950P: Ultra-Sensitive and Cost-effective Method for Early Stage Hepatocellular Carcinoma and Intrahepatic Cholangiocarcinoma Detection Using Plasma cfDNA Fragmentomic Profiles

Xiangyu Zhang[1], Zheng Wang[1], Xinyu Wang[1], Wanxiangfu Tang[2], Rui Liu[2], Hua Bao[2], Xin Chen[2], Shuyu Wu[2], Xue Wu[2], Yang Shao[2], Jia Fan[1,3,4], Jian Zhou[1,3,4]

[1]Department of Liver Surgery and Transplantation, Liver Cancer Institute, Zhongshan Hospital, Fudan University; Key Laboratory of Carcinogenesis and Cancer Invasion (Fudan University), Ministry of Education; Shanghai Key Laboratory of Organ Transplantation, Zhongshan Hospital, Fudan University, Shanghai, 200032, China; [2]Nanjing Geneseeq Technology Inc., Nanjing, Jiangsu, China; [3]Institute of Biomedical Sciences, Fudan University, Shanghai, 200032, China; [4]State Key Laboratory of Genetic Engineering, Fudan University, Shanghai, 200032, China.

Background

Primary liver cancer (PLC) is a leading cause of cancer mortality worldwide, with a total of over 840,000 new cases identified and 780,000 related deaths annually. Early detection of PLC, including hepatocellular carcinoma (HCC), intrahepatic cholangiocarcinoma (ICC), combined HCC-ICC (cHCC-ICC), is essential for patients' survival. Up-to-date, a fast, cost-effective, and accurate model is still needed for PLC early detection. This study aims to develop an accurate and cost-effective method for PLC early detection and differentiating ICC from HCC using plasma cell-free DNA (cfDNA) fragmentomic profiles.

Method

Whole-genome sequencings (WGS) were performed using plasma cfDNA samples from 192 PLC patients (159 HCC, 26 ICC, 7 cHCC-ICC) and 170 non-cancer controls (including 53 liver cirrhosis[LC] or hepatitis B virus[HBV]-positive) recruited in the training cohort. An ensembled stacked model for PLC detection was constructed using cfDNA Fragment Size Ratio (FSR) and Fragment Size Distribution (FSD) on the training cohort. The model performance was assessed in an independent test cohort (189 PLC patients [157 HCC, 26 ICC, 6 cHCC-ICC], 164 non-cancer controls [including 51 LC/HBV]).

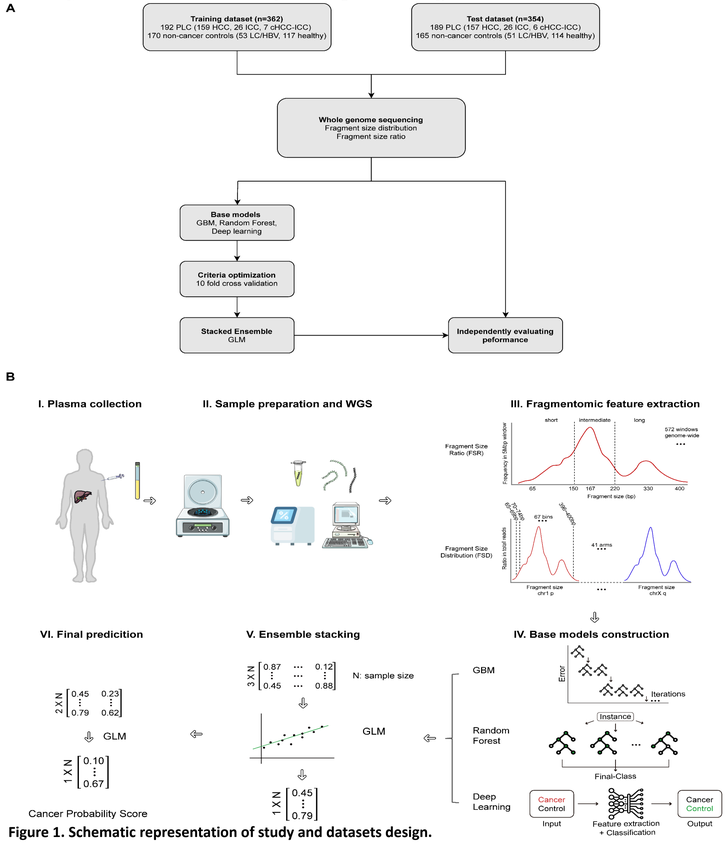

Figure 1. Schematic representation of study and datasets design.

Results

◆ **Our model showed excellent performance for cancer detection in the test cohort (Area Under the Curve [AUC]:0.995, 96.8% sensitivity at 98.8% specificity) (Table 1, Figure 2)**

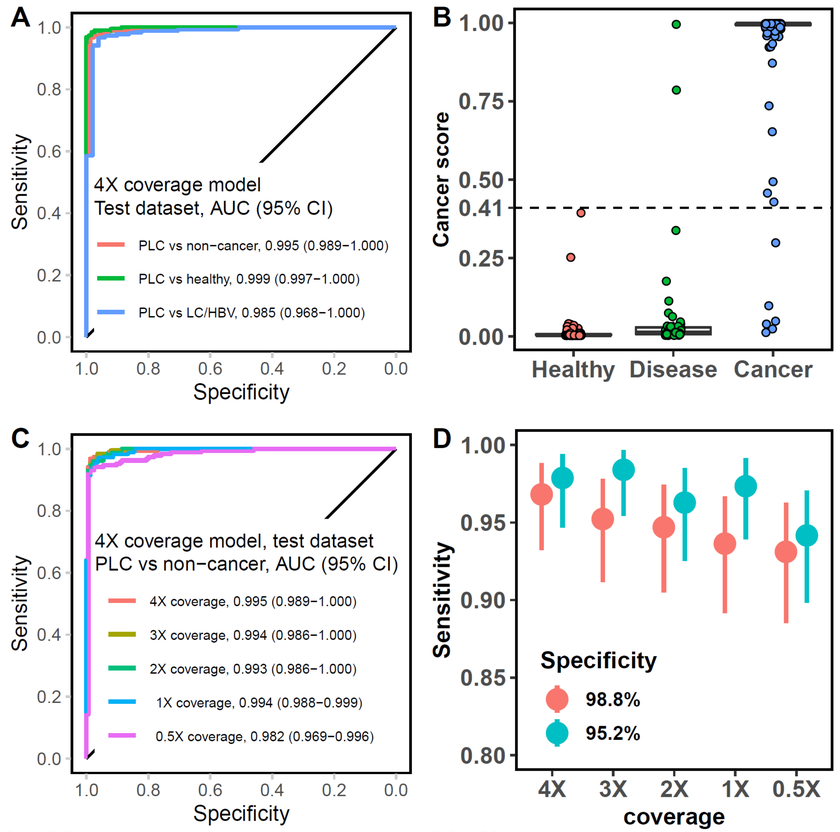

Figure 2. Evaluation of ensembled stacked model detecting early PLC. A) ROC curves evaluating the overall performance of the predictive model, which was constructed using unified 4x coverage WGS data, in distinguishing PLC patients from non-cancer controls (LC/HBV, healthy) in the test dataset. B) Boxplots illustrating cancer score distribution in the healthy, disease and cancer groups in the test dataset based on the 4x coverage model. The 98.8% specificity cutoff for cancer score was 0.41 as shown by the dotted line. C) ROC curves for distinguishing PLC from non-cancer controls of a limit of detection analysis, the 4X coverage model was evaluated using WGS data downsampled to 3X, 2X, 1X and 0.5X. D) Dot plot of 4X coverage model sensitivity in detecting PLC using 4X, 3X, 2X, 1X and 0.5X WGS data, at 98.8% (red) and 95.2% (green) specificities for non-cancer controls. The error bars represented 95% confidence interval.

◆ **Our model maintained consistent performances during downsampling process, even using 1X coverage data (AUC: 0.994, 93.7% sensitivity at 98.8% specificity) (Figure 2)**

◆ **A separate model showed great potential in distinguishing ICC from HCC (AUC: 0.776) (Figure 3)**

Table 1. Evaluating model performances using the test dataset.

Cancer vs Non-Cancer		Actual	
		Cancer	Non-Cancer
Predict	Cancer	183	2
	Non-Cancer	6	163
Sensitivity (95% CI)		96.8% (93.2-98.8%)	
Specificity(95% CI)		98.8% (95.7-99.9%)	
PPV (95% CI)		98.9% (96.1-99.9%)	
NPV (95% CI)		96.4% (92.4-98.7%)	
Accuracy (95% CI)		97.7% (95.6-99%)	
Cancer vs Healthy		**Actual**	
		Cancer	Healthy
Predict	Cancer	183	0
	Healthy	6	114
Sensitivity (95% CI)		96.8% (93.2-98.8%)	
Specificity(95% CI)		100% (96.8-100%)	
PPV (95% CI)		100% (98-100%)	
NPV (95% CI)		95% (89.4-98.1%)	
Accuracy (95% CI)		98% (95.7-99.3%)	
Cancer vs Disease		**Actual**	
		Cancer	Disease
Predict	Cancer	183	2
	Disease	6	49
Sensitivity (95% CI)		96.8% (93.2-98.8%)	
Specificity(95% CI)		96.1% (86.5-99.5%)	
PPV (95% CI)		98.9% (96.1-99.9%)	
NPV (95% CI)		89.1% (77.8-95.9%)	
Accuracy (95% CI)		96.7% (93.5-98.6%)	

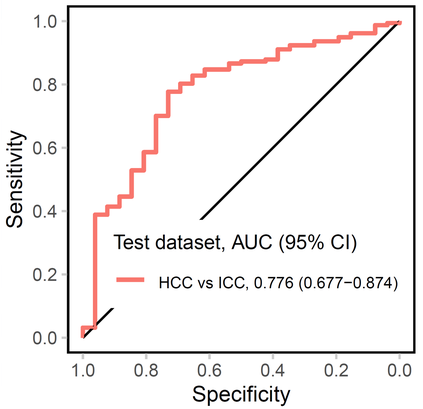

Figure 3. ROC curve of model distinguishing HCC from ICC. Ensembled stacked model for distinguishing HCC and ICC was trained using 159 HCC and 26 ICC from the training cohort and was evaluated using the 157 HCC and 26 ICC from the test cohort. The same FSR and FSD profiles and ensembled stacked machine learning approach was used for constructing the model

Conclusion

- ◆ We herein reported a predictive model using the comprehensive fragmentomic profiling of plasma cfDNA for PLC early detection.
- ◆ Our method, which is faster and more cost-effective by replying on only low coverage WGS data, has out-performed previously reported models.
- ◆ Our method has exhibited more significant potential in clinical practice for early detection of PLC and its different subtypes

Funding

◆ This work was supported by grants from the National Key R&D Program of China (2019YFC1315800 and 2019YFC1315802 to Jian Zhou), National Natural Science Foundation of China (81830102 and 81772578 to Jian Zhou) and Shanghai Municipal Key Clinical Specialty (to Jian Zhou).

Competing interests

◆ Wanxiangfu Tang, Rui Liu, Hua Bao, Xin Chen, Shuyu Wu, Xue Wu and Yang Shao are employees of Nanjing Geneseeq Technology Inc., Nanjing, Jiangsu, China. The remaining authors have nothing to declare.

CONTACT:

Xiangyu Zhang ,Email: tijmuzxy@163.com