The combination of circulating long noncoding RNAs AK001058, INHBA-AS1, MIR4435-2HG, and CEBPA-AS1 fragments in plasma serve as diagnostic markers for gastric cancer

Dong Ke[1,2,3], Hanwei Li[2], Yi Zhang[2,3], Yinghong An[4], Hanjiang Fu[2], Xuedong Fang[1,3], Xiaofei Zheng[2]

[1]General Surgery, The Second Hospital of Jilin University, Changchun, 130041, China

[2]Beijing Key Laboratory for Radiobiology, Beijing institute of Radiation Medicine, Beijing, 100850, China

[3]Gastrointestinal Colorectal and Anal Surgery, The China-Japan Union Hospital of Jilin University, Changchun, 130033, China

[4]Clinical Laboratory Center, Chinese PLA Air Force General Hospital, Beijing, 100142, China

Correspondence to: Hanjiang Fu, **email:** Fuhj75@126.com
                      Xuedong Fang, **email:** fangxuedong@medmail.com.cn
                      Xiaofei Zheng, **email:** xfzheng100@126.com

ABSTRACT

Background: Suitable diagnostic markers for cancers are urgently required in clinical practice. Long non-coding RNAs, which have been reported in many cancer types, are a potential new class of biomarkers for tumor diagnosis.

Results: Five lncRNAs, including AK001058, INHBA-AS1, MIR4435-2HG, UCA1 and CEBPA-AS1 were validated to be increased in gastric cancer tissues. Furthermore, we found that plasma level of these five lncRNAs were significantly higher in gastric cancer patients compared with normal controls. By receiver operating characteristic analysis, we found that the combination of plasma lncRNAs with the area under the curve up to 0.921, including AK001058, INHBA-AS1, MIR4435-2HG, and CEBPA-AS1, is a better indicator of gastric cancer than their individual levels or other lncRNA combinations. Simultaneously, we found that the expression levels of a series of MIR4435-2HG fragments are different in gastric cancer plasma samples, but most of them higher than that in healthy control plasma samples.

Materials and Methods: LncRNA gene expression profiles were analyzed in two pairs of human gastric cancer and adjacent non-tumor tissues by microarray analysis. Nine gastric cancer-associated lncRNAs were selected and assessed by quantitative real-time polymerase chain reaction in gastric tissues, and 5 of them were further analyzed in gastric cancer patients' plasma.

Conclusions: Our results demonstrate that certain lncRNAs, such as AK001058, INHBA-AS1, MIR4435-2HG, and CEBPA-AS1, are enriched in human gastric cancer tissues and significantly elevated in the plasma of patients with gastric cancer. These findings indicate that the combination of these four lncRNAs might be used as diagnostic or prognostic markers for gastric cancer patients.

INTRODUCTION

Gastric cancer (GC) is one of the most common malignant tumors [1] and the second most frequent cancer associated with mortality, partially owing to the fact that it is often diagnosed at an advanced stage, and not at the early stage [2, 3]. It is vitally important from a clinical

perspective to seek effective tools for the early detection of GC.

Long non-coding RNAs (lncRNAs) are a newly discovered class of ncRNAs longer than 200 nucleotides in length [4]. Changes in lncRNA expression levels have been increasingly reported in a variety of cancer types, suggesting a correlation between lncRNAs and

carcinogenesis [5, 6]. Therefore, changes in lncRNA expression levels can serve as a novel target for cancer diagnosis and therapy. For example, lncRNAs can serve as potent biomarkers for the diagnosis and prognosis of clear cell renal cell carcinoma [7]. Overexpression of lncRNA ZFAS1 is associated with intrahepatic and extrahepatic metastasis and poor prognosis of HCC [8]. Moreover, lncRNA HIF1A-AS1, regulates proliferation and migration of oesophageal adenocarcinoma cells [9]. Level of H19 was higher in GC patient plasma than healthy controls with the area under curve(AUC) up to 0.838 [10]. However, Some lncRNAs, highly expressed in GC tissues, served as markers with low AUCs [11, 12]. Given that single lncRNA management, the AUC is too low to discriminate GC patients from healthy controls. Therefore, it is important to discover lncRNAs with higher sensitivity and specificity that can be used as a biomarker for GC diagnosis.

In this study, nine lncRNAs, dysregulated in GC tissues identified by lncRNA microarray, were selected to evaluate their expression in gastric tissues and plasma samples. The purpose of our study was to determine whether these lncRNAs could discriminate GC patients from healthy controls. In addition, we analyzed the expression characteristics of diverse lncRNA fragments in plasma samples, and the potential relationship between circulating lncRNA levels and the clinicopathological features of GC.

RESULTS

Expression profile of lncRNAs in gastric cancer

LncRNA gene expression profiles were analyzed in two pairs of human GC and adjacent non-tumor (NT) tissues by microarray analysis. Fold-change (Tumor vs. Normal) and P value were calculated from the normalized expression (Fold-change ≥ 2 or ≤ 0.5, $p < 0.05$). The microarray data has been deposited in NCBI Gene Expression Omnibus (GEO) and the GEO accession number is GSE93512. In total, 154 lncRNAs were identified to be consistently increased (Supplementary Figure 1A) in all two GC groups, and 238 lncRNAs were consistently decreased (Supplementary Figure 1B). Among these, 9 lncRNAs, showing significant difference in both tissue microarrays, were selected for further validation (Supplementary Table 1). Of these 9 lncRNAs, INHBA-AS1, MIR4435-2HG, UCA1, AK001058, LOC100133091, and MGC12916 were increased, where as CEBPA-AS1, FLJ37453, and LINC01184 were decreased in GC tissues.

Five lncRNAs were increased in GC tissues

Based on the gastric tissue microarray results, we validated the expression of the 9 lncRNAs in 49 GC tissues and adjacent NT tissues using qRT-PCR. Selection

of an appropriate reference gene is crucial to the analysis. RNA expression was normalized to that of β-actin [13, 14] or 18S rRNA as described previously [15, 16]. In this study, 18S rRNA was selected as the reference gene, because the expression level of 18S rRNA was not significantly different between GC tissues and adjacent NT tissues. We first examined 18 paired gastric tissues, but of the 9 selected lncRNAs, lncRNA FLJ37453, LINC01184, LOC100133091, and MGC12916 did not show marked changes (results not shown). Next, we examined the other five lncRNAs in the remaining 31 paired gastric tissues. LncRNAs INHBA-AS1, MIR4435-2HG, CEBPA-AS1, UCA1, and Ak001058 were increased in 37 (75.51%), 41 (83.67%), 39 (75.59%), 39 (75.59%), and 47 (95.92%) of the 49 GC tissues, respectively (Figure 1A–1E). The relationship between lncRNA levels in tissues and the clinicopathological features of GC patients was also analyzed (Table 1). The expression levels of INHBA-AS1, MIR4435-2HG, CEBPA-AS1, and AK00108 were associated with tumor grade (Supplementary Figure 2A–2D); AK001058 had a higher expression level in GC tissues with lymph node metastasis compared to that with no lymph node metastasis (Supplementary Figure 2E), and the expression level of UCA1 was higher in GC I stage than that in GC II-IV stage (Supplementary Figure 2F). The AUCs for INHBA-AS1, MIR4435-2HG, CEBPA-AS1, UCA1, and AK001058 were 0.740, 0.770, 0.741, 0.722, and 0.957, respectively (Supplementary Figure 3A). The AUC value of the combination of 5-lncRNA was up to 0.976 (95%CI: 0.000–1.000) (Supplementary Figure 3B), when the AUC value of a single lncRNA was lower than that of the 5-lncRNA signature.

Correlation of antisene lncRNAs expression and their corresponding mRNAs expression in gastric cancer tissues

Most protein coding genes (PCGs) have their associated antisense RNA, which can interact with nearby associated PCGs. LncRNAs are reportedly able to regulate all steps of the gene expression process [17]. Numerous studies have focused on the analysis of the expression patterns of lncRNAs and their possible crosstalk with adjacent protein-coding genes. The antisense lncRNA Khps1 activates SPHK1 transcription by targeting chromatin modifying enzymes to the SPHK1 promoter and changing chromatin structures [18]. RBM15-AS1, transcribed in the opposite direction within exon 1 of RBM15 was increased in megakaryocyte and activated megakaryocyte differentiation and may play a regulatory role in leukemogenesis by enhancing RBM15 protein translation[19]. INHBA-AS1 and CEBPA-AS1 are the antisense RNAs of INHBA and CEBPA, respectively. CEBPA-AS1 and CEBPA were both increased in 23 (95.8%) and decreased in 1 (4.17%) GC tissues (Supplementary Figure 4A). INHBA-AS1 and INHBA were

Table 1: Correlation between lncRNA-INHBA-AS1, MIR4435-2HG, CEBPA-AS1, UCA1, and AK001058 panel expression levels in gastric tissues and clinical parameters

Clinical parameters	Number of cases	*p*-value				
		INHBA-AS1	MIR4435-2HG	CEBPA-AS1	UCA1	AK001058
Age(years)						
< 60	19	0.2067	0.0251 (*)	0.8578	0.0571	0.2048
> 60	30					
Gender						
female	18	0.2906	0.0589	0.2672	0.8692	0.5503
male	31					
Tumor grade						
Low	28	0.0114(*)	0.0416(*)	0.0313(*)	0.4828	0.0330(*)
middle or high	21					
LN metastasis						
No	17	0.9298	0.4592	0.8496	0.2886	0.0350(*)
Yes	32					
Depth of invasion						
T1-2	11	0.7742	0.6702	0.2251	0.5407	0.7243
T3-4	38					
TNM stage						
I	19	0.7255	0.7011	0.7996	0.0185(*)	0.2222
II-IV	30					

Abbreviations: lncRNA, long noncoding RNA; LN, lymph node. TNM, tumor node metastasis; *Statistically Significant ($p < 0.05$).

both increased in 19 (78.2%) (Supplementary Figure 4B) among 24 paired GC tissues. According these results, we found that the changing trend of CEBPA, INHBA, and their antisense RNA basically identical. It would be valuable to study the functional relationship between INHBA-AS1, CEBPA-AS1 and their related PCGs.
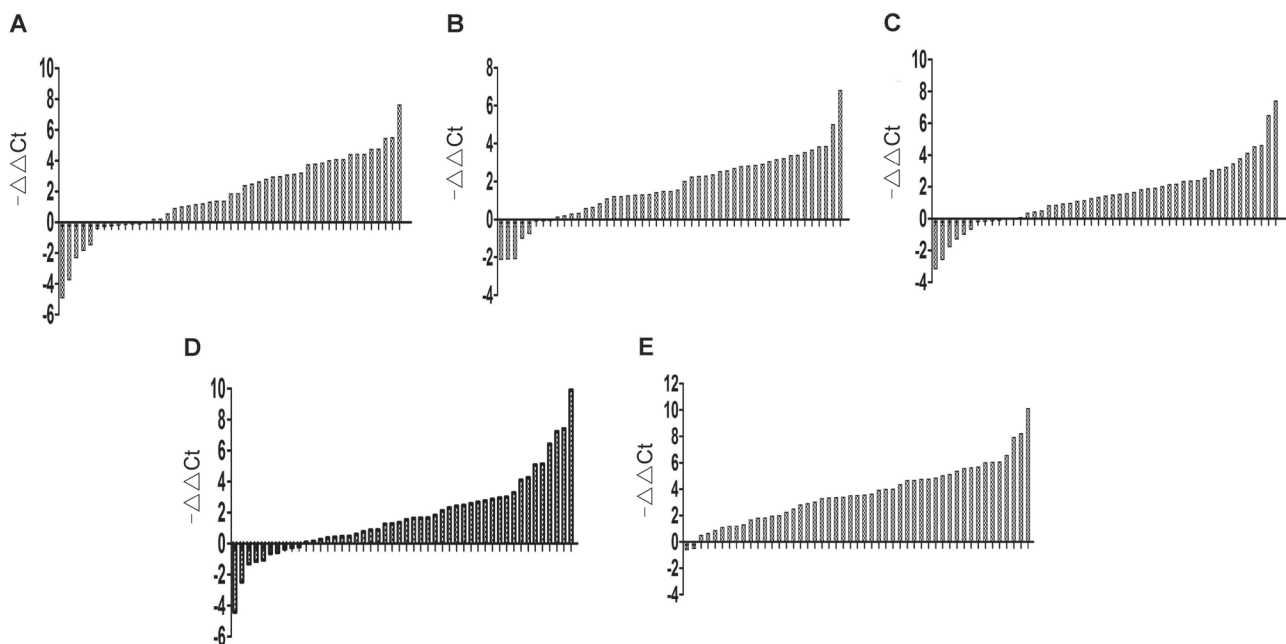

Figure 1: Gene expression levels in gastric tissues. Expression level of lncRNA INHBA-AS1 (**A**), MIR4435-2HG (**B**), CEBPA-AS1 (**C**), UCA1 (**D**), and AK001058 (**E**) in 49 paired gastric tissues. $-\Delta\Delta Ct = (Ct_{\text{target gene in normal}} - Ct_{\text{18S rRNA in normal}}) - (Ct_{\text{target gene in cancer}} - Ct_{\text{18S rRNA in cancer}})$.

Detection of plasma expression of the 5 lncRNAs fragments in the training set

We validated the 5 lncRNAs that were increased in the GC tissues in 51 GC patients and 53 healthy controls, the plasma samples were collected from Air Force General Hospital, PLA, Beijing, China. All 5 lncRNAs, including INHBA-AS1, MIR4435-2HG, CEBPA-AS1, UCA1 and AK001058 were increased in the GC plasma (Figure 2A–2C, 2E), exhibiting the same trend as in the GC tissues. With an AUC of 0.855, 0.882, 0.785, 0.728, and 0.852 for INHBA-AS1, MIR4435-2HG, CEBPA-AS1, UCA1, and AK001058, respectively (Figure 2F), the AUC for UCA1 was lowest in the five lncRNAs. Furthermore, the AUC of the 5-lncRNA panel was 0.920 (95%CI: 0.860–0.981) (Figure 2G), when the AUC of a single lncRNA was lower than that of the 5-lncRNA signature. And a five-minus-one lncRNA signature was constructed, the five-minus-UCA1 lncRNA signature has a higher AUC (AUC = 0.921; 95%CI: 0.862–0.981) than the others and the 5-lncRNA signature (Figure 2G). We also constructed a 3-lncRNA signature (Figure 2H, Supplementary Table 2) and 2-lncRNA signature (Supplementary Table 3), randomly chose from the 5 lncRNAs, their AUCs were also lower than the five-minus-UCA1 lncRNA signature. So a new 4-lncRNA signature, including INHBA-AS1, MIR4435-2HG, CEBPA-AS1, and AK001058, was constructed. The relationship between lncRNA levels in the plasma and the clinicopathological features of GC is shown in Table 2. The expression level of INHBA-AS1 was associated with tumor size (Figure 3A) and tumor grade (Figure 3B); There is a correlation between the expression level of AK001058 and depth of tumor invasion (Figure 3C) and TNM stage (Figure 3D).

Validation of the utility of candidate lncRNAs in the testing set

We validated the 4-lncRNA panel using the same method in a test cohort of 47 GC patients and 52 healthy controls, the plasma samples were collected from the China-Japan Union Hospital of Jilin University, Changchun, China. In the testing set, all the four lncRNAs, including INHBA-AS1, MIR4435-2HG, CEBPA-AS1, and AK001058 were also increased in the plasma samples from GC patients (Figure 4A–4D), and the predictor was also remarkably stable, with AUCs of 0.752, 0.817, 0.819, and 0.820 for INHBA-AS1, MIR4435-2HG, CEBPA-AS1, and AK001058, respectively (Figure 4E). In addition, the AUC value of the 4-lncRNA panel was 0.902 (Figure 4F). Additionally, we constructed 4-minus-one lncRNA signatures by excluding each lncRNA from the set individually and comparing the AUCs of these three lncRNA signatures
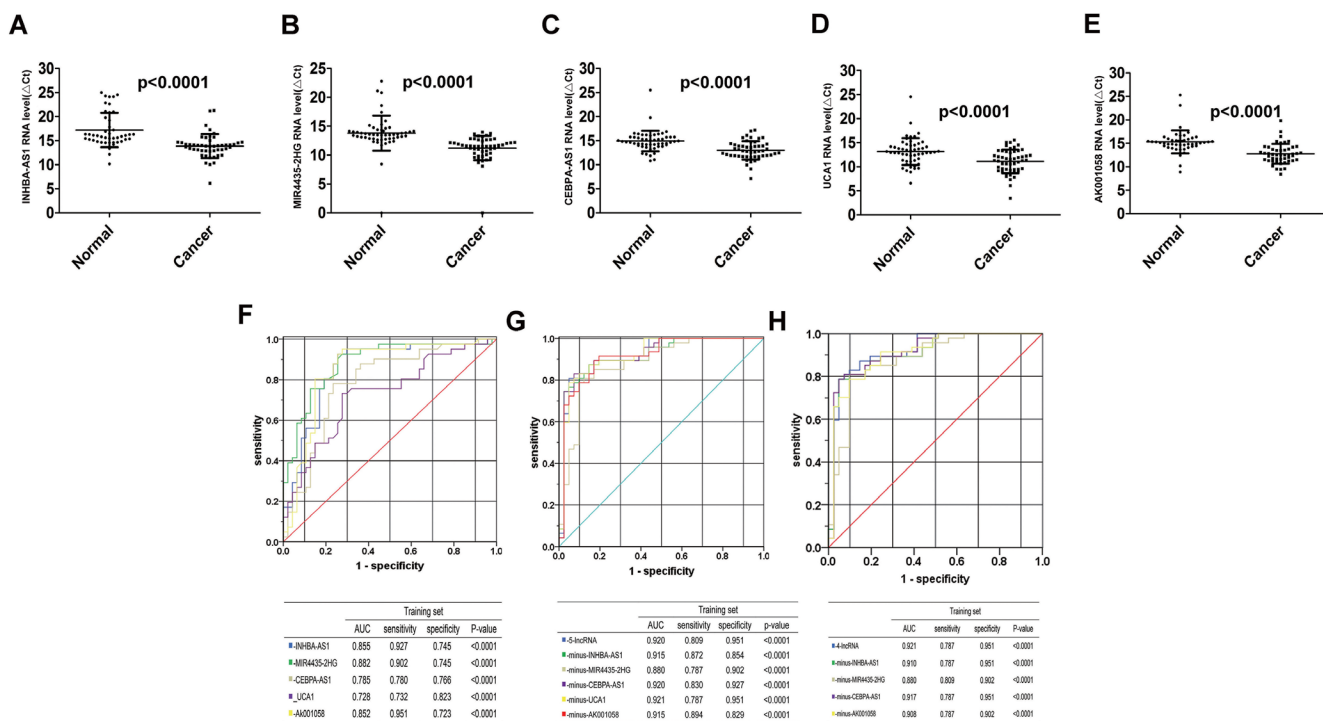

Figure 2: Distribution of lncRNAs in plasma and AUC for lncRNA panels in training set. Distribution of lncRNA-INHBA-AS1 (**A**), MIR4435-2HG (**B**), CEBPA-AS1 (**C**), UCA1 (**D**), and AK001058 (**E**) levels in the plasma from patients and healthy controls in the training set. $\Delta Ct_{target\ gene\ in\ cancer} = Ct_{target\ gene\ in\ cancer} - Ct_{18S\ rRNA\ in\ cancer}$; $\Delta Ct_{target\ gene\ in\ normal} = Ct_{target\ gene\ in\ normal} - Ct_{18S\ rRNA\ in\ normal}$. And the ROCs of the 5 lncRNA (**F**), 5-lncRNA and five-minus-one lncRNA (**G**), and a new 4-lncRNA signature and four-minus-one lncRNA (**H**) in training set expressed in plasma, respectively. The AUCs and *p*-values are listed in the picture.

Table 2: Correlation between lncRNA-INHBA-AS1, MIR4435-2HG, CEBPA-AS1, UCA1, and AK001058 panel expression levels in GC plasma and clinical parameters

Clinical parameters	Number of cases	*p*-value				
		INHBA-AS1	MIR4435-2HG	CEBPA-AS1	UCA1	AK001058
Age(years)						
< 60	23	0.0380(*)	0.5922	0.0764	0.1033	0.6995
> 60	28					
Gender						
female	16	0.3604	0.4214	0.3135	0.1274	0.5364
male	35					
Tumor size(cm)						
≥ 4	21	0.0181(*)	0.8777	0.7252	0.9587	0.8750
< 4	12					
Tumor grade						
Low	26	0.0479(*)	0.6296	0.3396	0.5124	0.1324
middle or high	25					
LN metastasis						
No	14	0.2922	0.4839	0.2764	0.2712	0.1322
Yes	29					
Depth of invasion						
T1-2	21	0.0599	0.0556	0.0614	0.0542	0.0053(**)
T3-4	27					
TNM stage						
I	18	0.3756	0.4137	0.5555	0.6802	0.0281(*)
II-IV	30					

Abbreviations: lncRNA, long noncoding RNA; LN, lymph node; TNM, tumor node metastasis; *Statistically Significant ($p < 0.05$).

with that of the original 4-lncRNA signature. None of the four-minus-one lncRNA signatures had a higher AUC value in the training set (Figure 2H) and testing sets (Figure 4F) than those of the 4-lncRNA signature. The AUCs of a 2-lncRNA signature, randomly chose from the 4 lncRNAs were also lower than the 4-lncRNA signature (Supplementary Table 4) in testing sets.

The expression level of each lncRNA fragment is different in plasma

Different from miRNAs, complete lncRNA molecule in plasma were little. Plasma lncRNAs exist mainly in the form of RNA fragments. The contents of lncRNA fragments in plasma might be different from each other. Therefore, it is necessary to evaluate the expression level of different RNA fragments from the same lncRNA in plasma to employ lncRNAs as biomarkers. LncRNA MIR4435-2HG, which was increased in GC plasma, was selected for the following test. MIR4435-2HG is a lncRNA with 3887bp, we designed 13 pairs of primers to detect the expression levels of different RNA fragments from MIR4435-2HG (Figure 5A, Supplementary Table 5)

by qRT-PCR. Each of fragments is ~300bp, among these possible RNA fragments, i and j were not detected in all GC and healthy control plasma samples. Fragment a could be detected, but its expression level in GC plasma is not significantly different from that in healthy control samples. The levels of fragment b, c, d, g, k, l, m in GC plasma samples are much higher than those in healthy controls (Figure 5B). According to the results, fragment b, c, d, g, k, l or m may be more potential to serve as a biomarker. These data suggest that it is necessary to identify the most stable and differential expressed RNA fragment as diagnostic and prognostic biomarker.

DISCUSSION

Gastroscopy [20] and computed tomography (PET/CT) [21] are the commonly used tools for the diagnosis of GC. Carcinoembryonic antigen (CEA), CA19-9 [22], and CA-724 [23] are classic tumor markers commonly used in the management of GC [24]. Given that these markers with low diagnostic sensitivity and specificity have limited utility [25], discovery of new biomarkers has attracted attention for early detection of GC.

An increasing number of studies have demonstrated that circulating nucleic acids (CNAs) can be found in cell-free plasma, serum, and other body fluids. For example, miRNAs are differentially expressed in various types of cancers and possess good diagnostic value in cancer screening [26–28]. In some cases, circulating cell-free DNAs (cfDNAs) are also dysregulated in the plasma and serum of cancer patients [29, 30]. In non-small cell lung cancer, cfDNA could be served as a prognostic and predictive biomarker [31].

Given that plasma is relatively easy to obtain and genome-wide screening approaches provide us new opportunities to develop novel diagnostic or prognostic markers, blood based testing is ideal for evaluating biomarkers in cancer. Some lncRNAs or miRNAs that stably exist in plasma may be protected by exosome encapsulation, microvesicles, apoptotic bodies [32, 33] and complex formation with proteins [34]. Therefore, lncRNAs can serve as promising markers for early diagnosis of cancers.

In our study, we systematically determined the expression levels of nine cancer-associated lncRNAs dysregulated in two gastric tissue microarrays and detected
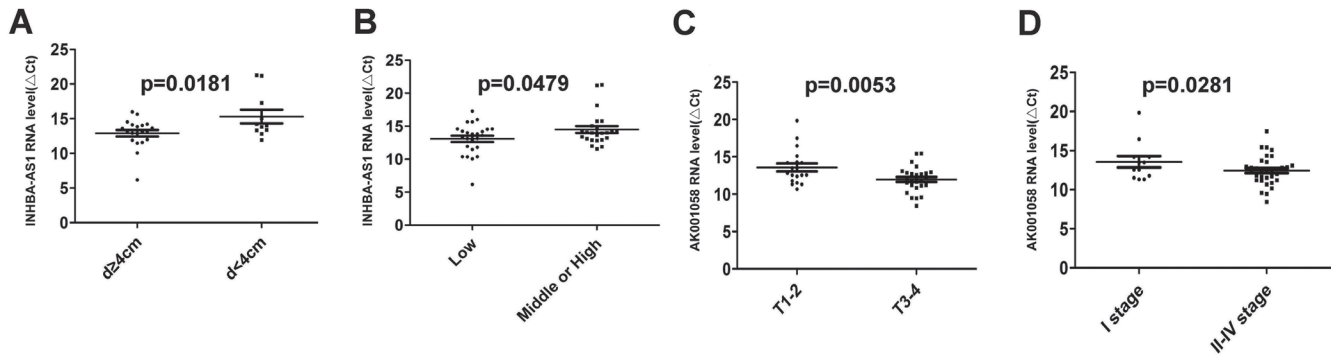

Figure 3: The relationship between lncRNA levels and the clinicopathological features of GC patients in training set. The relationship between the expression level of INHBA-AS1 and tumor size (**A**) and tumor grade (**B**); the relationship between the expression level of AK001058 and the depth of tumor invasion (**C**) and TNM stage (**D**). p-values are listed in the picture. $\Delta Ct_{target\ gene\ in\ cancer} = Ct_{target\ gene\ in\ cancer} - Ct_{18S\ rRNA\ in\ cancer}$.
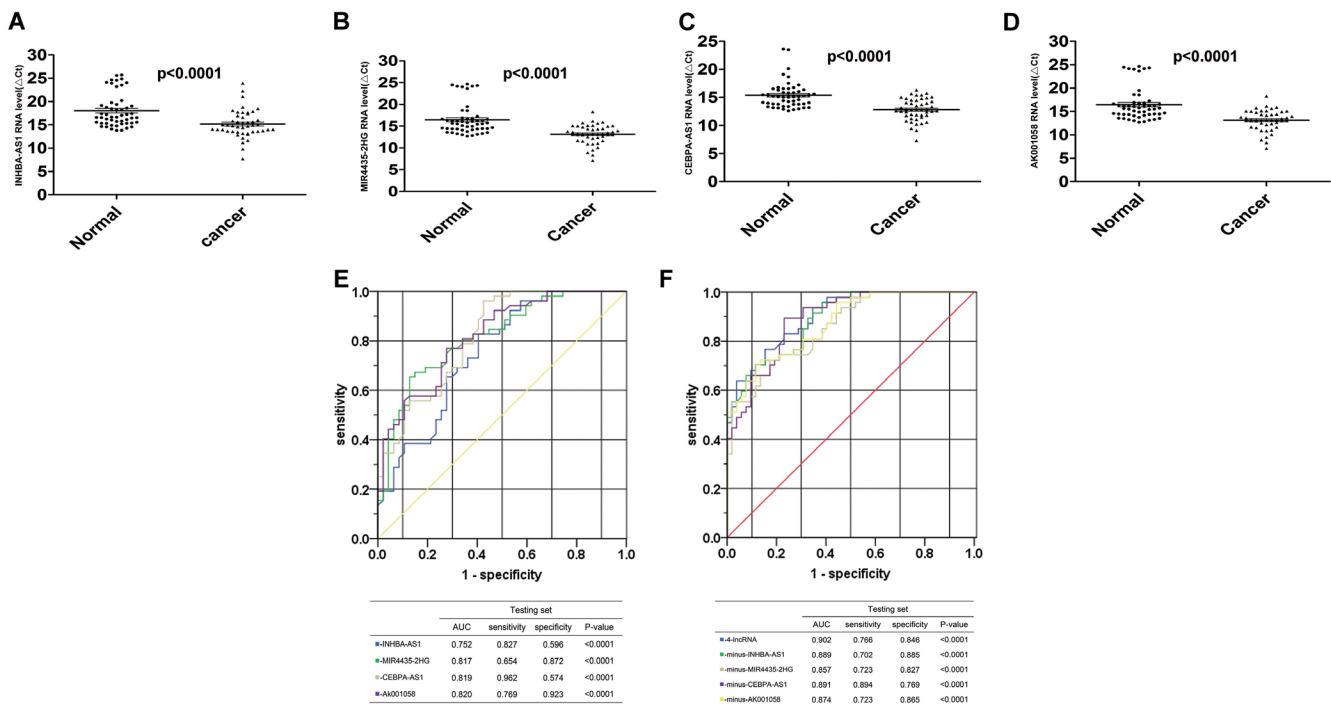

Figure 4: Distribution of lncRNAs in plasma and AUC for lncRNA panels in testing set. Distribution of lncRNA-INHBA-AS1 (**A**), MIR4435-2HG (**B**), CEBPA-AS1 (**C**) and AK001058 (**D**) levels in the plasma from patients and healthy controls in the testing set. $\Delta Ct_{target\ gene\ in\ cancer} = Ct_{target\ gene\ in\ cancer} - Ct_{18S\ rRNA\ in\ cancer}$; $\Delta Ct_{target\ gene\ in\ normal} = Ct_{target\ gene\ in\ normal} - Ct_{18S\ rRNA\ in\ normal}$. The ROCs of the 4 lncRNA (**E**), 4-lncRNA and four-minus-one lncRNA (**F**) expressed in plasma in testing set. The AUCs and p-values are listed in the picture.

their relative expression levels in paired gastric tissues by qRT-PCR. Expression of CEBPA protein was increased in HCCs [35], and epigenetic aberrations in regulating CEBPA expression contributed to leukemic transformation in acute myeloid leukemia[36]. INHBA mRNA and protein expression is commonly elevated in primary human NSCLC, and it promotes tumor metastasis [37]. Here, we found that there is a positive relation between the RNA levels of INHBA-AS1 and CEBPA-AS1 with their PCGs in GC tissues, respectively (Supplementary Figure 4A, 4B). These two lncRNAs might influence the expression of their nearby PCGs through certain mechanism. Interestingly, CEBPA-AS1 was shown to be decreased in two GC tissues by microarray analysis, it is increased in most GC tissues in fact (Figure 1C). That maybe be because it is decreased in the two GC tissues by chance. Furthermore, Ak001058 is increased in 47 (95.92%) of the 49 GC tissues (Figure 1E), and the AUC value for AK001058 is 0.957 (Supplementary Figure 3A). These results suggest that Ak001058 is very sensitive, might be developed into a novel biomarker for diagnosis of GC.

With lower AUC value, UCA1 expression level was not changed remarkably in GC plasma. Notably, we established a 4-lncRNA signature not including UCA1,

and our results showed that the 4-lncRNA panel was highly indicative of GC diagnosis. The AUC values of this 4-lncRNA panel were 0.921 and 0.902 in the training and testing sets, respectively, which was higher than that of any panels (Figures 2, 4). Among the four lncRNAs, the RNA level of AK001058 was higher in GC advanced stage (Figure 3), suggesting that this gene can predict GC at a relatively early stage.

LncRNA MIR4435-2HG is related to cell-cycle in lung cancer cells [38]. Here, we found that MIR4435-2HG was also increased in GC tissues. Interestingly, our results demonstrate that different fragments of MIR4435-2HG have different expression levels, and the level of the same fragment was different in GC plasma and healthy controls. The primer of MIR4435-2HG in this study is complementary with fragment b. All these data claims that lncRNAs exist mainly in the form of RNA fragments in plasma. Therefore, it is necessary to identify the most stable and differential expressed RNA fragment as diagnostic biomarker.

Taken together, the results show that the plasma 4-lncRNA panel constructed could not only discriminate GC patients from healthy controls, but also displayed potential use for diagnosing GC at an early stage. Not
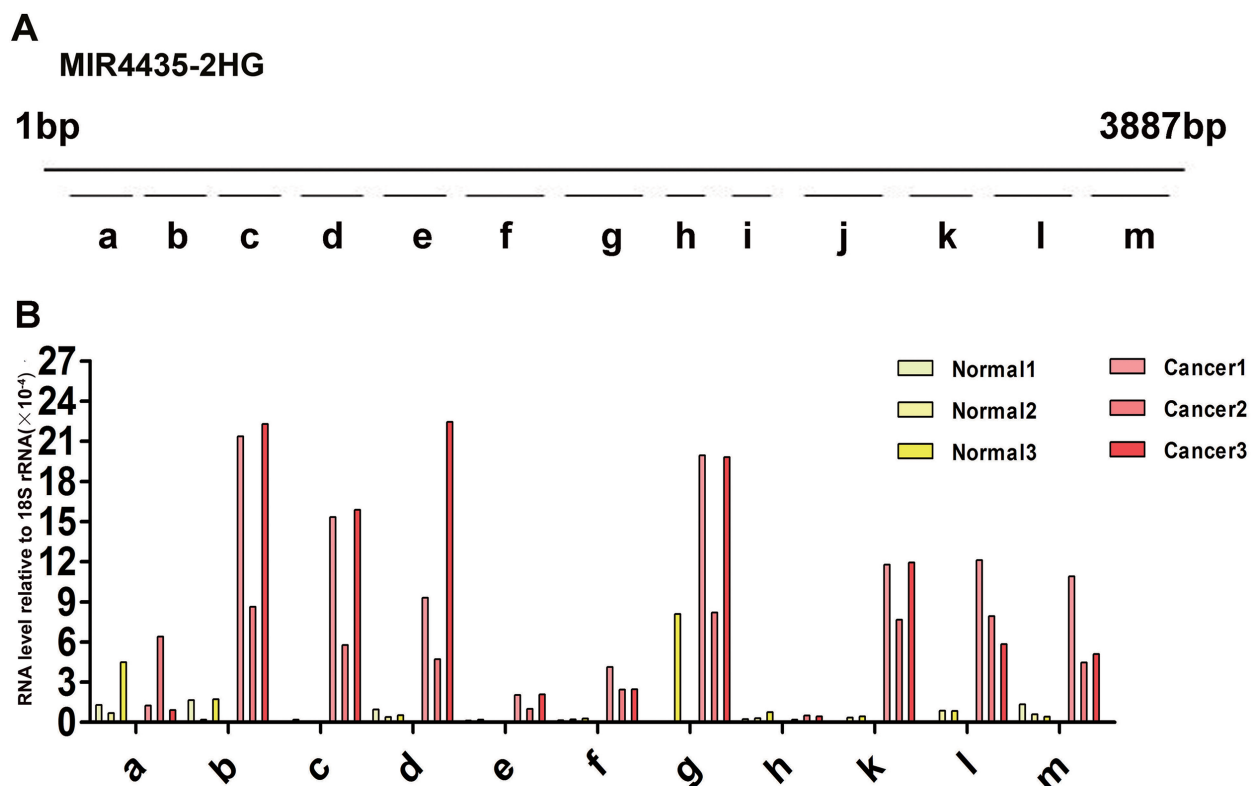

Figure 5: Expression of diverse MIR4435-2HG fragments in health control and gastric cancer plasma. (A) 13 different primer pairs of MIR4435-2HG gene, a-m fragments are the qRT-PCR products with different length. a: 186–291(106 bp); b: 424–529(106 bp); c: 728–833(106 bp); d: 1035–1151(117 bp); e: 1257–1356(100 bp);f: 1530–166(136bp); g: 1955–2079(125 bp); h: 2270-2359(90 bp); i: 2609–2692(84 bp); j: 2716–2841(126 bp); k: 3164–3273(110 bp); l: 3319-3457(139bp); m: 3657–3791(135 bp). (B) Expression level of MIR4435-2HG fragments in three GC patient and health control samples relative to 18S rRNA. Fragment i, j cannot be detected among the six plasma samples, and fragment l can be detected excluding one sample of three healthy controls.

all, but some specific fragments of lncRNAs in plasma could be used as an appropriate biomarker. Our work may facilitate the detection of GC and serve as the basis for further studies of plasma lncRNAs in predicting personalized treatment strategies and efficacy of GC patients.

MATERIALS AND METHODS

Patients and clinical samples

A total of 49 primary cancer tissues and the paired adjacent non-tumor tissue were collected from patients who underwent surgery for GC at the China-Japan Union Hospital of Jilin University. These tissues were flash frozen in liquid nitrogen immediately after surgery and subsequently stored at −80°C. The plasma samples were collected from the China-Japan Union Hospital of Jilin University and Air Force General Hospital, PLA. All plasma samples were stored at −80°C. No patients received anticancer treatments before surgery in this study. All samples were staged in accordance with the tumor node metastasis (TNM) classification and criteria of the Union for International Cancer Control (UICC), and tumor grade was assessed according to the UICC criteria. Written informed consent was obtained from all patients(Supplementary Tables 6, 7, 8). The Ethics Committee of the China-Japan Union Hospital of Jilin University and Air Force General Hospital, PLA approved the use of samples for this study.

Sample preparation and tissue microarray

Two groups of gastric tissue samples, including GC tissue and adjacent non-tumor tissue, were prepared. RNA was extracted from the frozen tissue using TRI reagent (Sigma, USA), RNA integrity was analyzed by denatured agarose gel electrophoresis. Microarray hybridization was performed by Kangchen Biotech, Shanghai P. R. China.

Tissue and plasma RNA extraction

Total tissue RNA was extracted from the frozen tissue block using TRI reagent (Sigma, USA) according to the manufacturer's protocol. Plasma cell-free RNA was extracted from 200 μL of plasma using the miRNeasy Serum/Plasma Kit (QIAGEN, German) according to the manufacture's protocol. All RNA and cDNA products were stored at −80°C until use.

Reverse transcription (RT) and quantitative real-time polymerase chain reaction (qRT-PCR) validation

Total RNA was reverse-transcribed using ImProm II Reverse Transcriptase (Promega, USA) according to the manufacturer's instructions. qRT-PCR was performed with SYBR Premix ExTaq (TaKaRa, Japan) on an MX3000p instrument (Agilent Technologies, USA) according to the manufacturer's protocol. For tissues, RT products were diluted for 100 times, and the dilutions were used as new templates to detect the expression level of 18S rRNA. Amplification of the appropriate product was confirmed by a melting curve analysis following amplification. The relative expression of each lncRNA was calculated using comparative cycle threshold (CT) ($2^{-\Delta Ct}$) method with 18S rRNA as the endogenous control for data normalization. The method of analyzing relative gene expression is described in detail in this paper [39].

Statistical analysis

The statistical analysis was performed using the SPSS 17.0 and GraphPad Prism 5.0 softwares. Student's *t*-test was used to evaluate differences in the expression of the chosen lncRNAs in tissues and plasmas from the GC patients and healthy controls. The specificity, sensitivity, and AUC for the lncRNA levels were determined using an ROC analysis. Using the binary outcome of GC and healthy control samples as dependent variables, a logistic regression model was established using the stepwise model selection method. All of the statistical tests were two tailed, and *P* value < 0.05 was considered statistically significant.

CONCLUSIONS

One of the major strengths of this study is that we found that the combination of plasma lncRNAs is a better indicator of GC, and these four lncRNAs might be used as diagnostic or prognostic markers and therapeutic targets for GC patients.

We also identified that the expression levels of a series of MIR4435-2HG fragments are different in GC plasma samples, these results imply that the expression level of different RNA fragments from the same lncRNA in plasma to employ lncRNAs as biomarkers.

Abbreviations

GC: gastric cancer; AUC: Area under the curves Cl: Confidence interval, ROC: Receiver operating characteristic; lncRNAs: long noncoding RNAs, RT: Reverse transcription; TNM: tumor node metastasis, qRT-PCR: quantitative real-time polymerase chain reaction; NT: non-tumor, UICC: the Union for International Cancer Control; Ct: cycle threshold, miRNAs: microRNAs, CNAs: circulating nucleic acids; cfDNAs: circulating cell-free DNAs, UCA1: urothelial cancer associated 1, CEBPA: CCAAT/enhancer binding protein alpha, CEBPA-AS1: CCAAT/enhancer binding protein alpha antisense RNA 1, INHBA: inhibin beta A subunit, INHBA-AS1:

inhibin beta A subunit antisense RNA1, MIR4435-2HG: microRNA 4435-2 host gene.

Authors' contributions

H Fu, X Fang, and X Zheng conceived of the project and contributed to study design; D Ke, H Li, Y Zhang and Y An performed data collection; D ke, H Fu, X Fang, and X Zheng analyzed and interpreted the data; D Ke, H Fu, and X Zheng wrote the manuscript with input from all authors; all authors read and approved the final version of the manuscript.

REFERENCES

1. Shao Y, Ye M, Jiang X, Sun W, Ding X, Liu Z, Ye G, Zhang X, Xiao B, Guo J. Gastric juice long noncoding RNA used as a tumor marker for screening gastric cancer. Cancer. 2014; 120:3320–3328.

2. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin. 2011; 61:69–90.

3. Hartgrink HH, Jansen EP, van Grieken NC, van de Velde CJ. Gastric cancer. Lancet. 2009; 374:477–490.

4. Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. Nature. 2012; 482:339–346.

5. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem. 2012; 81:145–166.

6. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. Cell. 2009; 136:629–641.

7. Wu Y, Wang YQ, Weng WW, Zhang QY, Yang XQ, Gan HL, Yang YS, Zhang PP, Sun MH, Xu MD, Wang CF. A serum-circulating long noncoding RNA signature can discriminate between patients with clear cell renal cell carcinoma and healthy controls. Oncogenesis. 2016; 5:e192.

8. Li T, Xie J, Shen C, Cheng D, Shi Y, Wu Z, Deng X, Chen H, Shen B, Peng C, Li H, Zhan Q, Zhu Z.

9. Yang X, Song JH, Cheng Y, Wu W, Bhagat T, Yu Y, Abraham JM, Ibrahim S, Ravich W, Roland BC, Khashab M, Singh VK, Shin EJ, et al. Long non-coding RNA HNF1A-AS1 regulates proliferation and migration in oesophageal adenocarcinoma cells. Gut. 2014; 63:881–890.

10. Zhou X, Yin C, Dang Y, Ye F, Zhang G. Identification of the long non-coding RNA H19 in plasma as a novel biomarker for diagnosis of gastric cancer. Sci Rep. 2015; 5:11516.

11. Zhao Y, Guo Q, Chen J, Hu J, Wang S, Sun Y. Role of long non-coding RNA HULC in cell proliferation, apoptosis and tumor metastasis of gastric cancer: a clinical and *in vitro* investigation. Oncol Rep. 2014; 31:358–364.

12. Lin X, Yang M, Xia T, Guo J. Increased expression of long noncoding RNA ABHD11-AS1 in gastric cancer and its clinical significance. Med Oncol. 2014; 31:42.

13. Dong L, Qi P, Xu MD, Ni SJ, Huang D, Xu QH, Weng WW, Tan C, Sheng WQ, Zhou XY, Du X. Circulating CUDR, LSINCT-5 and PTENP1 long noncoding RNAs in sera distinguish patients with gastric cancer from healthy controls. Int J Cancer. 2015; 137:1128–1135.

14. Xiang JF, Yin QF, Chen T, Zhang Y, Zhang XO, Wu Z, Zhang S, Wang HB, Ge J, Lu X, Yang L, Chen LL. Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. Cell Res. 2014; 24:513–531.

15. March-Villalba JA, Martinez-Jabaloyas JM, Herrero MJ, Santamaria J, Alino SF, Dasi F. Cell-free circulating plasma hTERT mRNA is a useful marker for prostate cancer diagnosis and is associated with poor prognosis tumor characteristics. PLoS One. 2012; 7:e43470.

16. Lledo SM, Garcia-Granero E, Dasi F, Ripoli R, Garcia SA, Cervantes A, Alino SF. Real time quantification in plasma of human telomerase reverse transcriptase (hTERT) mRNA in patients with colorectal cancer. Colorectal Dis. 2004; 6:236–242.

17. Albrecht AS, Orom UA. Bidirectional expression of long ncRNA/protein-coding gene pairs in cancer. Brief Funct Genomics. 2016; 15:167–173.

18. Postepska-Igielska A, Giwojna A, Gasri-Plotnitsky L, Schmitt N, Dold A, Ginsberg D, Grummt I. LncRNA Khps1 Regulates Expression of the Proto-oncogene SPHK1 via Triplex-Mediated Changes in Chromatin Structure. Mol Cell. 2015; 60:626–636.

19. Tran NT, Su H, Khodadadi-Jamayran A, Lin S, Zhang L, Zhou D, Pawlik KM, Townes TM, Chen Y, Mulloy JC, Zhao X. The AS-RBM15 lncRNA enhances RBM15 protein translation during megakaryocyte differentiation. EMBO Rep. 2016; 17:887–900.

20. Liu DY, Gan T, Rao NN, Xing YW, Zheng J, Li S, Luo CS, Zhou ZJ, Wan YL. Identification of lesion images from gastrointestinal endoscope based on feature extraction of

combinational methods with and without learning process. Med Image Anal. 2016; 32:281–294.

21. Nakajo M, Kajiya Y, Tani A, Jinguji M, Nakajo M, Yoshiura T. FLT-PET/CT diagnosis of primary and metastatic nodal lesions of gastric cancer: comparison with FDG-PET/CT. Abdom Radiol (NY). 2016; 41:1891–8.

22. Sun Z, Zhang N. Clinical evaluation of CEA, CA19-9, CA72-4 and CA125 in gastric cancer patients with neoadjuvant chemotherapy. World J Surg Oncol. 2014; 12:397.

23. Chen XZ, Zhang WK, Yang K, Wang LL, Liu J, Wang L, Hu JK, Zhang B, Chen ZX, Chen JP, Zhou ZG, Mo XM. Correlation between serum CA724 and gastric cancer: multiple analyses based on Chinese population. Mol Biol Rep. 2012; 39:9031–9039.

24. Kim HJ, Lee KW, Kim YJ, Oh DY, Kim JH, Im SA, Lee JS. Chemotherapy-induced transient CEA and CA19-9 surges in patients with metastatic or recurrent gastric cancer. Acta Oncol. 2009; 48:385–390.

25. He CZ, Zhang KH, Li Q, Liu XH, Hong Y, Lv NH. Combined use of AFP, CEA, CA125 and CAl9-9 improves the sensitivity for the diagnosis of gastric cancer. BMC Gastroenterol. 2013; 13:87.

26. Samsonov R, Burdakov V, Shtam T, Radzhabovsmall a CZ, Vasilyev D, Tsyrlina E, Titov S, Ivanov M, Berstein L, Filatov M, Kolesnikov N, Gil-Henn H, Malek A. Plasma exosomal miR-21 and miR-181a differentiates follicular from papillary thyroid cancer. Tumour Biol. 2016; 37:12011–12021.

27. Cao Z, Liu C, Xu J, You L, Wang C, Lou W, Sun B, Miao Y, Liu X, Wang X, Zhang T, Zhao Y. Plasma microRNA panels to diagnose pancreatic cancer: Results from a multicenter study. Oncotarget. 2016; 7:41575–41583. doi: 10.18632/oncotarget.9491.

28. Huang Z, Huang D, Ni S, Peng Z, Sheng W, Du X. Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer. Int J Cancer. 2010; 127:118–126.

29. Liao W, Yang H, Xu H, Wang Y, Ge P, Ren J, Xu W, Lu X, Sang X, Zhong S, Zhang H, Mao Y. Noninvasive detection of tumor-associated mutations from circulating cell-free DNA in hepatocellular carcinoma patients by targeted deep sequencing. Oncotarget. 2016; 7:40481–40490. doi: 10.18632/oncotarget.9629.

30. Pathak AK, Bhutani M, Kumar S, Mohan A, Guleria R. Circulating cell-free DNA in plasma/serum of lung cancer patients as a potential screening and prognostic tool. Clin Chem. 2006; 52:1833–1842.

31. Ai B, Liu H, Huang Y, Peng P. Circulating cell-free DNA as a prognostic and predictive biomarker in non-small cell lung cancer. Oncotarget. 2016; 7:44583–44595. doi: 10.18632/oncotarget.10069.

32. Kogure T, Yan IK, Lin WL, Patel T. Extracellular Vesicle-Mediated Transfer of a Novel Long Noncoding RNA TUC339: A Mechanism of Intercellular Signaling in Human Hepatocellular Cancer. Genes Cancer. 2013; 4:261–272. doi: 10.1177/1947601913499020.

33. Hasselmann DO, Rappl G, Tilgen W, Reinhold U. Extracellular tyrosinase mRNA within apoptotic bodies is protected from degradation in human serum. Clin Chem. 2001; 47:1488–1489.

34. Irion U, St Johnston D. bicoid RNA localization requires specific binding of an endosomal sorting complex. Nature. 2007; 445:554–558.

35. Lu GD, Ang YH, Zhou J, Tamilarasi J, Yan B, Lim YC, Srivastava S, Salto-Tellez M, Hui KM, Shen HM, Nguyen LN, Tan BC, Silver DL, et al. CCAAT/enhancer binding protein alpha predicts poorer prognosis and prevents energy starvation-induced cell death in hepatocellular carcinoma. Hepatology. 2015; 61:965–978.

36. Hackanson B, Bennett KL, Brena RM, Jiang J, Claus R, Chen SS, Blagitko-Dorfs N, Maharry K, Whitman SP, Schmittgen TD, Lubbert M, Marcucci G, Bloomfield CD, et al. Epigenetic modification of CCAAT/enhancer binding protein alpha expression in acute myeloid leukemia. Cancer Res. 2008; 68:3142–3151.

37. Wamsley JJ, Kumar M, Allison DF, Clift SH, Holzknecht CM, Szymura SJ, Hoang SA, Xu X, Moskaluk CA, Jones DR, Bekiranov S, Mayo MW. Activin upregulation by NF-kappaB is required to maintain mesenchymal features of cancer stem-like cells in non-small cell lung cancer. Cancer Res. 2015; 75:426–435.

38. Yang Q, Xu E, Dai J, Liu B, Han Z, Wu J, Zhang S, Peng B, Zhang Y, Jiang Y. A novel long noncoding RNA AK001796 acts as an oncogene and is involved in cell growth inhibition by resveratrol in lung cancer. Toxicol Appl Pharmacol. 2015; 285:79–88.

39. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001; 25:402–408.