smu@163.com) or Hong-Yang Wang (hywangk@vip.sina.com)

REFERENCES

- Villanueva, A. N. *Engl. J. Med.* **380**, 1450–1462 (2019).
- Branco, M. R., Ficiz, G. & Reik, W. *Nat. Rev. Genet.* **13**, 7–13 (2011).
- Zhao, Y. et al. *Nucleic Acids Res.* **47**, D163–D169 (2019).
- Jiang, P. et al. *Cancer Discov.* **10**, 664–673 (2020).
- Cristiano, S. et al. *Nature* **570**, 385–389 (2019).
- Lapin, M. et al. *J. Transl. Med.* **16**, 300 (2018).
- Yamamoto, Y. et al. *Cancer Sci.* **110**, 617–628 (2019).
- Cai, J. et al. *Gut* **68**, 2195–2205 (2019).
- Pang, Y. et al. *Hepatology* **68**, 1308–1318 (2018).
- Chayanupatkul, M. et al. *J. Hepatol.* **66**, 355–362 (2017).

THE PRECAR TEAM

Ying-Chao Wang⁵, Xiao-Tang Fan¹³, Guo-Qing Jiang¹⁴, Guo-Hong Deng¹⁵, Chun-Ying Wang¹⁶, Qiang Gao¹⁷, Feng Shen¹⁸, He-Ping Hu¹⁹, Qing-Zheng Zhang⁶, Yi-Nong Ye²⁰, Jing Zhang⁶, Yan-Hang Gao²¹, Jie Xia¹⁵, Hua-Dong Yan⁹, Min-Feng Liang²⁰, Yan-Long Yu²², Fu-Ming Sun⁶, Yu-Jing Gao¹⁶, Jian Sun¹⁰, Chun-Xiu Zhong¹⁰, Yin Wang⁶, Hui Wang¹⁹, Fei Kong²¹, Jin-Ming Chen²², Hao Wen²³, Dan Zheng²⁴, Yuan Yang^{24,25} and Chuan-Xin Wang^{7,8}

¹³Department of Hepatology, First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang 830000, China. ¹⁴Department of Hepatobiliary Surgery, Clinical Medical College, Yangzhou University, Yangzhou, Jiangsu 225001, China. ¹⁵Department of Infectious Diseases, Southwest Hospital, Third Military Medical University (Army Medical University), Chongqing 400038, China. ¹⁶Xuzhou Infectious Diseases Hospital, Xuzhou, Jiangsu 221004, China. ¹⁷Department of Liver Surgery and Transplantation, Liver Cancer Institute, Zhongshan Hospital, and Key Laboratory of Carcinogenesis and Cancer Invasion (Ministry of Education), Fudan University, Shanghai 200032, China. ¹⁸Department of Hepatic Surgery IV, Eastern Hepatobiliary Surgery Institute, Second Military Medical University, Shanghai 200438, China. ¹⁹Department of Hepatobiliary Medicine, Shanghai Eastern Hepatobiliary Surgery Hospital, Shanghai 210822, China. ²⁰The Department of Infectious Disease, the First People's Hospital of Foshan, Foshan, Guangdong 528000, China. ²¹The First Hospital of Jilin University, Changchun, Jilin 130021, China. ²²Chifeng Clinical Medical School of Inner Mongolia Medical University, Chifeng, Inner Mongolia 024000, China. ²³State Key Laboratory of Pathogenesis, Prevention and Treatment of High Incidence Diseases in Central Asia, First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang 830000, China. ²⁴Department of Gastroenterology, The Central Hospital of Wuhan, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei 430014, China. ²⁵Eastern Hepatobiliary Surgery Hospital, Second Military Medical University, Shanghai 200438, China.