Circulating exosome-derived bona fide long non-coding RNAs predicting the occurrence and metastasis of hepatocellular carcinoma

Yunjie Lu1 | Yunfei Duan1 | Qinghua Xu2 | Li Zhang2 | Weibo Chen1 | Zhen Qu1 | Baoqiang Wu1 | Wensong Liu1 | Longqing Shi1 | Di Wu1 | Yan Yang3 | Donglin Sun1 | Xuemin Chen1

1Department of Hepatobiliary Surgery, The First People’s Hospital of Changzhou, The Third Hospital Affiliated to Soochow University, Changzhou, China
2Department of General Surgery, Liyang People’s Hospital, Liyang, China
3Department of Nephrology, The First People’s Hospital of Changzhou, The Third Hospital Affiliated to Soochow University, Changzhou, China

Correspondence
Donglin Sun and Xuemin Chen, Department of Hepatobiliary Surgery, The First People’s Hospital of Changzhou, The Third Hospital Affiliated to Soochow University, Changzhou, Jiangsu, China
Emails: czsundl@yeah.net (DS); czcxm007@yeah.net (XC)

Funding information
Changzhou Medical Science Fund, Grant/Award Number: BRA2017118; science and technology project of health and family planning commission of Changzhou City, Grant/Award Number: ZD201502; Changzhou High-Level Medical Talents Training Project, Grant/Award Number: 2016CZBJ044

Abstract
Although the diagnosis and therapy approach developed, techniques for the early diagnosis of HCC remain insufficient which results in poor prognosis of patients. The traditional biomarker AFP, however, has been proved with low specificity. Circulating exosomal ncRNAs revealed different profiles reflecting the characteristics of tumour. In this study, we mainly focused on circulating exosomal ncRNAs which might be the fingerprint for HCC, especially for the diagnosis or metastasis prediction. A high throughput IncRNA microarray in exosomes extracted from cell-free plasma was applied. The risk score analysis was employed to screen the potential exosome-derived IncRNAs in two independent sets based on different clinical parameters in 200 paired HCC patients. After a multi-stage validation, we finally revealed three IncRNAs, ENSG00000248932.1, ENST00000440688.1 and ENST00000457302.2, increased in HCC comparing with the both chronic hepatitis (CH) patients and cancer-free controls. ROC curve revealed a higher sensitivity and specificity in predicting the occurrence of HCC from cancer-free controls and CH patients with the area under curve (AUC) of 0.905 and 0.879 by combining AFP. The three IncRNA panel combined with AFP also indicted a fingerprint function in predicting the metastasis of HCC with the AUC of 0.870. In conclusion, ENSG00000248932.1, ENST00000440688.1 and ENST00000457302.2 might be the potential biomarker for the tumorigenesis prediction from CH patients or healthy controls and may also be applied for dynamic monitoring the metastasis of HCC.

KEYWORDS
exosome, fingerprint, IncRNA, plasma, risk score analysis
1 | INTRODUCTION

Hepatocellular carcinoma, also known as HCC, with the highly infection of HBV or HCV background in Chinese han population, is now listed as the most frequent liver cancer, with an increasing incidence over the last decades. However, the diagnosis technologies for HCC in different stages has been improved. For example, imaging techniques such as US, CT and MRI, with a high resolution and was also recommended by guidelines. However, for the nodules smaller than 2 cm in diameter in patients with liver cirrhosis, it was really difficult to distinguish. During the past decades, various biomarkers such as Alpha Fetal Protein (AFP), circulating miRNAs and IncRNAs have been proved as biomarker for HCC prediction. AFP was the traditional marker for the diagnosis or monitoring the recurrence for HCC; however, multiple studies have demonstrated the low specificity of AFP detection despite its high sensitivity. A novel, minimally invasive indicator with high sensitivity and specificity is necessary for the HCC especially for the early diagnosis or dynamic monitoring.

Exosomes were defined as small membranous vesicles with a size of 30-120 nm. Circulating exosomes could be released into the extracellular environment through the fusion of multivesicular bodies approach with the membrane of certain body fluid such a serum or plasma. It has been proved that ncRNA extracted from exosome could present a stable form to avoid the degradation induced by RNase, indicating they might be mediators for communications in different cells. Exosome-derived non-coding RNAs (ncRNAs) have been proved with different expression profiles which could indicate the characteristics of a certain tumour, or their function in tumour progression and metastasis. Various researches have proved that exosome-derived ncRNAs could act as markers for the diagnosis and prognosis prediction in various kinds of human cancers. However, most of the research mainly focused on the miRNAs. Although various proteins biomarkers such AFP, CEA, miRNA or IncRNAs have been isolated from exosomes as potential fingerprint for different approach, little is known about plasma exosome-derived IncRNAs in HCC, little systematic study was conducted by using the high throughput detection of IncRNA. Also many IncRNAs in exosomes have not been adequately investigated.

2 | MATERIALS AND METHODS

2.1 | Patients

Patients diagnosed with HCC and admitted to The First People's Hospital of Changzhou between October 2010 and September 2017 were enrolled in this study. Plasma samples were collected from these patients before surgical operation. Patients whom pathological diagnosis as either intrahepatic metastasis or portal vein tumour thrombus were regarded as metastasis HCC. This study was performed in accordance with the ethical guidelines of the Helsinki Declaration of 1975, as revised in 2008. All patients included in the study provided informed consent, and the study protocol was approved by the Institutional Review Board of The First People's Hospital of Changzhou. Experiments were undertaken with the understanding and written consent of each subject. Peripheral blood samples of patients were collected before operation. Blood samples were collected in a separate vacuum cube, followed by centrifugation at 1000g for 10 minutes.

2.2 | Isolation and identification of exosome-derived RNAs from cell-free plasma samples

The exosomes were extracted from plasma samples using the ExoQuick Exosome Precipitation Solution (System Biosciences). Briefly, 500 μL plasma was mixed with solution and then incubated for 30 minutes at 4°C. To validate the isolation of the exosomes from plasma, we performed the immunoblotting by using the exosome biomarkers such as CD9, CD63 and TSG101 for investigation. We also employed the calnexin as a negative control based on the little endogenous expression in exosome. The extracted exosomes were further detected on NanoSight LM10 instrument for the basic volume and concentration.

2.3 | RNA extracting, microarray detection and data analysis workflow

The mirNeasy serum/plasma micro kit (Qiagen) was used for total RNAs isolation. The synthetic Caenorhabditis elegans miRNA (cel-miR-39; Applied Biosystems) was added to each sample as external reference.

A multiphase case-control study was designed to identify the exosome-derived IncRNA profile as a signature for HCC. In the screening stage, RNA extracted from three HCC patients without metastasis (HCC-N), three HCC patients with metastasis (HCC-M), three patients diagnosed with chronic hepatitis (HCC) and three paired controls samples were also subjected for Human LncRNA Array v3.0 (Arraystar; Agilent) to screen the IncRNAs in different groups. Based on this, the RT-qPCR was firstly used in the training set to further filter signals of the screened IncRNAs. The detailed steps for candidate screening and data filtering parameters were presented in Figure S1. The detailed primer sequence has been listed in Table S6.

2.4 | Risk score analysis

The screening phase was divided into training set and validation set.

Training set

The training set contained 20 samples with random selection. The expression levels of these candidates were analysed in these samples, and the algorithms comparative 2−ΔΔCt method were applied for analysing.

Validation set

The validation set enrolled 180 samples in each group. The case-control study was designed to measure relative expression levels of the selected potential biomarkers candidates.
Risk score analysis was a traditional analysis to validate a certain biomarker. Here, data in the training set were performed to evaluate the associations between the concentrations of the plasma lncRNA expression levels. The upper 95% reference interval of each lncRNA value in controls was set as the threshold to code the expression level of the corresponding lncRNA for each sample as 0 and 1 in the training set. A risk score function (RSF) was defined according to a linear combination of the expression level for each lncRNA. For example, the RSF for sample i using information from four lncRNAs was as follows: 

$$rsfi = \sum_{j=1}^{4} W_j \cdot sij.$$ 

In the above equation, sij is the risk score for lncRNA j on sample i, and Wj is the weight of the risk score of lncRNA j. We conducted the ROC analysis by using the total RSF value according to the case-control group in the training set. We chose the value as the cut-off value because the value of sensitivity + specificity was maximal.

### 2.5 Statistical analysis

If no special circumstances were declared, data were presented as mean ± SD. Chi-square tests and Student’s t test analysis of variance were used to evaluate statistical differences in demographic and clinical characteristics, respectively. Statistical analysis was performed using STATA 10.0 and presented with GraphPad Prism 5.0 software. Results were considered statistically significant at $P < .05$.

### 3 RESULTS

#### 3.1 Clinical parameters analysis of patients enrolled in this study

A total of 600 patients including 200 HCC patients, 200 CH patients and 200 healthy controls were enrolled in this study. The detailed clinical information was listed in Table S1. All the HCC patients and CH patients was confirmed with HBV or HCV infection background. The tumours were smaller than 5 cm in 116 patients, 111 patients had multiple tumours, half patients were diagnosed with metastasis while only 14 patients with tumour capsular incomplete. The median AFP level was 32.7 ng/mL (Table S1). Besides, among the HCC patients, we divided into two groups according to the metastasis (HCC-N and HCC-M), each group enrolled 100 HCC patients. The clinical parameters were calculated and analysed including the tumour size, tumour number, tumour capsular, AFP value and TNM stage. As presented in Table S2, none of these was proved as significant.

#### 3.2 Plasma exosomes identification and expression detection

Firstly, samples extracted from human plasma were detected by using exosome specific biomarker. As presented in Figure 1A, CD9, CD63 and TSG101 were used as positive controls for exosomes, and for calnexin, which is an integral protein of the endoplasmic reticulum and is not expressed in the exosome.

![Figure 1A](image-url)

[Figure 1A](image-url) The expression landscape of circulating exosome-derived lncRNAs in HCC patients. A, Expression of CD9, CD63, TSG101 and calnexin by Western blot. The expression of CD9, CD63, TSG101 and calnexin in isolated exosomal pellets from the sera of patients were compared with those of the supernatant. The expression of CD9, CD63, and TSG101 was present in the isolated exosomes, whereas calnexin, a negative marker of exosomes, was only expressed in the supernatant. B, Cluster analysis of the different expression of the lncRNAs in different groups. NC: normal control; CH: chronic hepatitis; HCC-N: HCC without metastasis; HCC-M: HCC with metastasis.
(c) detection rate >75%. A total of 81 IncRNA transcripts were specifically increased in CH group comparing with NC group, 179 IncRNAs were collected in HCC-N group by comparing with CH group while 212 IncRNAs were up-regulated in HCC-M group comparing with the HCC-N group. In order to screen the biomarker predication the tumorigenesis and metastasis of HCC, the Venny analysis was applied and finally yielded 6 IncRNA candidates as listed in Figure 2A.

### 3.3 Identification of IncRNA genes from the training set and validation set

To further analyse the aberrant expression of the six candidates, as summarized in the workflow, all analyses were performed in the training data set first and then validated in another set. The 20 samples in the training set were randomly selected from the 200 patients’ samples. We further examined these differentially expressed IncRNAs by RT-qPCR in all the samples we enrolled. As presented in Figure 2B, the increased level of ENSG00000248932.1, ENSTO0000440688.1 and ENSTO0000457302.2 was confirmed in the larger sample size while the XLOC_001120, ENSG00000243766.2, AC058791.2 and TCONS_00003661 was revealed as no difference. Another 180 clinical samples were applied as the validation set. The three-panel IncRNAs were also detected in these set, as presented in Figure 3, ENSG00000248932.1, ENSTO0000440688.1 and ENSTO0000457302.2 was confirmed with continuous increasing level in NC, CH and HCC groups. In addition, by comparing the HCC-M with non-metastasis group, we also obtained a up-regulated level of all the three IncRNAs.

### 3.4 The association of three-ncRNA signature and patient’s occurrence/metastasis in the training set and validation set

With the risk score formula, we measured the three-ncRNA expression signature risk score in the training set. The patients were then ranked according to their risk scores. By using the median risk score
as cut-off in the training set, the patients were divided into low-risk and high-risk groups. In the following analysis, we compared the function of the three lncRNA in predicting HCC from healthy controls, HCC from CH patients and HCC-M from non-metastasis patients. AFP was also used as an independent factor. The combination of the three lncRNAs was also calculated as an independent factor as well was the combination of the three lncRNAs and AFP.

Firstly, to explore the ability of three lncRNA in HCC patients from healthy controls, patients in the high-risk group were labelled as HCC group while low-risk was regarded as healthy controls. Based on this cut-off, which was defined as the value of sensitivity plus specificity considered to be maximal, PPV and NPV were 90% and 95% in the training set, respectively. Similarly, when the same cut-off value was applied to calculate the risk score of samples in the larger validation sets, the PPV and NPV were 84% and 80%, respectively (Table S3).

We also used the ROC curves analysis to investigate the diagnostic sensitivity and specificity of the three-lncRNAs signature for HCC. Each single lncRNA alone and the three merged factors were analysed, respectively. As presented in Figure 4A, the AUCs of ENSG00000248932.1, ENST00000440688.1, ENST00000457302.2, AFP, combination of lncRNAs and combination of lncRNAs and AFP were 0.530, 0.632, 0.883, 0.400, 0.818 and 0.833, respectively, in the training set. The values of AUC in validation set were 0.794, 0.571, 0.538, 0.510, 0.838 and 0.905, respectively, as presented in Figures S2 and S3.

Secondly, patients in the high-risk group were labelled as HCC group while low-risk was regarded as CH patients. Based on this cut-off, which was defined as the value of sensitivity plus specificity considered to be maximal, PPV and NPV were 90% and 95% in the training set, respectively. Similarly, when the same cut-off value was applied to calculate the risk score of samples in the larger validation sets, the PPV and NPV were 84% and 80%, respectively (Table S4).

We also used the ROC curves analysis to investigate the diagnostic sensitivity and specificity of the three-lncRNAs signature for HCC. Each single lncRNA alone and the three merged factors were analysed, respectively. As presented in Figure 4B, the AUCs of ENSG00000248932.1, ENST00000440688.1, ENST00000457302.2, AFP, combination of lncRNAs and combination of lncRNAs and AFP were 0.776, 0.612, 0.720, 0.470, 0.960 and 0.970, respectively, in training set. The values of AUC in validation set were 0.699, 0.632, 0.565, 0.468, 0.534 and 0.870, respectively, as presented in Figures S2 and S3.

Thirdly, to explore the ability of three lncRNA in HCC-M from non-metastasis HCC patients, the patients in the high-risk group were labelled as HCC-M while low-risk was regarded as HCC-N. PPV and NPV were 90% and 95% in training set while 90% and 96% in validation set (Table S5).

The AUCs of ENSG00000248932.1, ENST00000440688.1, ENST00000457302.2, AFP, combination of lncRNAs and combination of lncRNAs and AFP were 0.672, 0.710, 0.505, 0.408, 0.850 and 0.910, respectively, in training set, and in the validation set, the AUC of which was 0.729, 0.762, 0.588, 0.546, 0.721 and 0.893, respectively, as presented in Figures S2 and S3.
3.5 | Double-blind test

Another 100 independent plasma samples were detected in a double-blind fashion to validate the predictive ability of the three lncRNA. We used the same risk score formula to analyse the expression of the three exosome-derived lncRNAs in those plasma samples and classifying them into a high-risk group and a low-risk group. Based on the pathologic diagnosis, the accuracy rate of the three exosome-derived lncRNA profile as HCC signature from healthy controls, CH patients was 85.2%, 88.3%, while predicting the metastasis was 81.9%.

3.6 | Stability expression of ENSG00000248932.1, ENST00000440688.1, ENST00000457302.2 in human plasma

The expression of the three lncRNAs was detected in RNA sample extracted from four healthy controls and was incubated at room temperature for 12 hours, 24 hours, subjecting it to up to five cycles of freezing and thawing or under storage of −80°C for about 7 days. After that, the exosome was further extracted. All the process had minimal effects on the concentrations of the four lncRNAs, demonstrating that these lncRNAs were sufficiently stable in the exosome of human plasma (Figure 5).

4 | DISCUSSION

The exploration of biomarker for HCC has been undertaking by multiple researchers over the past decades. The annotated biomarkers extracted from tumour tissues or cell-free plasma has been proved might indicate the occurrence or recurrence of HCC. HCC is one of the most common human malignant tumours worldwide and was proved with high mortality rates.

Circulating ncRNAs have been well documented based on the crucial role during the pathogenesis and development of human cancer. Several ncRNAs have been developed as potential biomarkers for early screening. As we known exosome-derived ncRNAs could remain a stable condition in plasma and exhibit different expression profiles representing the properties of cancer cells, researches hypothesized these exosome-derived ncRNAs might serve as sensitive and non-invasive biomarkers for both diagnostic and prognostic...
purposes. In recent years, increasing studies have focused on this field, resulting in a subclass of extracellular vesicles involved in intercellular communication released by most cell types and various body fluid.\textsuperscript{23} Cells can trigger cancer-related disorders after the recognition and uptake of circulating exosome-derived lncRNAs, providing indications for early tumour biopsy, diagnosis and treatment.\textsuperscript{24} Accumulated studies have implicated that exosomes play critical roles in the development and progression of malignant tumours. It has been confirmed tumour-derived exosomes tumour antigens and promote tumour development.\textsuperscript{25} For example, Linc00974 has been identified as biomarker for HCC diagnosis which was involved in TGF-beta-associated pathway.\textsuperscript{17} Besides, long non-coding RNA POU3F3 in plasma has also proved as a novel biomarker for diagnosis of oesophageal squamous cell carcinoma.\textsuperscript{26} Here in this study, exosome from plasma was first isolated and identified. Based on this, the exosome-derived ncRNAs were extracted and were applied for the lncRNA microarray. The certain parameters were used for screening to guarantee the further validation. The risk score analysis was applied to test the diagnosis predicting ability for HCC from NC and CH patients as well as the metastasis of HCC. Finally, we obtained the ENSG00000248932.1, ENST00000440688.1, ENST00000457302.2 as candidate fingerprint. The ROC analysis with AFP as an independent biomarker also revealed that the three exosome-derived lncRNAs panel indicated a higher sensitivity and specificity than AFP.

In conclusion, we identified three exosome-derived lncRNAs, ENSG00000248932.1, ENST00000440688.1, ENST00000457302.2, as the potential fingerprints for the tumorigenesis prediction. Thus, we propose that this panel lncRNAs might be utilized to develop early diagnosis and invasive screening tools for HCC. More in-depth studies are required to confirm the potential mechanism of this lncRNA in the development of HCC.

CONFLICT OF INTEREST
The authors declared that they have no financial competing interest.

AUTHOR CONTRIBUTIONS
YJL, DLS and YFD designed the study and drafted the manuscript; QHX, WBC, ZQ and BQW participated in data organization; YJL, LZ, WSL, LQS and DW collected the patients’ information; YY and XMC conducted the statistical analysis. All authors read and approved the final manuscript.

ORCID
Donglin Sun \(\text{https://orcid.org/0000-0001-6295-9709}\)

REFERENCES
1. Bayo J, Fiore EJ, Dominguez LM, et al. A comprehensive study of epigenetic alterations in hepatocellular carcinoma identifies potential therapeutic targets. J Hepatol. 2019;71:78-90.
2. Arzumanyan A, Reis HM, Feitelson MA. Pathogenic mechanisms in HBV- and HCV-associated hepatocellular carcinoma. Nat Rev Cancer. 2013;13:123-135.
3. Cast A, Valanejad L, Wright M, et al. C/EBPalpha-dependent pre-neoplastic tumor foci are the origin of hepatocellular carcinoma and aggressive pediatric liver cancer. Hepatology. 2018;67:1857-1871.
4. Chaisaingmongkol J, Budhu A, Dang H, et al. Common molecular subtypes among Asian hepatocellular carcinoma and cholangiocarcinoma. Cancer Cell. 2017;32:57-70.e3.
5. Farazi PA, DePinho RA. Hepatocellular carcinoma pathogenesis: from genes to environment. Nat Rev Cancer. 2006;6:674-687.
6. Prieto J, Melero I, Sangro B. Immunological landscape and immunotherapy of hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol. 2015;12:681-700.
7. Rajaguru DJ, Borgert AJ, Smith AL, et al. Radiofrequency ablation versus stereotactic body radiotherapy for localized hepatocellular
carcinoma in nonsurgically managed patients: analysis of the national cancer database. J Clin Oncol. 2018;36:600-608.
8. Salem ESB, Fan GC. Pathological effects of exosomes in mediating diabetic cardiomyopathy. Adv Exp Med Biol. 2017;998:113-138.
9. Lee YR, Kim G, Tak WY, et al. Circulating exosomal noncoding RNAs as prognostic biomarkers in human hepatocellular carcinoma. Int J Cancer. 2019;144:1444-1452.
10. Fu Q, Zhang Q, Lou Y, et al. Primary tumor-derived exosomes facilitate metastasis by regulating adhesion of circulating tumor cells via SMAD3 in liver cancer. Oncogene. 2018;37:6105-6118.
11. Wang S, Xu M, Li X, et al. Exosomes released by hepatocarcinoma cells endow adipocytes with tumor-promoting properties. J Hematol Oncol. 2018;11:82.
12. Liu H, Chen W, Zhi X, et al. Tumor-derived exosomes promote tumor self-seeding in hepatocellular carcinoma by transferring miRNA-25-5p to enhance cell motility. Oncogene. 2018;37:4964-4978.
13. Wang X, Shen H, Zhangyuan G, et al. 14-3-3zeta delivered by hepatocellular carcinoma-derived exosomes impaired anti-tumor function of tumor-infiltrating T lymphocytes. Cell Death Dis. 2018;9:159.
14. Lou G, Chen Z, Zheng M, et al. Mesenchymal stem cell-derived exosomes as a new therapeutic strategy for liver diseases. Exp Mol Med. 2017;49:e346.
15. Yao H, Liu N, Lin MC, et al. Positive feedback loop between cancer stem cells and angiogenesis in hepatocellular carcinoma. Cancer Lett. 2016;379:213-219.
16. Lasser C, Eldh M, Lotvall J. Isolation and characterization of RNA-containing exosomes. J Vis Exp. 2012;59:e3037.
17. Tang J, Zhuo H, Zhang X, et al. A novel biomarker Linc00974 interacting with KRT19 promotes proliferation and metastasis in hepatocellular carcinoma. Cell Death Dis. 2014;5:e1549.
18. Zucman-Rossi J, Villanueva A, Nault JC, et al. Genetic landscape and biomarkers of hepatocellular carcinoma. Gastroenterology. 2015;149:1226-1239.e4.
19. Wang Y, Yang L, Chen T, et al. A novel IncRNA MCM3AP-AS1 promotes the growth of hepatocellular carcinoma by targeting miR-194-5p/FOXA1 axis. Mol Cancer. 2019;18:28.
20. Dong Q, Du Y, Li H, et al. EGFR and c-MET cooperate to enhance resistance to PARP inhibitors in hepatocellular carcinoma. Cancer Res. 2019;79:819-829.
21. Peng L, Jiang B, Yuan X, et al. Super-enhancer-associated long non-coding RNA HCCL5 is activated by ZEB1 and promotes the malignancy of hepatocellular carcinoma. Cancer Res. 2019;79:572-584.
22. Fang T, Lv H, Lv G, et al. Tumor-derived exosomal miR-1247-3p induces cancer-associated fibroblast activation to foster lung metastasis of liver cancer. Nat Commun. 2018;9:191.
23. Arbelaitz A, Azkargorta M, Krawczyk M, et al. Serum extracellular vesicles contain protein biomarkers for primary sclerosing cholangitis and cholangiocarcinoma. Hepatology. 2017;66:1125-1143.
24. Qi H, Liu C, Long L, et al. Blood exosomes endowed with magnetic and targeting properties for cancer therapy. ACS Nano. 2016;10:3323-3333.
25. Singer S, Zhao R, Barsotti AM, et al. Nuclear pore component Nup98 is a potential tumor suppressor and regulates posttranscriptional expression of select p53 target genes. Mol Cell. 2012;48:799-810.
26. Tong YS, Wang WX, Zhou XL, et al. Identification of the long non-coding RNA POU3F3 in plasma as a novel biomarker for diagnosis of esophageal squamous cell carcinoma. Mol Cancer. 2015;14:3.

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Lu Y, Duan Y, Xu Q, et al. Circulating exosome-derived bona fide long non-coding RNAs predicting the occurrence and metastasis of hepatocellular carcinoma. J Cell Mol Med. 2020;24:1311-1318. https://doi.org/10.1111/jcmm.14783