

Identification of diagnostic long non-coding RNA biomarkers in patients with hepatocellular carcinoma

GANG LI, HAO SHI, XINYI WANG, BEI WANG, QIANQIAN QU, HAIYANG GENG and HONGJUN SUN

Department of Medical Imaging, Qianfoshan Hospital Affiliated to Shandong University, Jinan, Shandong 250014, P.R. China

Received August 3, 2018; Accepted May 9, 2019

DOI: 10.3892/mmr.2019.10307

Abstract. Liver cancer is a leading cause of cancer-associated mortality worldwide. Hepatocellular carcinoma (HCC) is the most common subtype of liver cancer. The aim of the present study was to identify long non-coding RNA (lncRNAs) as diagnostic biomarkers for HCC. The lncRNA and mRNA expression profiles of a large group of patients with HCC were obtained from The Cancer Genome Atlas. The differentially expressed lncRNAs (DELncRNAs) and the differentially expressed mRNAs (DEmRNAs) were identified by bioinformatics analysis. Using feature selection procedure and a classification model, the optimal diagnostic lncRNA biomarkers for HCC were identified. Classification models, including random forests, decision tree and support vector machine (SVM), were established to distinguish between HCC and normal tissues. DEmRNAs co-expressed with the lncRNAs were considered as targets of DELncRNAs. Functional annotation of DEmRNAs co-expressed with these lncRNAs biomarkers was performed. Receiver operating characteristic curve analysis of lncRNAs biomarkers was conducted. A total of 3,177 lncRNAs and 15,183 mRNAs between HCC and normal tissues were obtained. RP11-486O12.2, RP11-863K10.7, LINC01093 and RP11-273G15.2 were identified as optimal diagnostic lncRNA biomarkers for HCC that were co-expressed with 273, 69, 76 and 1 DEmRNAs, respectively. The area under the curve values of the random forest model, decision tree model and SVM model were 0.992, 0.927 and 0.992, and the specificity and sensitivity of the three models were 100.0 and 95.6, 92.0 and 98.3 and 98.0 and 97.2%, respectively. 'PPAR signaling pathway' and

'retinol metabolism' were two significantly enriched target pathways of DELncRNAs. The present study identified four DELncRNAs, including RP11-486O12.2, RP11-863K10.7, LINC01093 and RP11-273G15.2, as potential diagnostic biomarkers of HCC. Functional annotation of target DEmRNAs provided novel evidence for examining the precise roles of lncRNA in HCC.

Introduction

Liver cancer is one of the leading causes of cancer-associated mortality worldwide, and hepatocellular carcinoma (HCC) is the primary type of liver cancer, accounting for ~90% of liver cancer cases (1). Previous studies have indicated that the incidence rate of HCC is increasing with a poor 5-year survival rate of ~7% (2,3). In the early stages of HCC, the tumor can be removed effectively with a relatively long survival time, while advanced HCC is associated with poor survival, due to distant organ metastasis (4,5). Therefore, further investigation of HCC diagnostic biomarkers is required. However, an early diagnosis of HCC is not always effaceable. In addition, to the best of our knowledge, an effective HCC treatment has not yet been established.

Long non-coding RNAs (lncRNAs) are a group of non-coding RNAs >200 nucleotides and have emerged as important regulators in the majority of biological processes (6). Recent advances have suggested that lncRNA dysregulation is linked to the development of pathophysiological processes, including development, invasion and metastasis of cancer (7,8).

Increasing evidence suggests that the expression of dysregulated lncRNA has been implicated in the development and progression of tumors (9-12). The emerging roles of lncRNAs in HCC have also been investigated in previous studies. For example, the downregulated expression of the lncRNA, growth arrest specific 5 (GAS5), promotes cell invasion in HCC, by regulating vimentin, and is associated with a relatively poor prognosis (9). Colon cancer associated transcript 1 (CARLo-5) has been considered as an independent risk factor for disease-free survival in HCC as it may accelerate the metastasis of HCC, making it a potential novel therapeutic biomarker (10). Prostate cancer associated transcript 1 (PCAT-1) has been reported to be significantly upregulated in HCC tissues, and significantly associated with the overall survival time of patients with HCC (13). ZNFx1 antisense RNA 1 (ZFAS1) may function as a cancer gene in

Correspondence to: Professor Hongjun Sun, Department of Medical Imaging, Qianfoshan Hospital Affiliated to Shandong University, 16766 Jingshi Road, Jinan, Shandong 250014, P.R. China
E-mail: hongjunsun_doctor1@163.com

Abbreviations: lncRNAs, long noncoding RNAs; DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; HCC, hepatocellular carcinoma

Key words: hepatocellular carcinoma, biomarker, long non-coding RNAs, mRNAs, the cancer genome atlas

HCC progression, as it has been proved to bind microRNA-150 and abolish its tumor suppressive function, promoting the expression of matrix metalloproteinase (MMP14), MMP16 and zinc finger E-box-binding homeobox 1 (14). The upregulated expression of the lncRNAs metastasis associated lung adenocarcinoma transcript 1 and hepatocellular carcinoma upregulated long non-coding RNA (HULC) in HCC tissue may represent a good prognostic biomarker for curative resected HCC (15).

In the present study, lncRNA and mRNA expression data of a large number of patients with HCC was investigated using data from The Cancer Genome Atlas (TCGA). In addition, the study attempted to identify the optimal diagnostic lncRNA biomarkers by using feature selection and classification models. The functions of the potential lncRNA diagnostic biomarkers in HCC were further analyzed by the functional annotation of their co-expressed mRNAs.

Materials and methods

Eligible lncRNA and mRNA gene expression profiles of HCC in TCGA. From the database set-up until April 28th 2018, TCGA was searched for the genomic data of HCC. From TCGA data portal (tcga-data.nci.nih.gov/), lncRNA and mRNA gene expression profiles and clinical data of HCC and normal samples were downloaded. The present study included only patients who were histologically diagnosed as HCC. Finally, 361 HCC tissues and 50 normal adjacent samples from patients with HCC were included in this study.

Identification of differentially expressed mRNAs and lncRNAs between HCC and normal tissues. Undetectable lncRNAs and mRNAs, with a read count quantification of 0 in >20% HCC case or in >20% normal tissues were filtered and deleted. The differentially expressed lncRNAs (DElncRNAs) and mRNAs (DEmRNAs) in HCC compared to normal tissues were calculated using the R (version 3.3.3) package DESeq2 1.28.0 (16). The Benjamini and Hochberg multiple testing method (17) was applied to acquire the false discovery rate (FDR). FDR thresholds <0.05 and a log2-fold change >1 were used to define DElncRNAs and DEmRNAs. Using the R package pheatmap 0.7.4 (<https://cran.r-project.org/web/packages/pheatmap/index.html>), the hierarchical clustering analysis of DElncRNAs and DEmRNAs was performed.

Identification of the optimal diagnostic lncRNA biomarkers for HCC. To identify optimal diagnostic lncRNA biomarkers for HCC, feature selection procedures were performed as follows: The LASSO algorithm analysis was conducted using the 'glmnet' package (<https://cran.r-project.org/web/packages/glmnet/>) to decrease data dimensions. Single 10-fold cross-validation cycles were carried out using the coordinate descent algorithm for each fold and regularization parameters were indicated, resulting in the smallest average mean squared error across all folds. The optimal DElncRNAs were selected in HCC and normal tissue.

To further identify the optimal diagnostic lncRNA biomarkers for HCC, feature selection procedures were performed as follows. The importance value of each lncRNA was ranked from large to small, according to

Table I. HCC patient characteristics (n=361).

Parameter	Patients (n=361)
Gender	
Male	244
Female	117
Age	
<60	165
≥60	195
Unknown	1
Race	
Asian	156
Black or African American	17
White	176
Unknown	12
Tumor histologic grade	
G1	53
G2	171
G3	121
G4	11
Unknown	5
Tumor histologic stage	
Stage I	167
Stage II	82
Stage III	84
Stage IV	4

the decrease of mean accuracy using the random forest algorithm (randomForest; <https://cran.r-project.org/web/packages/randomForest/>); the optimum number of features was indicated by adding one DElncRNA at a time in the top down forward-wrapper packaging method; by using a support vector machine (SVM) at each increment.

The randomForest package was used to establish the random forest model. The 'rpart' package (<https://cran.r-project.org/web/packages/rpart/>) was used to build the decision tree model. The e1071 package (<https://cran.r-project.org/web/packages/e1071/index.html>) in R was used to establish the SVM model. The diagnostic ability of these three models and each lncRNA biomarker was evaluated by the receiver operating characteristic (ROC) area under curve (AUC), sensitivity and specificity.

Correlation between the four optimal diagnostic lncRNAs and clinical features. Analysis of the correlation between the four optimal lncRNAs and the clinical features, age and gender, was performed using Pearson's correlation coefficient. Analysis of the correlation between the four optimal lncRNAs and the clinical features, stage, grade and race, was performed using Spearman's correlation. *r* represents the correlation coefficient. The threshold for correlation was *r*>0.5 and *P*<0.05.

Survival analysis of optimal diagnostic lncRNA biomarkers for HCC. To determine potential associations between identified DElncRNAs and the survival of patients with

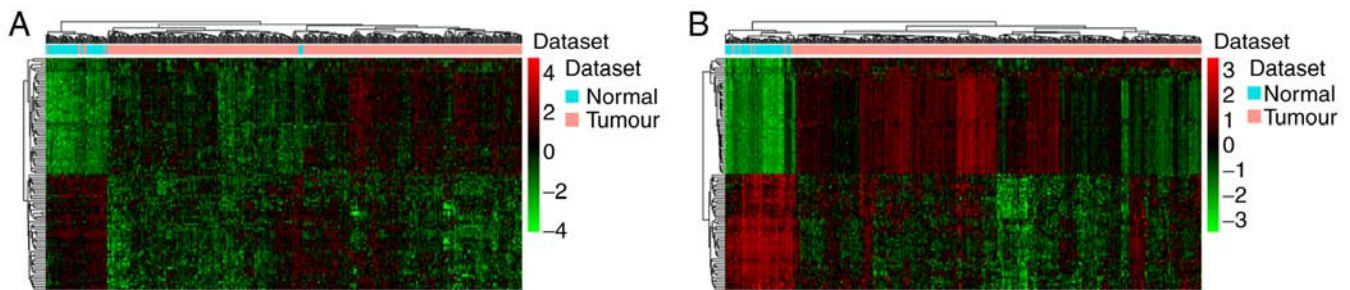


Figure 1. Hierarchical clustering analysis of DElncRNAs and the top 100 DEMRNAs between hepatocellular carcinoma and normal tissues. (A) DElncRNAs. (B) The top 100 DEMRNAs. Row and column represented DElncRNAs/DEMARNAs and tissue samples, respectively. The color scale represents the expression levels. Red and green colors represent up- and downregulation, respectively. DElncRNA, differentially expressed long non-coding RNA; DEMRNA, differentially expressed mRNA.

HCC, survival analysis was performed using the R survival package (version 2.44; <https://cran.r-project.org/web/packages/survival/index.html>). Univariate Cox regression analysis was performed for each DElncRNAs. $P < 0.05$ was considered statistically significant.

DEmRNAs co-expressed with the identified optimal diagnostic lncRNAs. The correlation between the optimal diagnostic lncRNAs and DEMRNAs were analyzed by the pairwise Pearson correlation coefficient. The threshold for DElncRNA-DEMARN co-expression pairs was $P < 0.05$ and $r > 0.6$. Cytoscape software 3.5.0 (cytoscape.org/) was used to construct the DElncRNA-DEMARN co-expression network.

Functional annotation. To disclose the biological functions and the potential pathways of the genes co-expressed with optimal DElncRNAs with diagnostic value for HCC, Gene Ontology (GO) classification (18,19) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (20-22) (<https://www.genome.jp/kegg/docs/relnote.html>) pathway enrichment analysis were performed using GeneCodis3 (genecodis.cnb.csic.es/analysis) online software. $FDR < 0.05$ was considered to indicate a statistically significant difference.

Results

Characteristics of patients with HCC. The clinical data of 361 patients with HCC were downloaded from TCGA. In the present study, 21.3 and 3.3% of patients had hepatitis B infection and hepatitis C infection, respectively. The incidence rate of HCC was indicated in 67.6 and 32.4%, in male and female patients, respectively. The incidence rate of HCC in Caucasians, Asians, and Black or African American was 48.7, 43.2 and 4.7%, respectively. The tumor stages were as follows: Stage I, 46.2%; stage II, 22.7%; stage III, 23.2%; and stage IV, 1.1% (Table I).

DEmRNAs and DElncRNAs between HCC and adjacent normal liver tissues. A total of 3,177 lncRNAs and 15,183 mRNAs were available for analysis following the exclusion of the minimally detected lncRNAs and mRNAs. The criteria of $FDR < 0.05$ and \log_2 -fold change > 1 were used to identify DElncRNAs and DEMARNAs. A total of 601 DElncRNAs, including 407 upregulated and 194 downregulated lncRNAs, and 2,920 DEMARNAs, including 1,930 upregulated and

990 downregulated mRNAs, were identified comparing HCC and adjacent normal tissues. Hierarchical clustering analysis of the top 100 DElncRNAs and DEMARNAs is presented in Fig. 1A and B, respectively.

Identification of the optimal diagnostic lncRNA biomarkers for HCC. Based on the reduced dimension of the data, comparing HCC and adjacent normal tissues identified 45 DElncRNAs using LASSO algorithm analysis (Table II). The random forest analysis was used to rank the 45 DElncRNAs, according to the decrease in mean accuracy (Fig. 2A). A 10-fold cross-validation result demonstrated that the average accuracy rate of four DElncRNAs, including RP11-486O12.2, RP11-863K10.7, LINC01093 and RP11-273G15.2, exhibited the highest score (Fig. 2B). Therefore, these four DElncRNAs were selected as the potential optimal diagnostic lncRNA biomarkers for HCC and were used to establish the random forests, decision tree and SVM models. Box-plot displayed the expression levels of four DElncRNAs between HCC and normal tissues (Fig. 2C-F). The AUC of the random forests model was 0.992 and the specificity and sensitivity of this model were 100.0 and 95.6%, respectively (Fig. 3A). The AUC of the decision tree model was 0.927 and the specificity and sensitivity of this model were 92.0% and 98.3, respectively (Fig. 3B). The AUC of the SVM model was 0.992, and the specificity and sensitivity of this model were 98.0 and 97.2% (Fig. 3C). The AUC of the combination of these four lncRNAs was > 0.85 , which indicated that the combination of these four lncRNAs was associated with HCC and serves a potential role in predicting the occurrence of HCC. The AUCs of these four lncRNAs, RP11-486O12.2, RP11-863K10.7, LINC01093 and RP11-273G15.2, were also > 0.85 (Fig. 3D-G). Correlation between these four lncRNAs and clinical features was also performed using Pearson's correlation coefficient or Spearman's correlation, and results showed that the four DElncRNAs were not associated with age, gender, stage, grade or race (Table III).

Co-expression of DEMARNAs and the identified optimal diagnostic lncRNA. A total of four optimal DElncRNA biomarkers for HCC were co-expressed with 388 DEMARNAs, accounting for 419 DElncRNA-DEMARN co-expression pairs. RP11-486O12.2, RP11-863K10.7, LINC01093 and RP11-273G15.2 were co-expressed with 273, 69, 76 and 1 DEMARNAs, respectively (Fig. 4).

Table II. Differentially expressed lncRNAs between HCC and normal tissues following reduced dimensions of the data.

lncRNA	log2 fold change	P-value	Regulation
CTD-2547G23.4	1.661885351	3.18x10 ⁻³⁸	Up
CTC-526N19.1	-2.226501065	8.58x10 ⁻³⁸	Down
RP11-250B2.6	-1.735228252	4.38x10 ⁻³⁴	Down
RP11-486O12.2	1.085632129	1.24x10 ⁻²⁹	Up
RP11-863K10.7	-2.44537751	5.53x10 ⁻²⁸	Down
AF131217.1	-2.550790337	7.72x10 ⁻²⁸	Down
LINC01093	-3.43026059	8.33x10 ⁻²⁶	Down
LINC00853	2.046868315	1.74x10 ⁻²⁵	Up
RP5-994D16.9	-1.016054876	7.17x10 ⁻²⁴	Down
LINC00907	-2.889956524	4.49x10 ⁻²²	Down
MIR3945HG	-2.002666478	1.58x10 ⁻¹⁹	Down
CTC-537E7.3	-2.630756292	2.58x10 ⁻¹⁹	Down
RP11-733O18.1	-1.727966554	1.26x10 ⁻¹⁸	Down
ASMTL-AS1	1.304698293	1.80x10 ⁻¹⁸	Up
AL161668.5	-2.049712877	2.28x10 ⁻¹⁸	Down
RP11-106M3.3	1.079926794	3.88x10 ⁻¹⁸	Up
CTA-228A9.3	1.604725390	4.56x10 ⁻¹⁸	Up
RP11-96D1.6	-1.922229332	3.17x10 ⁻¹⁶	Down
DCST1-AS1	1.241184606	1.48x10 ⁻¹⁵	Up
LINC00893	1.211220275	5.90x10 ⁻¹⁵	Up
RP11-365O16.3	-1.828187085	4.13x10 ⁻¹⁴	Down
CTA-984G1.5	1.347562311	4.19x10 ⁻¹⁴	Up
LINC00950	1.227184171	5.67x10 ⁻¹⁴	Up
LINC01389	1.558778315	2.04x10 ⁻¹³	Up
RP11-676J12.7	-2.447116555	8.57x10 ⁻¹³	Down
RP11-1348G14.8	1.017320557	2.62x10 ⁻¹²	Up
AD000684.2	1.15389113	1.22x10 ⁻¹¹	Up
RP11-273G15.2	-1.877345455	5.25x10 ⁻¹¹	Down
RP4-545L17.11	1.052121463	1.04x10 ⁻¹⁰	Up
RP11-666A8.8	1.026612682	1.81x10 ⁻¹⁰	Up
RP11-613D13.8	-1.567841389	3.61x10 ⁻¹⁰	Down
RP11-214O1.3	-1.537151619	1.54x10 ⁻⁹	Down
RP11-256L6.3	-1.692081916	1.58x10 ⁻⁹	Down
LINC01979	-1.709561819	3.31x10 ⁻⁹	Down
bP-21264C1.2	1.063155352	7.78x10 ⁻⁹	Up
FAM99A	-1.893766593	2.04x10 ⁻⁸	Down
LINC00702	1.138170496	2.97x10 ⁻⁸	Up
AC139887.4	1.064667365	6.72x10 ⁻⁸	Up
CERNA1	1.430153282	8.35x10 ⁻⁸	Up
AC011899.9	-1.006302718	5.80x10 ⁻⁷	Down
HULC	1.117705678	1.11x10 ⁻⁶	Up
USP2-AS1	1.009672100	6.87x10 ⁻⁶	Up
PCAT29	1.048101691	2.05x10 ⁻⁴	Up
AJ006998.2	1.090865825	7.06x10 ⁻⁴	Up
RP11-714M23.2	-1.009911294	2.28x10 ⁻⁸	Down

Bold highlight indicates four optimal diagnostic lncRNA biomarkers for HCC. lncRNA, long non-coding RNA; HCC, hepatocellular carcinoma.

Functional annotation. The aforementioned co-expression of 388 DE mRNAs with four optimal DE lncRNA biomarkers

was used to perform the GO and KEGG enrichment analysis. According to GO enrichment analysis, 'mitotic cell cycle'

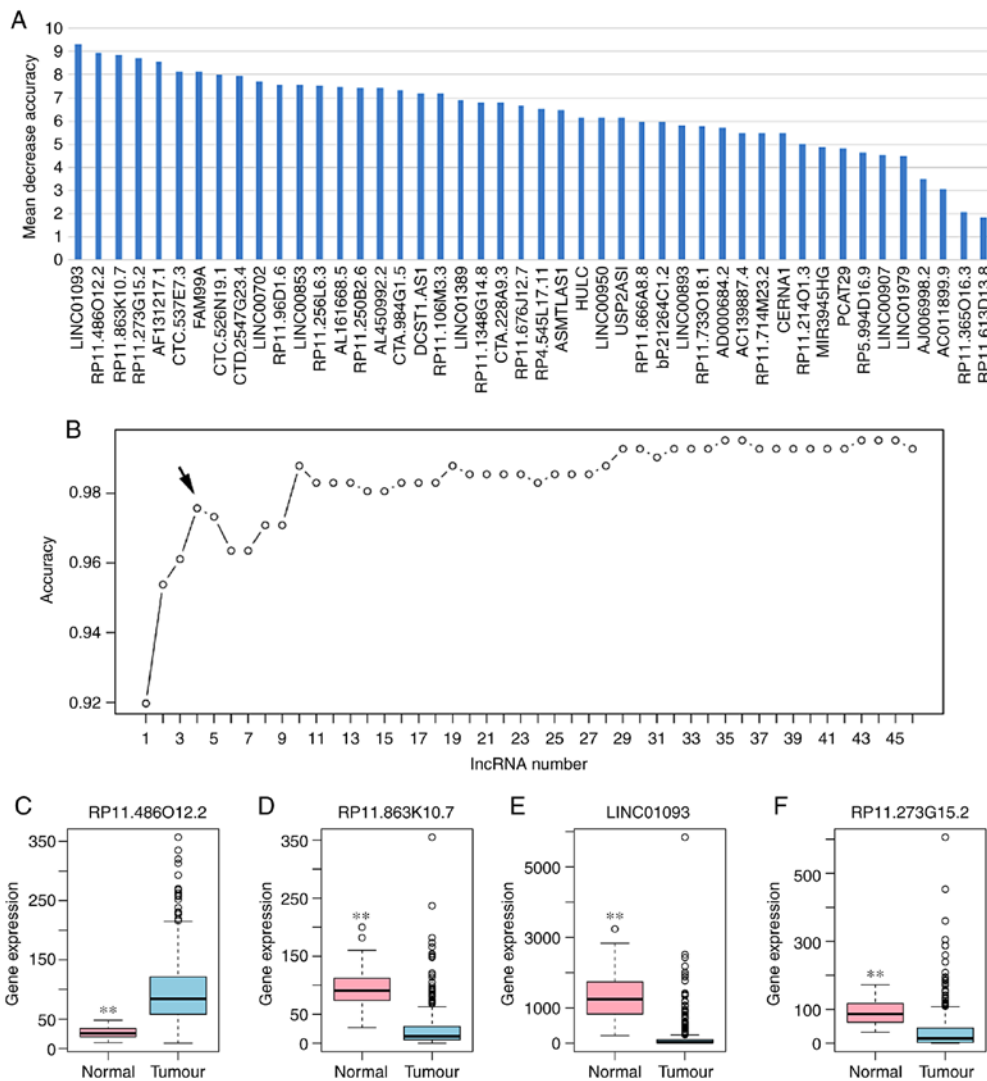


Figure 2. Identification of lncRNA biomarkers for HCC. (A) Importance value of each DElncRNA ranked according to the mean decrease in accuracy by using the random forest analysis. (B) Variance rate of classification performance when increasing numbers of the predictive DElncRNAs. Box-plot displayed the expression levels of (C) RP11-486O12.2, (D) RP11-863K10.7, (E) LINC01093 and (F) RP11-273G15.2 between HCC and normal tissues. HCC, hepatocellular carcinoma; lncRNA, long non-coding RNA.

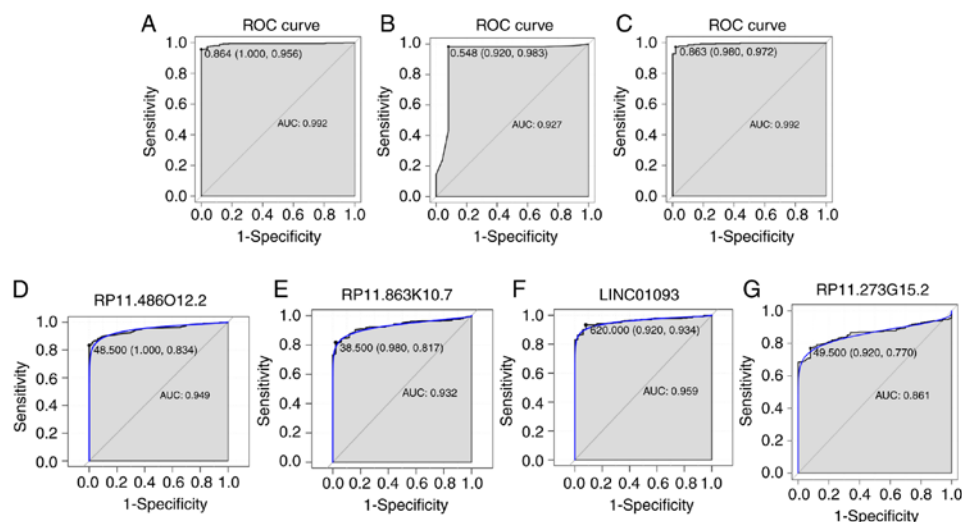
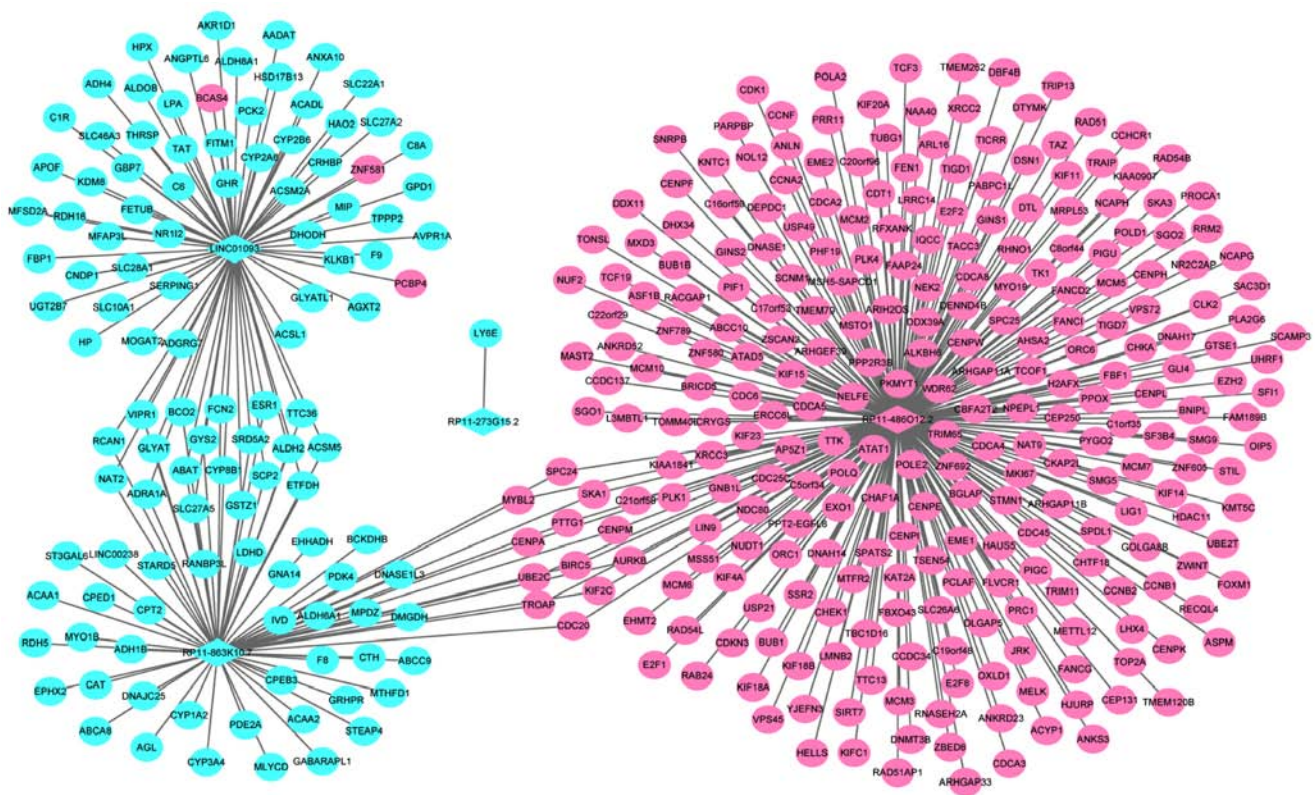


Figure 3. ROC analysis of five hepatocellular carcinoma-specific lncRNA biomarkers. The ROC results of these four diagnostic lncRNA biomarkers (RP11-486O12.2, RP11-863K10.7, LINC01093 and RP11-273G15.2) their combination based on (A) support random forest, (B) decision tree model and (C) support vector machine model, and individual (D) RP11-486O12.2, (E) RP11-863K10.7, (F) LINC01093 and (G) RP11-273G15.2. ROC, receiver operating characteristic; AUC, area under the curve; lncRNA, long non-coding RNA.

lncRNA	r (P-value)				
	Age	Gender	Stage	Grade	Race
LINC0109	0.2181856 (5.793x10 ⁻⁵) ^a	0.01921931 (0.7264)	-0.135419 (0.01325) ^a	-0.3284388 (7.68x10 ⁻¹⁰) ^a	-0.124623 (0.02273) ^a
RP11-486O12.2	0.0327529 (0.5508)	-0.1374603 (0.01191) ^a	0.1646208 (0.002545) ^a	0.1652273 (0.002452) ^a	-0.08335334 (0.1284)
RP11-863K10.7	0.02148461 (0.6956)	0.1633948 (0.002743) ^a	-0.07297289 (0.1834)	-0.244541 (6.148x10 ⁻⁶) ^a	-0.1271082 (0.02014) ^a
RP11-273G15.2	-0.07088287 (0.1963)	0.02640504 (0.6306)	0.06047835 (0.2704)	0.07930304 (0.1481)	0.06394619 (0.2438)



signaling pathway'. A total of nine DEmRNAs, including RDH5, CYP2A6, CYP3A4, RDH16, CYP1A2, CYP2B6, ADH4, UGT2B7 and ADH1B were enriched in the 'retinol metabolism pathway'.

HCC is the most common subtype of liver cancer, with a high worldwide incidence and mortality rate (1). Taking into

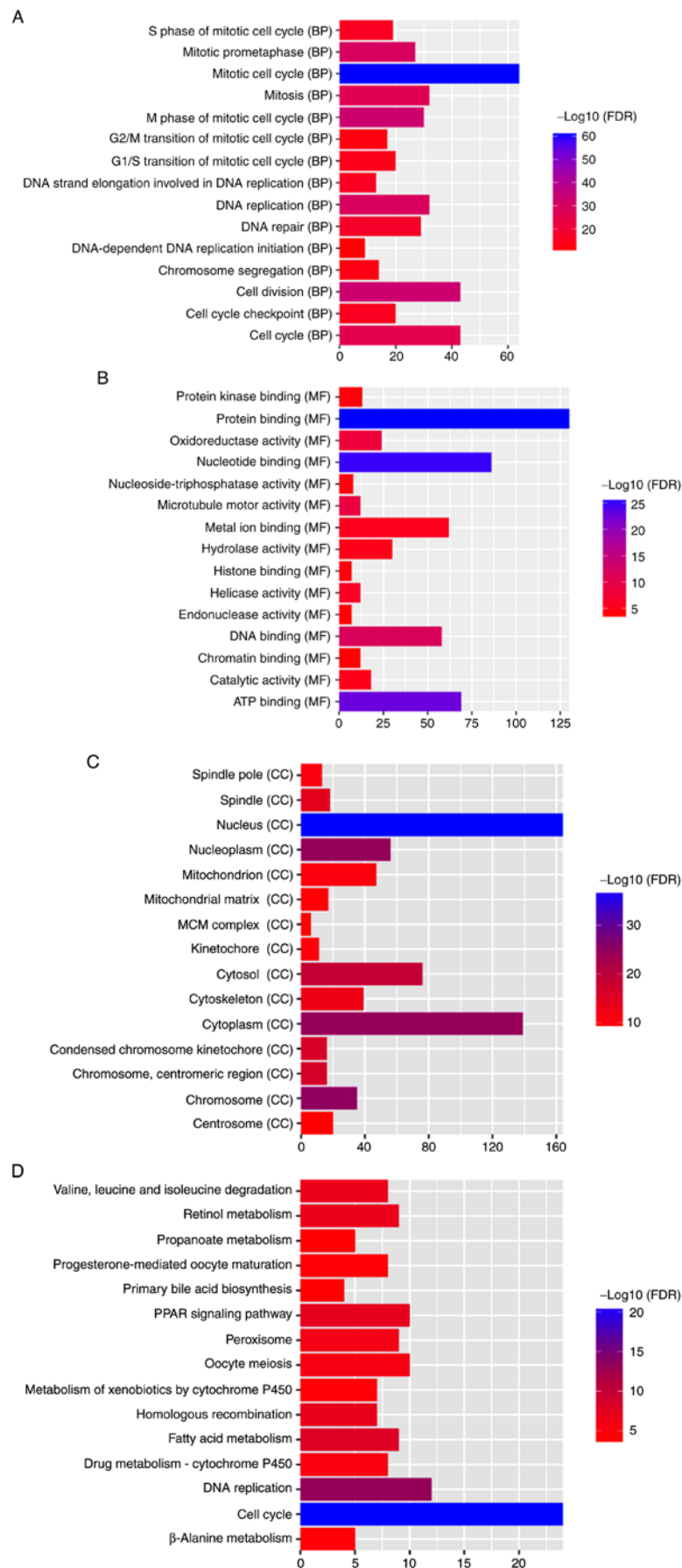


Figure 5. Significant enrichment GO terms and KEGG pathways of DElncRNAs co-expressed with DEmRNAs in hepatocellular carcinoma. The x-axis shows -log FDR and y-axis shows GO terms or KEGG pathways. (A) Biological process. (B) Molecular function. (C) Cellular component. (D) KEGG pathways. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; BP, biological process; FDR, false discovery rate; MF, molecular function; CC, cellular component; MCM, minichromosome maintenance protein complex; PPAR, peroxisome proliferator-activated receptor.

consideration the importance of early diagnosis on survival, further examination for the identification of accurate and specific biomarkers of HCC are required. Increasing evidence suggest that lncRNAs have a vital role in the progress of HCC (7-9). However, studies on lncRNAs as predictive biomarkers and therapeutic targets remain limited.

In the present study, 601 DElncRNAs and 2,920 DEmRNAs were identified between HCC and adjacent tissues, based on the data downloaded from TCGA from 361 patients with HCC. A total of four optimal diagnostic lncRNA biomarkers, including RP11-486O12.2, RP11-863K10.7, LINC01093 and RP11-273G15.2, for HCC were identified by feature selection and classification models.

The ROC results displayed the important diagnostic value of these four lncRNA biomarkers, and their combination, for HCC. The integrated analysis of the four diagnostic lncRNA biomarkers from the GEO database indicated that the lncRNA LINC01093 is downregulated in HCC (23). It has been reported that the downregulation of lncRNA LINC01093 is associated with poor prognosis of patients with HCC (24). Therefore, LINC01093 may be used as a prognostic indicator for HCC. In the present study, not only LINC01093 was indicated to be decreased in HCC tissues compared with adjacent tissues, the diagnostic value of HCC was also revealed with an AUC of 0.959, and a specificity and a sensitivity of 92.0 and 93.4%, respectively.

To the best of our knowledge, with the exception of LINC01093, the present study was the first to identify RP11-486O12.2 as an upregulated DElncRNA, and RP11-863K10.7 and RP11-273G15.2 as downregulated DElncRNAs in HCC. However, their underlying biological function remains unclear. The relationship between these four lncRNAs (LINC01093, RP11-486O12.2, RP11-863K10.7 and RP11-273G15.2) and tumor grade is unclear. The four lncRNAs previously reported to be associated with HCC (GAS5, CARLo-5, PCAT-1 and ZFAS1) are not included in Table II, with the following potentially explanation: i) The data sets used may be different; ii) the criteria for defining differential expression may be different; iii) the algorithm used may be different.

Co-expression analysis of lncRNA-mRNA is the most common approach for identifying potential target genes of lncRNAs and to further investigate the biological function of lncRNAs in a number of diseases. A co-expression network of DElncRNAs-DEmRNAs was constructed using these four specific lncRNA biomarkers of HCC, and the functional annotation of DEmRNAs co-expressed with four lncRNAs was carried out. A total of six DEmRNAs, including SLC27A5, ACSL1, SCP2, SLC27A2, ACADL and PCK2, co-expressed with LINC01093, and four DEmRNAs, including EHHADH, ACAA1, CPT2 and CYP8B1 co-expressed with RP11-863K10.7, were significantly enriched in the PPAR signaling pathway. In addition, three DEmRNAs, including SLC27A5, SCP2 and CYP8B1, enriched in the PPAR pathway were co-expressed with LINC01093 and RP11-863K10.7. PPARs is a transcription factor activated by ligands, and is a member of the nuclear hormone receptor superfamily (25). PPAR signaling pathways are involved in the development and progression of multiple cancers (26-28). Previous studies have confirmed that PPAR signaling pathway serves a crucial

role in regulating HCC tumorigenesis, and has the ability to mediate HCC apoptosis depending on the regulation of the PI3K pathway (29,30). ACSL1 has been reported to serve as a classic target gene of PPAR α (31,32). lncRNAs HULC activates ACSL1 by upregulating the transcription factor PPAR α in HCC (33). Therefore, the present study suggests that LINC01093 and RP11-863K10.7 may be involved in the occurrence of HCC by regulating the PPAR signaling pathway.

A total of five DEmRNAs, including CYP2A6, RDH16, CYP2B6, ADH4 and UGT2B7, co-expressed with LINC01093 and four DEmRNAs, including RDH5, CYP3A4, CYP1A2 and ADH1B co-expressed with RP11-863K10.7 were significantly enriched in the 'retinol metabolism' signaling pathway. Cytochrome P450 proteins (CYPs) are a family enzymes localized in the endoplasmic reticulum and the mitochondrial membrane with important roles in metabolism (34). The most significant drug metabolizing CYPs in humans, include CYP2A6, CYP3A4 and CYP1A2, which are responsible for the metabolism of >90% of medicines and for the metabolic activation of procarcinogens (35,36). CYP2A6 is an enzyme that has a crucial role in the metabolic activation of a number of procarcinogens, including 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, 1,3-butadiene, aflatoxin B1 and N-nitrosodiethylamine (37,38). The upregulated expression of CYP2A6 is associated with an increased risk of liver cancer in individuals with frequent aflatoxin exposure (39). CYP2A6 is also a specific biomarker for colorectal cancer (40). CYP2A6 gene deletion has been reported to be significantly lower in patients with lung cancer compared with healthy individuals (41). It was therefore hypothesized in the present study that LINC01093 and RP11-863K10.7 may serve an important role in HCC by regulating retinol metabolism. Cui *et al* (42) is a more comprehensive study, and their study predicted lncRNAs associated with the prognosis and metastasis of patients with HCC, and the lncRNAs were validated using cell lines *in vitro*. HCC samples are currently being collected to validate the expression of the identified key lncRNAs in subsequent research with larger sample size. Then, the biological significances of key lncRNAs will be investigated in model systems or cell lines.

In conclusion, the present study has indicated that the lncRNA expression profiles are altered in HCC compared with normal samples. A total of four DElncRNAs were identified as potential biomarkers of HCC. Functional annotation of DEmRNAs co-expression with the four lncRNA biomarkers of HCC provided novel insight for examining the pathogenesis of HCC. There are limitations of the current study. These four lncRNAs (LINC01093, RP11-486O12.2, RP11-863K10.7 and RP11-273G15.2) were not associated with survival (data not shown). Moreover, the expression of specific lncRNAs in HCC induced by different factors, including HBV, was not investigated. Finally, identifying diagnostic biomarkers for HCC is currently in a pilot study, and further experiments are required to determine the biological significance of important lncRNAs in HCC.

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

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

Authors' contributions

HSu and GL designed the project. HSh, XW, BW, QQ and HG analyzed and interpreted the data. HSh and GL were major contributors in drafting the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- Ma H, Yuan L, Li W, Xu K and Yang L: The lncRNA H19/miR-193a-3p axis modifies the radio-resistance and chemotherapeutic tolerance of hepatocellular carcinoma cells by targeting PSEN1. *J Cell Biochem* 119: 8325-8335, 2018.
- Merion RM: Current status and future of liver transplantation. *Semin Liver Dis* 30: 411-421, 2010.
- Chen G, Li X, He G, Yu Z, Luo J, He J and Huang Z: Low expression of GNAI3 predicts poor prognosis in patients with HCC. *Int J Clin Exp Med* 8: 21482-21486, 2015.
- Pelus LM and Fukuda S: Peripheral blood stem cell mobilization: The CXCR2 ligand GROβ rapidly mobilizes hematopoietic stem cells with enhanced engraftment properties. *Exp Hematol* 34: 1010-1020, 2006.
- Song X, Wang Z, Jin Y, Wang Y and Duan W: Loss of miR-532-5p in vitro promotes cell proliferation and metastasis by influencing CXCL2 expression in HCC. *Am J Transl Res* 7: 2254-2261, 2015.
- Yang X, Xie X, Xiao YF, Xie R, Hu CJ, Tang B, Li BS and Yang SM: The emergence of long non-coding RNAs in the tumorigenesis of hepatocellular carcinoma. *Cancer Lett* 360: 119-124, 2015.
- Qu Z, Yuan CH, Yin CQ, Guan Q, Chen H and Wang FB: Meta-analysis of the prognostic value of abnormally expressed lncRNAs in hepatocellular carcinoma. *Onco Targets Ther* 9: 5143-5152, 2016.
- Sun J, Wei X and Xu L: Upregulation of lncRNA Sox2ot indicates a poor prognosis for patients with hepatocellular carcinoma and promotes cell invasion. *Oncol Lett* 16: 1189-1195, 2018.
- Chang L, Li C, Lan T, Wu L, Yuan Y, Liu Q and Liu Z: Decreased expression of long non-coding RNA GAS5 indicates a poor prognosis and promotes cell proliferation and invasion in hepatocellular carcinoma by regulating vimentin. *Mol Med Rep* 13: 1541-1550, 2016.
- Wang F, Xie C, Zhao W, Deng Z, Yang H and Fang Q: Long non-coding RNA CARLo-5 expression is associated with disease progression and predicts outcome in hepatocellular carcinoma patients. *Clin Exp Med* 17: 33-43, 2017.
- Zhao J, Greene CM, Gray SG and Lawless MW: Long noncoding RNAs in liver cancer: What we know in, 2014. *Expert Opin Ther Targets* 18: 1207-1218, 2014.
- Yu W, Qiao Y, Tang X, Ma L, Wang Y, Zhang X, Weng W, Pan Q, Yu Y, Sun F and Wang J: Tumor suppressor long non-coding RNA, MT1DP is negatively regulated by YAP and Runx2 to inhibit FoxA1 in liver cancer cells. *Cell Signal* 26: 2961-2968, 2014.
- Yan TH, Yang H, Jiang JH, Lu SW, Peng CX, Que HX, Lu WL and Mao JF: Prognostic significance of long non-coding RNA PCAT-1 expression in human hepatocellular carcinoma. *Int J Clin Exp Pathol* 8: 4126-4131, 2015.
- Li T, Xie J, Shen C, Cheng D, Shi Y, Wu Z, Deng X, Chen H, Shen B, Peng C, *et al*: Amplification of long noncoding RNA ZFAS1 promotes metastasis in hepatocellular carcinoma. *Cancer Res* 75: 3181-3191, 2015.
- Sonohara F, Inokawa Y, Hayashi M, Yamada S, Sugimoto H, Fujii T, Kadera Y and Nomoto S: Prognostic value of long non-coding RNA HULC and MALAT1 following the curative resection of hepatocellular carcinoma. *Sci Rep* 7: 16142, 2017.
- Love MI, Huber W and Anders S: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550, 2014.
- Benjamini Y and Hochberg Y: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc: Ser B (Methodol)* 57: 289-300, 1995.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al*: Gene ontology: Tool for the unification of biology. The gene ontology consortium. *Nat Genet* 25: 25-29, 2000.
- The Gene Ontology Consortium: The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Res* 47D: D330-D338, 2019.
- Kanehisa M, Sato Y, Furumichi M, Morishima K and Tanabe M: New approach for understanding genome variations in KEGG. *Nucleic Acids Res* 47D: D590-D595, 2019.
- Kanehisa M, Furumichi M, Tanabe M, Sato Y and Morishima K: KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 45D: D353-D361, 2017.
- Kanehisa M and Goto S: KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28: 27-30, 2000.
- Jin B, Wang W, Du G, Huang GZ, Han LT, Tang ZY, Fan DG, Li J and Zhang SZ: Identifying hub genes and dysregulated pathways in hepatocellular carcinoma. *Eur Rev Med Pharmacol Sci* 19: 592-601, 2015.
- Dai M, Chen S, Wei X, Zhu X, Lan F, Dai S and Qin X: Diagnosis, prognosis and bioinformatics analysis of lncRNAs in hepatocellular carcinoma. *Oncotarget* 8: 95799-95809, 2017.
- Antoniosante A, d'Angelo M, Castelli V, Catanesi M, Iannotta D, Giordano A, Ippoliti R, Benedetti E and Cimini A: The involvement of PPARs in the peculiar energetic metabolism of tumor cells. *Int J Mol Sci* 19: pii: E1907, 2018.
- Ricci M, Miola M, Multari C, Borroni E, Canuto RA, Congiusta N, Vernè E, Follenzi A and Muzio G: PPARs are mediators of anti-cancer properties of superparamagnetic iron oxide nanoparticles (SPIONs) functionalized with conjugated linoleic acid. *Chem Biol Interact* 292: 9-14, 2018.
- Sanchez DJ, Steger DJ, Skuli N, Bansal A and Simon MC: PPARγ is dispensable for clear cell renal cell carcinoma progression. *Mol Metab* 14: 139-149, 2018.
- Li Q, Peng YS, Chen PJ, Wang ML, Cao C, Xiong H, Zhang J, Chen MH, Peng XB and Zeng K: Peroxisome proliferator-activated receptor-γ agonist-mediated inhibition of cell growth is independent of apoptosis in human epidermoid carcinoma A431 cells. *Oncol Lett* 15: 6578-6584, 2018.
- Xiao YB, Cai SH, Liu LL, Yang X and Yun JP: Decreased expression of peroxisome proliferator-activated receptor alpha indicates unfavorable outcomes in hepatocellular carcinoma. *Cancer Manag Res* 10: 1781-1789, 2018.
- Liu Z, Wang Y, Dou C, Sun L, Li Q, Wang L, Xu Q, Yang W, Liu Q and Tu K: MicroRNA-1468 promotes tumor progression by activating PPAR-γ-mediated AKT signaling in human hepatocellular carcinoma. *J Exp Clin Cancer Res* 37: 49, 2018.
- Phillips CM, Goumidi L, Bertrais S, Field MR, Cupples LA, Ordoas JM, Defoort C, Lovegrove JA, Drevon CA, Gibney MJ, *et al*: Gene-nutrient interactions with dietary fat modulate the association between genetic variation of the ACSL1 gene and metabolic syndrome. *J Lipid Res* 51: 1793-1800, 2010.
- Ong KT, Mashek MT, Bu SY, Greenberg AS and Mashek DG: Adipose triglyceride lipase is a major hepatic lipase that regulates triacylglycerol turnover and fatty acid signaling and partitioning. *Hepatology* 53: 116-126, 2011.

33. Cui M, Xiao Z, Wang Y, Zheng M, Song T, Cai X, Sun B, Ye L and Zhang X: Long noncoding RNA HULC modulates abnormal lipid metabolism in hepatoma cells through an miR-9-mediated RXRA signaling pathway. *Cancer Res* 75: 846-857, 2015.
34. Korobkova EA: Effect of natural polyphenols on CYP metabolism: Implications for diseases. *Chem Res Toxicol* 28: 1359-1390, 2015.
35. Zanger UM and Schwab M: Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol Ther* 138: 103-141, 2013.
36. Zhou J, Wen Q, Li SF, Zhang YF, Gao N, Tian X, Fang Y, Gao J, Cui MZ, He XP, *et al*: Significant change of cytochrome P450s activities in patients with hepatocellular carcinoma. *Oncotarget* 7: 50612-50623, 2016.
37. Gonzalez FJ and Gelboin HV: Role of human cytochrome P-450s in risk assessment and susceptibility to environmentally based disease. *J Toxicol Environ Health* 40: 289-308, 1993.
38. Messina ES, Tyndale RF and Sellers EM: A major role for CYP2A6 in nicotine C-oxidation by human liver microsomes. *J Pharmacol Exp Ther* 282: 1608-1614, 1997.
39. Gullstén H, Agúndez JA, Benítez J, Läärä E, Ladero JM, Díaz-Rubio M, Fernandez-Salguero P, Gonzalez F, Rautio A, Pelkonen O and Raunio H: CYP2A6 gene polymorphism and risk of liver cancer and cirrhosis. *Pharmacogenetics* 7: 247-250, 1997.
40. Nowell S, Sweeney C, Hammons G, Kadlubar FF and Lang NP: CYP2A6 activity determined by caffeine phenotyping: Association with colorectal cancer risk. *Cancer Epidemiol Biomarkers Prev* 11: 377-383, 2002.
41. Kamataki T, Nunoya K, Sakai Y, Kushida H and Fujita K: Genetic polymorphism of CYP2A6 in relation to cancer. *Mutat Res* 428: 125-130, 1999.
42. Cui H, Zhang Y, Zhang Q, Chen W, Zhao H and Liang J: A comprehensive genome-wide analysis of long noncoding RNA expression profile in hepatocellular carcinoma. *Cancer Med* 6: 2932-2941, 2017.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.