RNA-Seq profiling revealed PBMC RNA as potential biomarker for hepatocellular carcinoma

Zhiyi Han
Shenzhen Traditional Chinese Medicine Hospital
Wenxing Feng
Shenzhen Traditional Chinese Medicine Hospital
Rui Hu
Shenzhen Traditional Chinese Medicine Hospital
Qinyu Ge
Southeast University
Wenfeng Ma
Shenzhen Traditional Chinese Medicine Hospital
Wei Zhang
Shenzhen Traditional Chinese Medicine Hospital
Shaomin Xu
Shenzhen Traditional Chinese Medicine Hospital
Bolin Zhan
Shenzhen Traditional Chinese Medicine Hospital
Lai Zhang
Shenzhen Traditional Chinese Medicine Hospital
Xinfeng Sun (714310278@qq.com)
Shenzhen Traditional Chinese Medicine Hospital
Xiaozhou Zhou
Shenzhen Traditional Chinese Medicine Hospital https://orcid.org/0000-0002-4758-0060

Primary research

Keywords: Hepatocellular carcinoma, Peripheral Blood Mononuclear Cells, RNA sequencing, Cancer biomarker, Differential expressed genes

Posted Date: August 12th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-55341/v1

Abstract

Background

Hepatocellular carcinoma is one of the most common malignancies with extremely high incidence and mortality rates. Although there have been many studies focus on biomarkers study, few have been reported on PBMC RNA profiles of hepatocellular carcinoma.

Methods

In this study, we attempted to profile the expression of Peripheral Blood Mononuclear Cells (PBMCs) RNA by using RNA-seq technology and compared the transcriptome between hepatocellular carcinoma patients and the healthy controls. 17 patients and 17 healthy controls involved in this study, PBMCs RNA were sequenced. The sequencing data were analyzed with bioinformatics tools and qRT-PCR was used for selected differential expressed gene validation.

Results

It is showed that 1578 dysregulated genes found including 1334 upregulated genes and 244 downregulated genes. GO enrichment and KEGG analysis denoted most of the differential expressed genes (DEGs) involved in immune response are closely related to hepatocellular carcinoma. Expression of the 6 selected genes (DEGs, SELENBP1, SLC4A1, SLC26A8, HSPA8P4, CALM1, and RPL7p24) were confirmed by qRT-PCR, and higher sensitivity and specificity obtained by ROC analysis of the 6 genes. CALM1 was found gradually decreasing along with the tumor enlarged.

Conclusions

It is suggested potential biomarker for diagnosis, classification and therapeutic target of hepatocellular carcinomas. This study provided new visions into development of liver cancer and potential efficient clinical diagnosis in the future.

1. Introduction

Hepatocellular carcinoma (HCC) is one of the leading causes of cancer mortality worldwide [El-Serag and Rudolph, 2007]. Although there are many recent advances in cancer diagnosis and treatment with respect to surgery, radiotherapy, chemotherapy and biotherapy, majority of it remains incurable once it has become metastatic and has a very poor prognosis, primarily due to diagnostic delays or omissions [Attwa and El-Etreby, 2015; Befeler and Di Bisceglie, 2002; Budhu et al., 2006]. The majority of liver cancer occurs in patients with underlying liver disease, such as hepatitis B virus (HBV) infection and cirrhosis [Hu et al., 2019]. Over half of patients with HCC are diagnosed at advanced stages, preventing the possibility of

curative therapies. Imaging diagnosis, such as positron emission tomography (PET), is a highly specific tool in hepatocellular carcinoma diagnosis, but in small metastasis or micro-metastasis, typical imaging characteristics are lacking. Alpha-fetoprotein (AFP) and alkaline phosphatase (ALP or AKP), are widely used, yet imperfect, biomarker for hepatocellular carcinoma early diagnosis in clinical practice. It has been reported that AFP (at a threshold level of 20 ng/mL) showed low sensitivity of 40–60% with specificity of 80–90% [Edoo et al., 2019; Forner and Bruix, 2012; Zhu et al., 2013]. Low sensitivity, false negativity (e.g., a small HCC with normal AFP level), and false positivity (e.g., liver function damage and certain gastrointestinal tumors) of AFP could lead to decreased chance of early diagnosis and thus poor clinical outcomes, highlighting the requirement for more effective approaches for HCC detection.

Here, we investigated peripheral blood mononuclear cells (PBMCs) transcriptomes in HCC and evaluated the diagnostic value of PBMCs transcripts. PBMCs, an easily accessible and minimally invasive sample, could be isolated from patients with hepatocellular carcinoma and healthy control in this present study [Varela-Martinez et al., 2018]. With simple components, it would be beneficial for increasing the reliability of the results through the decline of intra-tumor heterogeneity. Initial screening of dys-regulated mRNAs was conducted using RNA-seq [Shen et al., 2019]. Then, PCA were performed to group the samples according to their similarity on the gene expression, which could decrease the inter-tumor heterogeneity through reducing dimension of effectors. Finally, the dysregulated pathways or biological processes were selected and novel potential mRNAs were confirmed by qRT-PCR. The design of the study could be seen in **Figure.1**. This work may help to understand the progression of tumor development and potential efficient clinical diagnosis in the future.

Transcriptomes of PBMCs from HCC and controls were profiled by RNA-Seq and then analyzed by bioinformatics methods. Subsequently, the proposed (screened) genes were validated by qRT-PCR in the validation cohort with 50 hepatocellular carcinoma patients and 50 healthy controls.

2. Material And Methods

2.1. Separation of PBMCs and RNA extraction

We collected peripheral blood (about 2ml) from 34 individuals (17 samples from HC patients and the 17 from the healthy control people who have similar gender and age with these cancer patients, additional 33 patient samples were also collected for validation by qRT-PCR, the clinical characteristics of samples collected shown in **Figure.2**), the PBMCs was isolated from 2 ml fresh EDTA-blood of each human subject by Ficoll-Paque^{TM} PREMIUM according to its commercial protocols. And then the PBMCs samples were used for RNA extraction using Trizol reagent (Invitrogen, Carlsbad, CA) following standard procedures as previously described [Marioni et al., 2008]. RNA quality was accessed by the absorbance at 260 nm (A260) and 280 nm (A280) using NanoDrop ND-1000 (Thermo Fisher Scientific, Waltham, MA), and RNA integrity was determined by RNA integrity number (RIN; Agilent 2100 RIN Beta Version Software).

Ethics approval was obtained from the Ethics Committee of Shenzhen Traditional Chinese Medicine Hospital. All experiments were performed in accordance with relevant guidelines and regulations set out by the ethical committee.

2.2 RNA-seq

34 RNA samples were submitted for sequencing. Among them, 17 were extracted from the hepatocellular carcinoma patients, and 17 were extracted from the matched healthy people. Based on the previous studies, the sequencing library was prepared after removal of rRNA, according to the Illumina® TruSeq® RNA Sample Preparation Guide (Illumina, San Diego, CA, USA). Ligation of the indexed adaptor was carried out following the double strands cDNA synthesis. After size selection by using Agencourt AMPure XP (Beckman), The libraries were quantified and qualitied using Qubit® 2.0 Fluorometer with the Qubit® dsDNA HS assay kit (Invitrogen™, Eugene, OR, USA) and Agilent bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), respectively, and it was submitted to Illumina HiSeq X-10 (Illumina, San Diego, CA, USA) for sequencing with 2×150 bp pair-end technology.

2.3 Bio-informatic Analysis:

The raw reads were filtered by SOAPnuke (version 1.0.1), and then were mapped to the human (hg19) genomes provided by Illumina iGenomes (downloaded from cufflinks.cbcb.umd.edu/igenomes.html) with Tophat2 (version 2.0.7) calling Bowtie2 (version 2.1.0) using the default settings. The alignment and differentially expression genes analysis were performed with Cufflinks (version 2.0.2) [Trapnell et al., 2012]. The DEGs with p-value less than 0.05 and a fold change greater than 2-fold were regarded as significantly altered.

For functional analyses, GO analysis was carried out with the PANTHER (protein annotation through evolutionary relationship) classification system (http://www.pantherdb.org/) [Mi et al., 2019]. The statistical overrepresentation test was performed. It was based on the Mann-Whitney test and used to determine whether any ontology class or pathway had numeric values that were non-randomly distributed with respect to the entire list of values.

After that Weighted Gene Correlation Network Analysis (WGCNA) was used to analysis co-expression modules of DEGs in HCC [Wan et al., 2018; Zhang et al., 2017], and explored the correlation between different modules and clinical traits, such as age, gender, tumor-sizes, and classification. Finally, in order to obtain insights into the interaction between DEGs, the protein–protein interaction (PPI) network of DEGs was visualized by Cytoscape software (version 3.7.2).

2.4 qRT-PCR

99 RNA samples were submitted for qRT-PCR validation. Among them, 50 were extracted from the hepatocellular carcinoma patients, and 49 were extracted from the matched healthy people. For reverse transcription reactions, 500 ng total RNA was used with a PrimeScript™ RT Master Mix (Perfect Real

Time) (Takara Bio, Inc.) at 37˚C for 15 min and 85 ˚C for 5 second with a final volume of 20 μL. The following qPCR was performed using SYBR® Premix Ex Taq™ II (Perfect Real Time; Takara, Bio., Inc.) on an Applied Biosystems 7500 real-time PCR machine (Life Technologies) by using 2 μL of the cDNA obtained in the RT reaction. The primers were synthesized by Sangon Biotech, Inc. and shown in Table 1. The PCR reaction was performed at 95˚C for 3 min, followed by 40 cycles at 95˚C for 5 sec and 60˚C for 40 sec. Each reaction was repeated three times, and the mean value of Ct for each triplicate was calculated. The ΔCt target cDNA was the difference between Ct target gene and Ct of reference gene (GAPDH). And the ΔΔCt value between ΔCt of HCC and ΔCt of NC was used to calculate the amplification fold change in gene expression (ΔΔCt = ΔCt HC -ΔCt NC; ΔCt = Ct target – Ct reference). The quality of the amplification products was accessed by 2% agarose gel and dissociation curve.

2.5 Statistical analysis:

All of data analysis was performed with R (version 4.0.1) and SPSS (version 22.0). PCA, Pearson correlation analysis was performed to analyze the similarity of the samples, and Spearman correlation analysis was performed to analyze the correlations. Statistical differences were examined by a t-test or a one-way ANOVA. All statistical tests were performed as two-sided tests. And p values<0.05 were considered statistically significant.

3. Results

3.1 Baseline characteristics and RNA-seq information of samples

We used paired-end RNA-Seq to present the gene expression profiles of 17 PBMCs samples of the patient with hepatocellular carcinoma and 17 PBMCs samples of age-matched healthy people (as the control). RNA-Seq generated from 58,012,158 to 83,083,036 raw reads that were aligned to the human reference hg19, representing 22,894,689 to 42,821,652 reads mapped to exon (37.879%-57.238%). The baseline characteristics of the 17 patients with hepatocellular carcinoma were shown in **Figure.2**, including age, sex, BCLC classification and grade, tumor size and whether treatment when sampling, etc.

3.2. Characterizing the gene expression profiles of HCC

To gain insights into the characteristics of gene expression profiles of the liver cancer, PCA and the pearson's correlation coefficient were analyzed to generate an evaluation of the similarities or dissimilarities of the RNA-seq outputs. PCA is a linear projection method that allows visualization of high-dimensional data in a lower dimensional space. The results showed that the first principal component (PC1) accounted for 26% of the overall variance of the data, and the second principal component (PC2) accounted for 9%. As shown in Figure 3A, at PC1, the control samples and the hepatocellular carcinoma samples were gathered respectively, and each of them had the parallel contribution, it is demonstrating that they were similar to each other, the PC1 reflected general characters of gene expression profiles of the hepatocellular carcinoma; at PC2, the control samples were gathered into a group, the cancer samples

were also gathered, which indicating that there were differences in gene expression profiles between the two groups.

The pearson's correlation coefficient of significantly dysregulated mRNA expression showed in the heatmap of inter-sample correlation. It can be seen from the Figure 3B that there was an obvious difference of transcript expression levels between the HCC and control groups, and two groups found in the hepatocellular carcinoma, but the difference was slight within each group.

In addition, we classified of all differentially expressed genes (DEGs) base on the different location on the chromosomes (Figure.3C). The expression of the related genes could be found in all chromosomes, while there are no differentially expressed genes were found in chromosome 8, chromosome 16, chromosome 21 and Y chromosome. It could be seen from Figure 3D, the types of these DEGs mostly comes from protein coding region, and then lincRNA, antisense and proceeded pseudogene, etc. It is worth concerning that little DEGs were from miRNA, MT_tRNA and snoRNAs, it is suggested these molecules might have potential functions in the occurrence and development of hepatocellular carcinoma.

3.3. Different expressed genes in HCC

Total 60006 expressed genes were detected by RNA-seq of these 34 PBMC samples from the RNA-seq data. And we found that 1578 genes were differentially regulated between HC patients and normal controls by fold change ≥ 2.0 and $p < 0.05$ according to the volcano plot and scatter plot filtering (Figure 4A, 4B). And most of the DEGs are protein-coding genes. Among them, 1334 genes were upregulated, and 244 genes were downregulated in PBMCs of hepatic carcinoma patients (Figure.4C), the top 25 dysregulated expressed genes were shown in supplemental Table S1. Hierarchical clustering of these dysregulated genes indicated that the gene expression profiles between hepatic carcinoma and control samples were distinguishable (Figure 4D). These results suggested that the expression of PBMCs RNA in hepatocellular carcinoma is different from the healthy control.

3.4 GO and KEGG pathway analysis of differentially expressed genes

GO enrichment and KEGG pathway analysis were used to explore the functions of the DEGs. From Figure.5A and 5B we found the significant enriched GO terms in biologic process were neutrophil degranulation, neutrophil activation, neutrophil mediated immunity, neutrophil activation involved in immune response, granulocyte activation, T cell activation, hemostasis, blood coagulation, coagulation, leukocyte differentiation, and platelet degranulation; The significant enriched GO terms in cell component were secretory granule lumen, specific granule, cytoplasmic vesicle lumen, vesicle lumen, tertiary granule, secretory granule membrane, primary lysosome, azurophil granule, and azurophil granule lumen, and in molecular function was non-membrane spanning protein tyrosine kinase activity, actin binding, actin filament binding, SH2 domain binding, phosphatidylinositol binding, Toll-like receptor binding, phospholipid binding, GTPase activator activity, ad GTPase regulator activity. Among them, the significant enriched GO terms of the upregulated genes mainly involved in immune response, secretory granule lumen, Toll-like receptor binding and regulation of body fluid levels which has been validated

relating to cancers. The downregulated gene function enrichment involved in ribonucleoprotein complex biogenesis, T cell differentiation, T cell receptor complex, acetyltransferase activity are also closely related to hepatic carcinoma.

Based on KEGG annotation, pathway analysis showed that top 20 signaling pathways, including Lysosome, Endocytosis, Phagosome, Malaria, Platelet activation, FoxO signaling pathway, NOD-like receptor signaling pathway, T cell receptor signaling pathway, Hepatitis B, Apoptosis and so on, which closely associated with hepatic carcinoma (Figure.5C, D, E).

3.5 Co-expression modules analysis of HCC

The information contained in PBMCs RNA is still complex. In order to link more effectively information with HCC, DEGs from the 17 HC patients were used to construct the co-expression modules by WGCNA. Here, we defined the number of genes in each module at least to be 10 and the depth of cut is 0.8. As a result, 14 gene modules in hepatocellular carcinoma were identified, which were shown in different colors (Figure 6A), genes not assigned to any modules are returned to the grey module. The different number of DEGs in the 14 modules was shown in Figure 6B. The clinical traits including the age of patient, gender, tumor-size and BCLC classification were collected, and the correlation between co-expression module and clinic traits were identified. As shown in Figure 6C, we found that there were no co-expression modules have significant correlation with ascites, it is indicated that the differently expressed genes in hepatic carcinoma are not affected by ascites. From the figure, we can see that these clinical traits had little influence on DEGs of HCC. However, we found that the gene module with salmon related with several clinical features such as whether treated when sampling, the treatment method, whether cirrhosis and the classification of BCLC; Green module gene sets correlation with cirrhosis, treatment or not, treatment method and BCLC classification, and it was worth further analysis.

With WGCNA analysis, we discovered that genes in blue, turquoise and brown module occupy the dominant of all the differential expressed gene sets, indicated these genes, such as CXCR1, CXCR2, SELENBP1, SLC4A1, HSPA8P4, CALM1 and CAPN2 et.al, played the important roles in the generation, development and molecular regulation of HCC. It’s worth mentioning that the salmon module were showed correlated with several clinical features through the module-trait analysis, 13 genes contained in the this module including AHNAK, CALM1, CAPN2, EEF1A1, HNRNPA1, PPIA, 3 genes from RPS family (RPS13, RPS14, RPS26) and 4 genes related with RPL including RPL15, RPL29, RPL7p24 and RPLP0. Gene functional enrichment analysis result showed that most of the genes involved in neutrophil mediated immunity, leukocyte mediated immunity, leukocyte activation, immune system process and cell migration, adhesion and motility. The Protein–protein interaction (PPI) network of DEGs was constructed by the top 40 DEGs (ranked by fold change). It can be seen in Figure 7 that most of the genes interact with others, and only 5 genes have no interactions with others.

3.6 Validation with quantitative RT-PCR

To confirm the sequence results, we used qRT-PCR to validate the expression of DEGs in PBMC of HCC patients (n=50) and healthy controls (n=49). 6 genes were detected in validation cohort. These genes are SELENBP1, SLC4A1, SLC26A8, HSPA8P4, CALM1, and RPL7p24 which selected from the top dysregulated genes between HC and controls, and the trait related gene sets from WGCNA analysis result.

The sequencing results showed that all the 6 genes except CALM1 (*p*-value=0.063) were significantly differentially expressed in PBMCs of patients with HCC when comparing with control, as shown in Figure 8A, the qRT-PCR results showed similarly changes, and significant differences found between group HC and normal control in all the 6 genes.

Then the ROC curve analysis was performed to assess the diagnostic value of these DEGs in HC patients. It was showed in Figure 8B, The AUC was 0.977 (95% CI: 0.956–0.998 P < 0.001) for HSPA8P4, 0.975 (95% CI: 0.952–0.998 P < 0.001) for RPL7p24, 0.853 (95% CI: 0.776–0.931 P < 0.001) for SELENBP1, 0.850 (95% CI: 0.770–0.930 P < 0.001) for SLC26A8, 0.868 (95% CI: 0.794–0.942 P < 0.001) for SLC4A1, and 0.770 (95% CI: 0.676–0.863, P < 0.01) for CALM1. The results suggested the potential diagnostic value for these genes in HCC patients.

It should be noted that the expression of CALM1 detected by qRT-PCR was decreased significantly when comparing with control although no significant differences found in sequencing result. More interestingly, the expression level was gradually decreased with the increase of the tumor size (*p* value <0.05) (Figure 8C). It is suggested that CALM1 might closely related with the development of the tumor.

4. Discussion

In this study, PBMCs were separated for RNA-Seq to profile the different gene expression in HCC patient. PBMCs mainly contain lymphocyte and monocyte separated from blood sample which widely used in clinical diagnosis and academic researches. It had been confirmed that closely related with the occurrence, development, metastasis and prognosis of tumor because of the abnormal immune function. The detection of differential gene expression of PBMCs might supervise the development of tumor. Furthermore, PBMC is easy to access and low complexity, so it would be an ideal material for gene expression studying and tumor monitoring.

We utilized RNA-seq and identified 1334 genes were upregulated, and 244 genes were downregulated in PBMCs of HCC with fold changes ≥ 2.0 and p-value < 0.05. These identified DEGs were probably derived from circulating tumor cells, which shed into the vasculature from a primary cancer and circulated through the bloodstream [Maheswaran and Haber, 2010]. Previous studies demonstrated that CTCs contained a variety of mRNAs that play vital roles for subsequent growth of additional tumors in vital distant organs, triggering a mechanism for the vast majority of cancer-related deaths [Gazzaniga et al., 2015; Raimondi et al., 2011; Riethdorf et al., 2010], and their levels gradually increasing with the tumor stage. It indicated that the DEGs were closely related with the tumor size, classification et al, which could be used to supervise the progression of hepatocellular carcinoma [Cools-Lartigue et al., 2013; Hou et al., 2012].

After the GO, KEGG and co-expression analysis, six DEGs were selected for qPCR validation, 4 genes (SELENBP1, SLC4A1, SLC26A8 and HSPA8P4) from the top dysregulated genes and showed dominating the gene regulation network and 2 genes (CALM1 and RPL7p24) from WGCNA modules. CALM1 and RPL7p24 were selected from the salmon modules (Figure. 6C), which showed correlating with treatment method, cirrhosis and BCLC classification. Among them, SELENBP1, SLC4A1, SLC26A8 and RPL7p24 were showed upregulated, HSPA8P4 and CALM1 were showed downregulated from sequencing results. All the genes are involved in the regulatory pathway including immune response, granulocyte activation, T cell activation, Toll-like receptor binding, and GTPase regulator activity et al, which had been shown closely related to hepatocellular carcinoma [Greenhill, 2018a; Greenhill, 2018b; Yu et al., 2020].

All the six gene expression in hepatocellular carcinoma samples were confirmed by qRT-PCR. We found that no previous studies showed the expression of SLC4A1 and RPL7p24 were related with hepatocellular carcinoma, although the mutations had been reported with many diseases [2020; Wadhwa et al., 2015; Yi et al., 2019; Yu et al., 2019].Calmodulin related gene has been showed in many cancers and CALML3 was reported to be a potential biomarkers for pulmonary metastasis of hepatocellular carcinoma [Yang et al., 2018], while the dysregulation of CALM1 in hepatocellular carcinoma was firstly found in our study[Bhagwan et al., 2020]. The increasing expression of HSPA8 was reported in hepatocellular carcinoma and depressive disorder [Xiang et al., 2018], and previous studies also showed that SLC26A8 (solute carrier family 26 member 8) was related with many cancers including colorectal cancer, and mutation of SLC26A8 also related to many diseases [Dirami et al., 2013; El Khouri and Toure, 2014; Yu et al., 2018].

It was noteworthy that SELENBP1 (Selenium binding protein 1) has been reported downregulated in colorectal cancer, while upregulated in hepatocellular carcinoma in this study, and it was also confirmed in qRT-PCR in the validation cohort [Lee et al., 2020]. As shown in figure.8B, higher specificity and sensitivity was obtained when distinguishing liver cancer from normal samples by expression level of SELENBP1.

In this present study, CALM1 were found downregulated in hepatocellular carcinoma patient, while the *p* value is 0.063 that indicated no significant differences. However we still selected it for two reasons, one is for it was found in the gene sets from WGCNA modules, and the other is gradually changes found in patient with tumor sizes from small to large. As expected, significant differences obtained from qRT-PCR between liver cancer patients and normal controls. The gradually decreasing expression of CALM1 was also confirmed by RT-PCR as shown in Figure 8C.

It cannot be denied that there still have limits in this study. On the one hand, no result provided from hepatocellular carcinoma tissues for comparison. On the other hand, the patient cohort involved in this study was relative small. There are too many physiological differences among these patients with hepatocellular carcinoma although 17 samples were sequenced, and only several samples could be used for comparison when analyzing the correlations between traits and different gene modules which leading

to the current results is not to be reliable enough. However, potential biomarkers could be screened from the DEGs obtained by PBMC RNA sequencing of hepatocellular carcinoma.

5. Conclusions

The DEGs were profiled by RNA-Seq from PBMCs of hepatocellular carcinoma patients in this present study. 1578 DEGs founds between hepatocellular carcinoma and healthy controls including 1334 upregulated genes and 244 downregulated genes. Gene functional enrichment analysis results showed that most of the genes were involved in immune reactions related to hepatocellular carcinoma. Several DEGs selected (SELENBP1, SLC4A1, SLC26A8, HSPA8P4, CALM1 and RPL7p24) were confirmed by qRT-PCR, and to our knowledge, SLC4A1, RPL7p24, CALM1 and SLC26A8 were firstly found related to hepatocellular carcinoma, and it is suggested the potential biomarkers could be analyzed for classification, stages and therapeutic target of hepatocellular carcinoma in the future studies.

Abbreviations

Ethics approval and consent to participate

Patients involved in this study provided written informed consent. Ethics approval was obtained from the Ethics Committee of Shenzhen Traditional Chinese Medicine Hospital. All experiments were performed in accordance with relevant guidelines and regulations set out by the ethical committee.

Consent for publication

Not applicable

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files. And the raw sequencing datasets during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

Shenzhen Science and Technology Project (NO.JCYJ20170817094901026, JCYJ20180302173542393) & Sanming Project of Medicine in Shenzhen (SZSM201612074).

Authors' contributions

Zhou XZ and Han ZY contributed to the concept, design, experiment and analysis of data, in addition to the preparation of the manuscript. Feng WX; Hu R; Ma WF; Zhang W; Xu SM and Zhan BL contributed to sample collection. Han ZY and Ge QY contributed to the bio-informatic analysis. Feng WX and Hu R contributed to the acquisition of data. Sun XF and Zhang L contributed to specialized guidance. All the authors have approved the manuscript.

Acknowledgments

Thanks to the grant from the Shenzhen Science and Technology Project, and Sanming Project of Medicine in Shenzhen (NO.JCYJ20170817094901026, JCYJ20180302173542393, SZSM201612074).

References

2020 Expression of Concern: Activation of lncRNA lnc-SLC4A1-1 induced by H3K27 acetylation promotes the development of breast cancer via activating CXCL8 and NF-kB pathway. Artif Cells Nanomed Biotechnol 48:706.

Attwa MH, El-Etreby SA. 2015. Guide for diagnosis and treatment of hepatocellular carcinoma. World J Hepatol 7:1632-51.

Befeler AS, Di Bisceglie AM. 2002. Hepatocellular carcinoma: diagnosis and treatment. Gastroenterology 122:1609-19.

Bhagwan JR, Mosqueira D, Chairez-Cantu K, Mannhardt I, Bodbin SE, Bakar M, Smith JGW, Denning C. 2020. Isogenic models of hypertrophic cardiomyopathy unveil differential phenotypes and mechanism-driven therapeutics. J Mol Cell Cardiol 145:43-53.

Budhu A, Forgues M, Ye QH, Jia HL, He P, Zanetti KA, Kammula US, Chen Y, Qin LX, Tang ZY, Wang XW. 2006. Prediction of venous metastases, recurrence, and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment. Cancer Cell 10:99-111.

Cools-Lartigue J, Spicer J, McDonald B, Gowing S, Chow S, Giannias B, Bourdeau F, Kubes P, Ferri L. 2013. Neutrophil extracellular traps sequester circulating tumor cells and promote metastasis. J Clin Invest.

Dirami T, Rode B, Jollivet M, Da Silva N, Escalier D, Gaitch N, Norez C, Tuffery P, Wolf JP, Becq F, Ray PF, Dulioust E, Gacon G, Bienvenu T, Toure A. 2013. Missense mutations in SLC26A8, encoding a sperm-specific activator of CFTR, are associated with human asthenozoospermia. Am J Hum Genet 92:760-6.

Edoo MIA, Chutturghoon VK, Wusu-Ansah GK, Zhu H, Zhen TY, Xie HY, Zheng SS. 2019. Serum Biomarkers AFP, CEA and CA19-9 Combined Detection for Early Diagnosis of Hepatocellular Carcinoma. Iran J Public Health 48:314-322.

El-Serag HB, Rudolph KL. 2007. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology 132:2557-76.

El Khouri E, Toure A. 2014. Functional interaction of the cystic fibrosis transmembrane conductance regulator with members of the SLC26 family of anion transporters (SLC26A8 and SLC26A9): physiological and pathophysiological relevance. Int J Biochem Cell Biol 52:58-67.

Forner A, Bruix J. 2012. Biomarkers for early diagnosis of hepatocellular carcinoma. Lancet Oncol 13:750-1.

Gazzaniga P, Raimondi C, Nicolazzo C, Carletti R, di Gioia C, Gradilone A, Cortesi E. 2015. The rationale for liquid biopsy in colorectal cancer: a focus on circulating tumor cells. Expert Rev Mol Diagn 15:925-32.

Greenhill C. 2018a. New pathways in development of liver cancer. Nat Rev Endocrinol 15:2.

Greenhill C. 2018b. New pathways in development of liver cancer. Nat Rev Gastroenterol Hepatol 15:718.

Hou JM, Krebs MG, Lancashire L, Sloane R, Backen A, Swain RK, Priest LJ, Greystoke A, Zhou C, Morris K, Ward T, Blackhall FH, Dive C. 2012. Clinical significance and molecular characteristics of circulating tumor cells and circulating tumor microemboli in patients with small-cell lung cancer. J Clin Oncol 30:525-32.

Hu L, Zhu Y, Zhang J, Chen W, Li Z, Li L, Zhang L, Cao D. 2019. Potential circulating biomarkers of circulating chemokines CCL5, MIP-1beta and HA as for early detection of cirrhosis related to chronic HBV (hepatitis B virus) infection. BMC Infect Dis 19:523.

Lee YM, Kim S, Park RY, Kim YS. 2020. Hepatitis B Virus-X Downregulates Expression of Selenium Binding Protein 1. Viruses 12.

Maheswaran S, Haber DA. 2010. Circulating tumor cells: a window into cancer biology and metastasis. Curr Opin Genet Dev 20:96-9.

Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. 2008. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 18:1509-17.

Mi H, Muruganujan A, Huang X, Ebert D, Mills C, Guo X, Thomas PD. 2019. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). Nat Protoc 14:703-721.

Raimondi C, Gradilone A, Naso G, Vincenzi B, Petracca A, Nicolazzo C, Palazzo A, Saltarelli R, Spremberg F, Cortesi E, Gazzaniga P. 2011. Epithelial-mesenchymal transition and stemness features in circulating tumor cells from breast cancer patients. Breast Cancer Res Treat 130:449-55.

Riethdorf S, Muller V, Zhang L, Rau T, Loibl S, Komor M, Roller M, Huober J, Fehm T, Schrader I, Hilfrich J, Holms F, Tesch H, Eidtmann H, Untch M, von Minckwitz G, Pantel K. 2010. Detection and HER2 expression of circulating tumor cells: prospective monitoring in breast cancer patients treated in the neoadjuvant GeparQuattro trial. Clin Cancer Res 16:2634-45.

Shen Y, Bu L, Li R, Chen Z, Tian F, Ge Q. 2019. Expression And Biological Interaction Network Of RHOC For Hepatic Carcinoma With Metastasis In PBMC Samples. Onco Targets Ther 12:9117-9128.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7:562-78.

Varela-Martinez E, Abendano N, Asin J, Sistiaga-Poveda M, Perez MM, Reina R, de Andres D, Lujan L, Jugo BM. 2018. Molecular Signature of Aluminum Hydroxide Adjuvant in Ovine PBMCs by Integrated mRNA and microRNA Transcriptome Sequencing. Front Immunol 9:2406.

Wadhwa R, Ryu J, Ahn HM, Saxena N, Chaudhary A, Yun CO, Kaul SC. 2015. Functional significance of point mutations in stress chaperone mortalin and their relevance to Parkinson disease. J Biol Chem 290:8447-56.

Wan Q, Tang J, Han Y, Wang D. 2018. Co-expression modules construction by WGCNA and identify potential prognostic markers of uveal melanoma. Exp Eye Res 166:13-20.

Xiang X, You XM, Li LQ. 2018. Expression of HSP90AA1/HSPA8 in hepatocellular carcinoma patients with depression. Onco Targets Ther 11:3013-3023.

Yang B, Li M, Tang W, Liu W, Zhang S, Chen L, Xia J. 2018. Dynamic network biomarker indicates pulmonary metastasis at the tipping point of hepatocellular carcinoma. Nat Commun 9:678.

Yi T, Zhou X, Sang K, Huang X, Zhou J, Ge L. 2019. Activation of lncRNA lnc-SLC4A1-1 induced by H3K27 acetylation promotes the development of breast cancer via activating CXCL8 and NF-kB pathway. Artif Cells Nanomed Biotechnol 47:3765-3773.

Yu C, Hong H, Zhang S, Zong Y, Ma J, Lu A, Sun J, Zheng M. 2019. Identification of key genes and pathways involved in microsatellite instability in colorectal cancer. Mol Med Rep 19:2065-2076.

Yu L, Yin B, Qu K, Li J, Jin Q, Liu L, Liu C, Zhu Y, Wang Q, Peng X, Zhou J, Cao P, Cao K. 2018. Screening for susceptibility genes in hereditary non-polyposis colorectal cancer. Oncol Lett 15:9413-9419.

Yu QJ, Liang YZ, Mei XP, Fang TY. 2020. Tumor mutation burden associated with miRNA-gene interaction outcome mediates the survival of patients with liver hepatocellular carcinoma. EXCLI J 19:861-871.

Zhang C, Peng L, Zhang Y, Liu Z, Li W, Chen S, Li G. 2017. The identification of key genes and pathways in hepatocellular carcinoma by bioinformatics analysis of high-throughput data. Med Oncol 34:101.

Zhu K, Dai Z, Zhou J. 2013. Biomarkers for hepatocellular carcinoma: progression in early diagnosis, prognosis, and personalized therapy. Biomark Res 1:10.

Tables

Table 1 **DNA sequences of the primers used in this study**

Name	DNA sequences of primers
RPL7p24	Forward: CAAGGCTTCGATTAACATGCTGA
	Reverse: GCCATAACCACGCTTGTAGATT
CALM1	Forward: TTGACTTCCCCGAATTTTTGACT
	Reverse: GGAATGCCTCACGGATTTCTT
HSPA8P4	Forward: ATGCCAAACGTCTGATTGGAC
	Reverse: AGCATCATTCACCACCATAAAGG
SLC26A8	Forward: CATGGCACAGGTTCCTACGAT
	Reverse: GGCCAACACTTATACCAGCAAG
SLC4A1	Forward: CCTATACGCTTCCTCTTTGTGTT
	Reverse: CCATGTAGGCATCTATGCGGA
SELENBP1	Forward: ACCCAGGGAAGAGATCGTCTA
	Reverse: ACTTGGGGTCAACATCCACAG
GAPDH	Forward: ACAACTTTGGTATCGTGGAAGG
	Reverse: GCCATCACGCCACAGTTTC

Figures

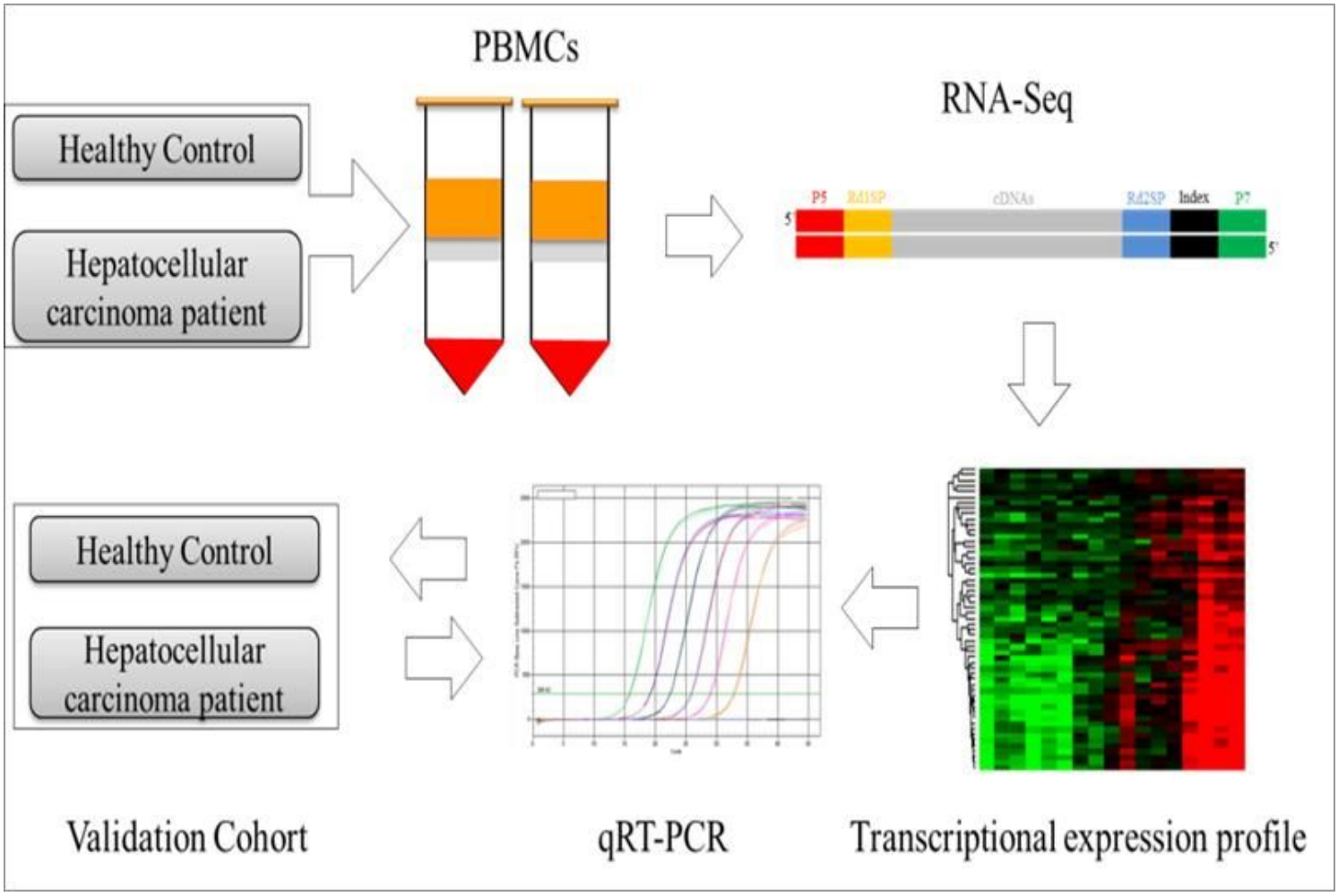


Figure 1

The experiment scheme of the study

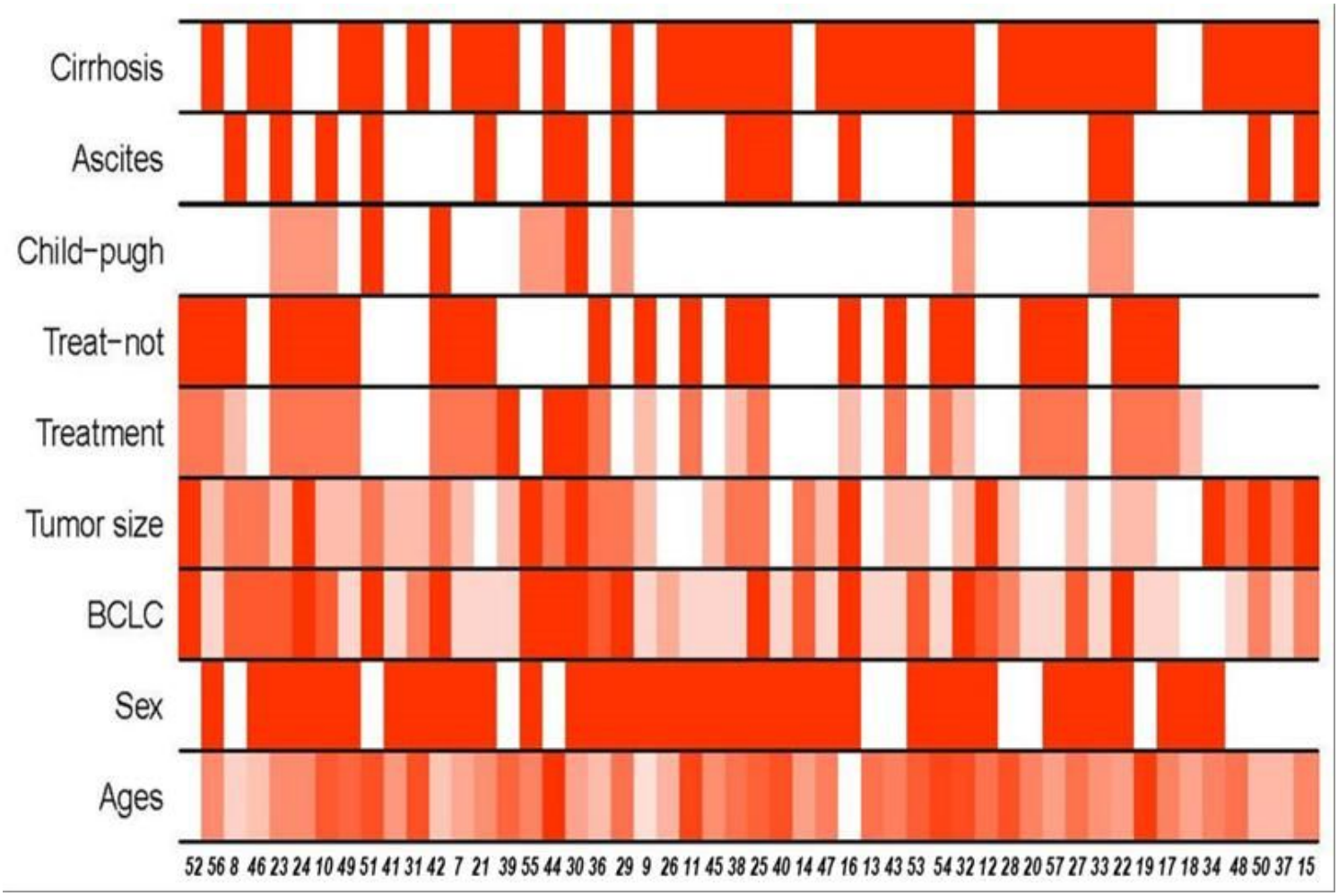


Figure 2

Clinical characteristics of HCC samples: For PBMCs was collected from 50 patients with HCC. Shown, in descending order, are whether cirrhosis, whether ascites, classification of Child-Pugh, whether treatment or not, treatment method, tumor size, BCLC classification, sex and ages of those patients at time of sampling.

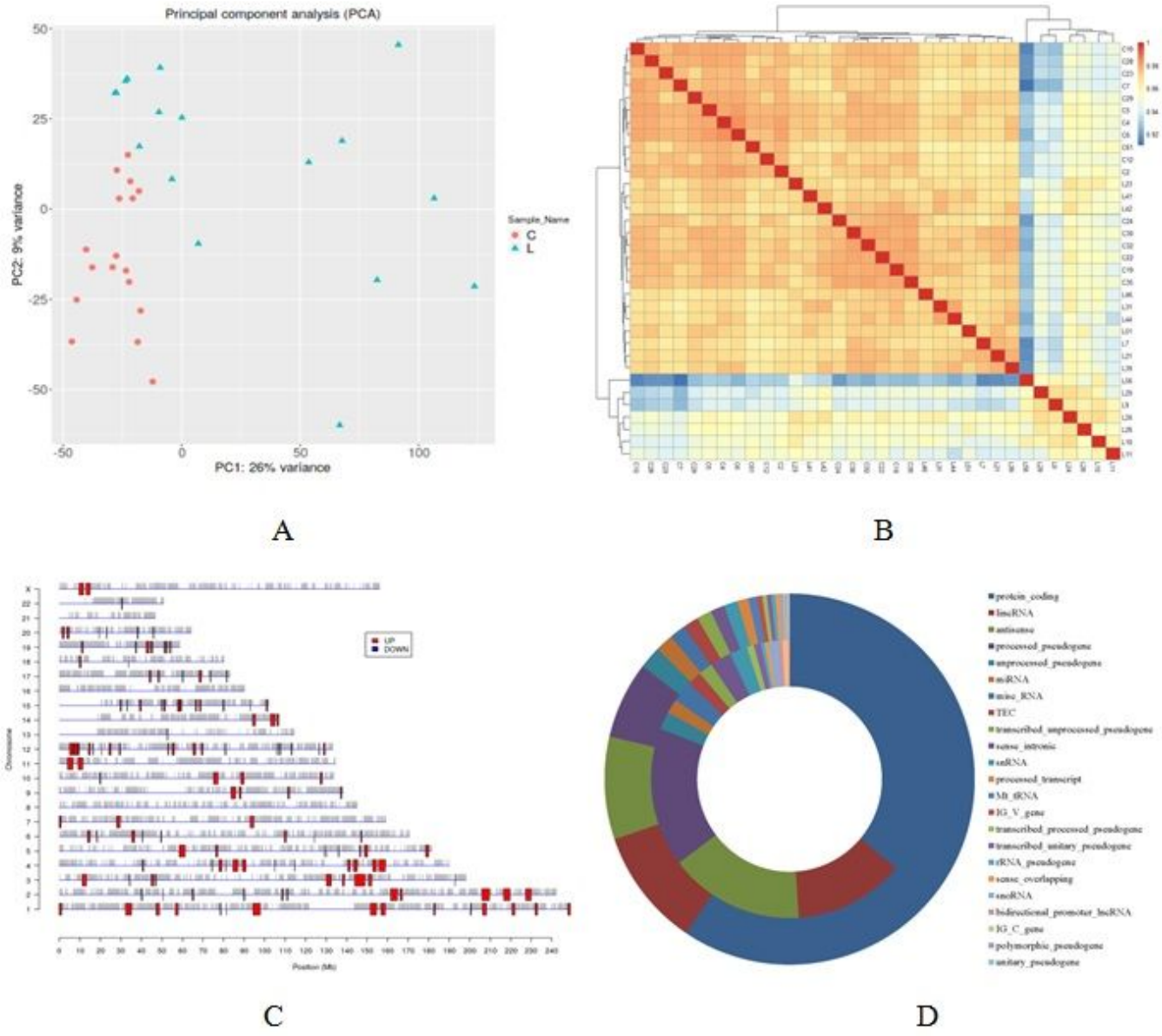


Figure 3

(A) Load plot of PCA; (B) The pearson's correlation coefficient of significantly dysregulated mRNA expression the heatmap of inter-sample correlation; (C) Distribution of the different expressed genes located in human chromosomes; (D) Types of these differential expressed genes sequenced, the outside ring indicated the upregulated genes and the inner ring indicated downregulated genes.

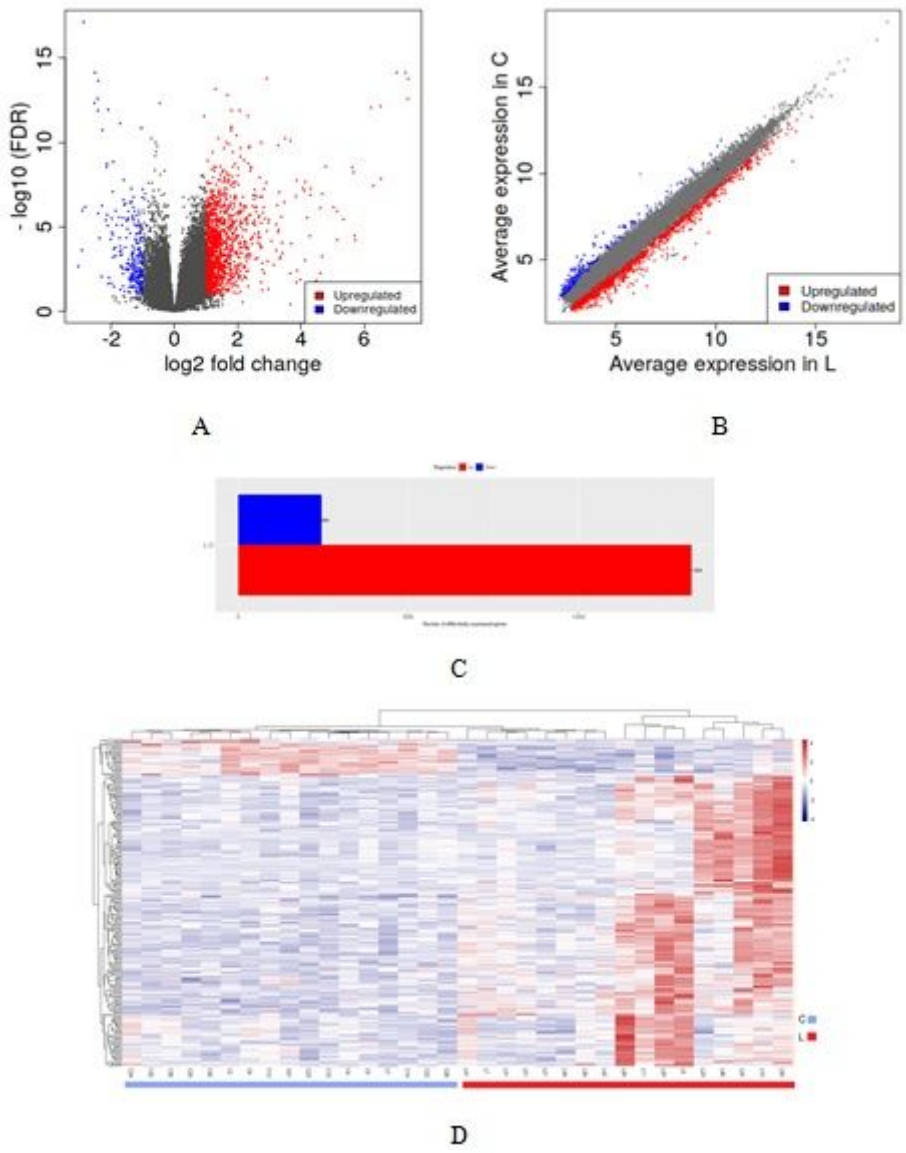


Figure 4

Differential gene expression profiles between HCC and healthy control. (A, B) Volcano plots and scatter plot showing differential expression of mRNAs between HCC and NC. The red (up) and blue (down) points represent the differentially expressed RNAs with fold change ≥ 2.0 and p-Value < 0.05; (C) The number of DEGs in PBMCs of HC was summarized and (D) Hierarchical clustering analysis of the DEGs in PBMCs of HC.

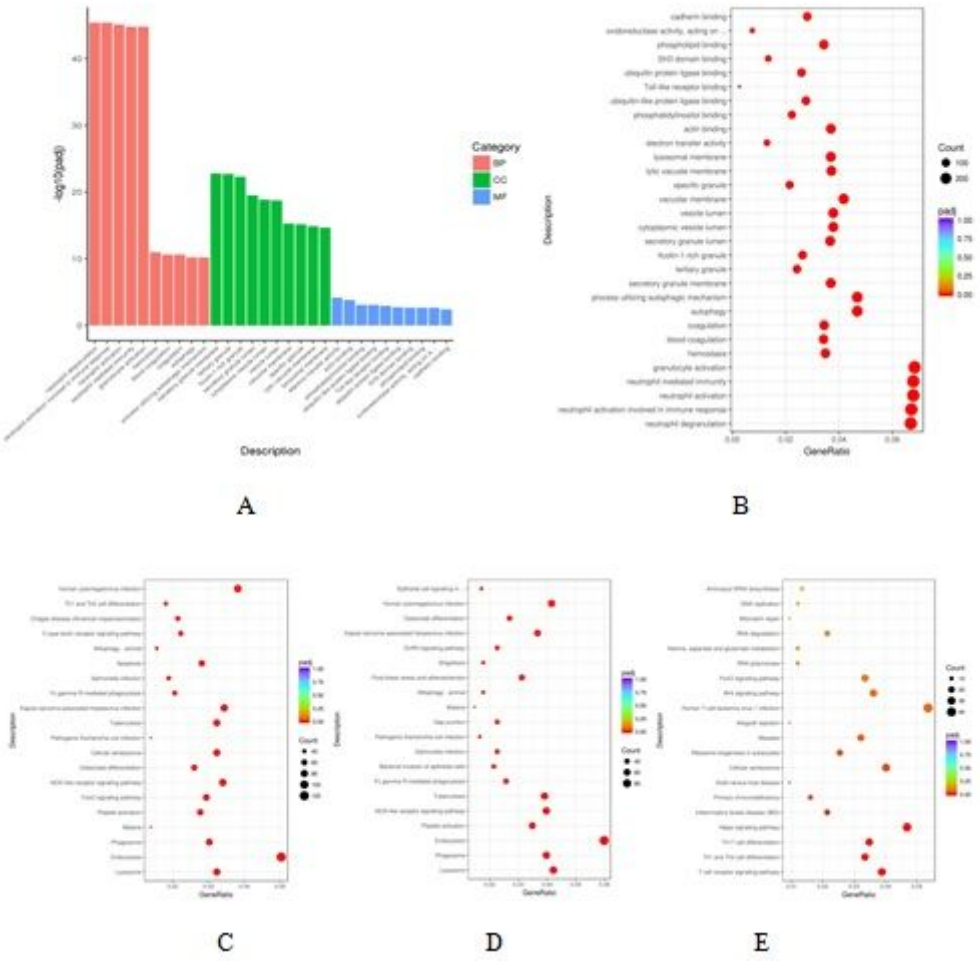
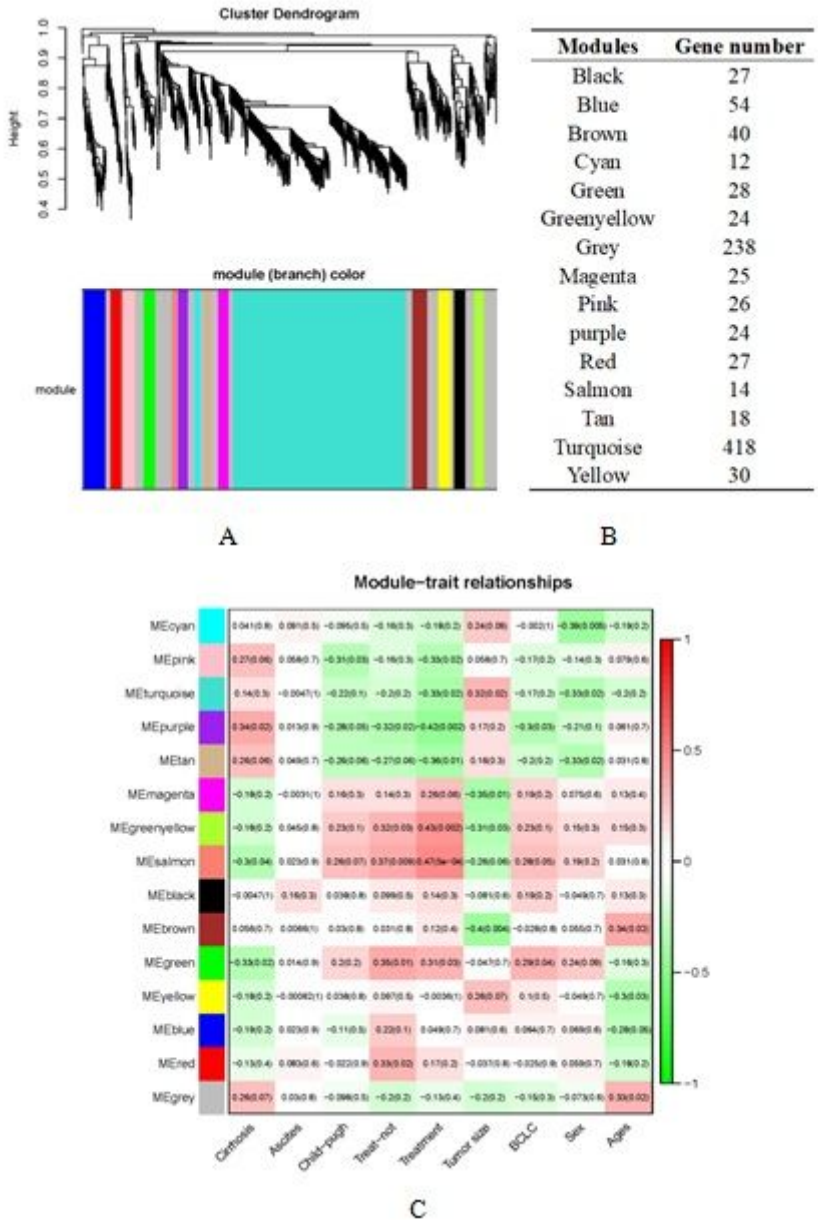


Figure 5

GO and KEGG terms analysis of DEGs in HCC PBMC. (A) GO enrichment analysis result of DEGs. From left to right are biological process, cellular component, molecular function, and the all GO term in the figure with p-value < 0.05; (B) GO term enrichment; (C) The top 20 KEGG pathways analysis result of DEGs, the KEGG term with p-value < 0.01; (D) The KEGG pathways analysis result of upregulated expressed genes; (E) The KEGG pathways analysis result of downregulated expressed genes.



Modules	Gene number
Black	27
Blue	54
Brown	40
Cyan	12
Green	28
Greenyellow	24
Grey	238
Magenta	25
Pink	26
purple	24
Red	27
Salmon	14
Tan	18
Turquoise	418
Yellow	30

Figure 6

Co-expression modules analysis with WGCNA. (A) Gene co-expression modules of DEGs in HCC. The identified modules have different colors below the dendrogram; (B) The number of genes in the 14 modules showed on the list; (C) Module-trait associations. Each row corresponds to a module gene, and a column as a trait. Each cell contains the corresponding correlation and p-value (in brackets). The table is color-coded by correlation according to the color legend.

Figure 7

PPI network of DEGs in HCC. The ball surrounded by shadows denoted upregulated genes and others showed downregulated genes.

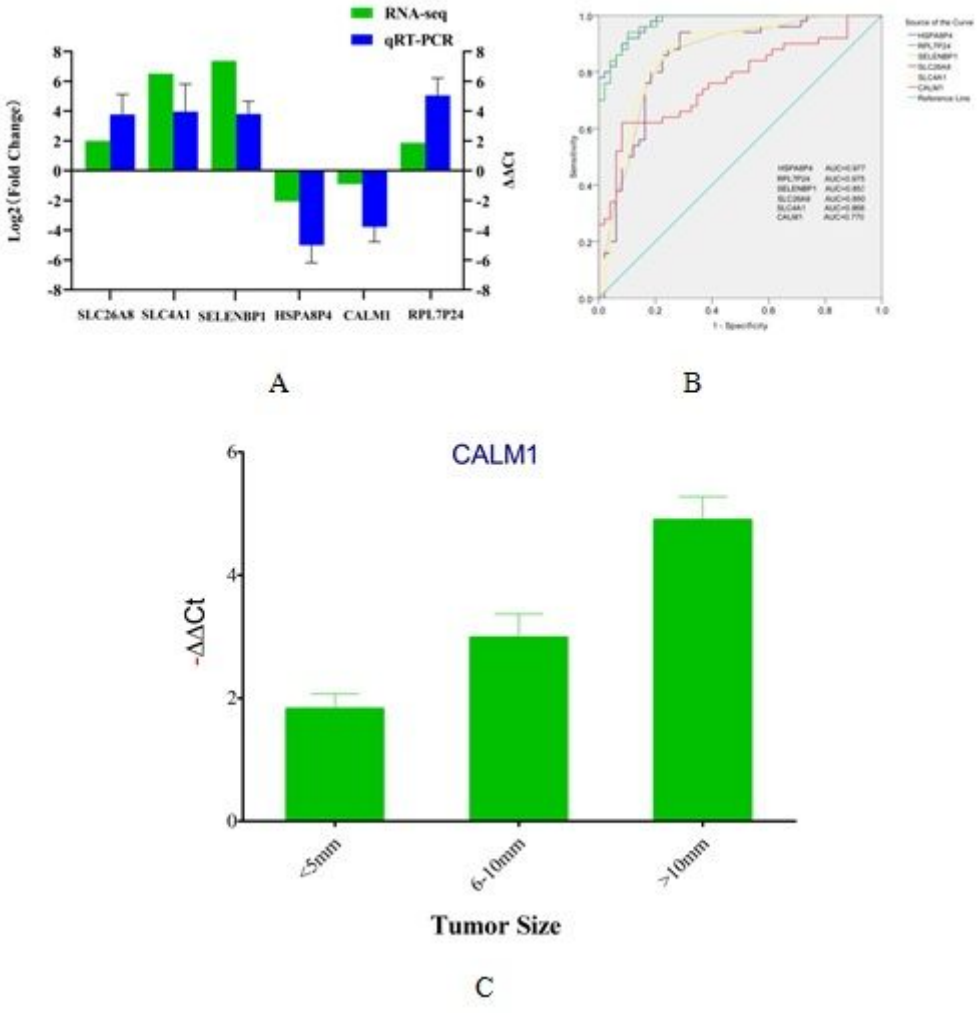


Figure 8

Validation of the PBMC RNA by qRT-PCR. (A) The expressions of selected dysregulated genes in HCC samples from RNA-Seq data were evaluated using qRT-PCR in the sample from 50 HCC patients and 49 normal controls. (B) ROC curve analysis of SELENBP1, SLC4A1, SLC26A8, HSPA8P4, CALM1, and RPL7p24, AUC values are given on the graphs. (C) Relationship of CALM1 gene expression and tumor sizes, the gene decreased gradually when tumor enlarged.