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
Shenzhen Traditional Chinese Medicine Hospital
Xiaozhou Zhou (zxz1006@gzucm.edu.cn)
Shenzhen Traditional Chinese Medicine Hospital

Research Article

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

DOI: https://doi.org/10.21203/rs.3.rs-280310/v1

Abstract

Background: Hepatocellular carcinoma with extremely high morbidity and mortality is one of the most common malignant tumors. Although many existing studies focus on the study of its biomarkers, little information has been released on the PBMC RNA profile of hepatocellular carcinoma.

Methods: We tried to make a profile throughout this analysis with 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 studies showed that hepatocellular carcinoma is a closely linked source of the most differentially expressed genes (DEGs), implicated in the immune response. Expression of the 6 selected genes (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 of hepatocellular carcinomas. This study provided new perspectives for liver cancer development and possible future successful clinical diagnosis.

1. Introduction

One of the major causes of cancer mortality in the world is hepatocellular carcinoma (HCC) [1]. Even though there is much new breakthrough in cancer diagnosis and treatment at surgical, radiological, chemotherapy and biological, it is worthwhile to note that when cancer cells metastasized and forecasts were unsatisfactory, the majority of cancers could not be targeted for treatment. The above cases are primarily due to factors like delay or absence in diagnosis [2-4]. Most candidates for liver cancer conditions are patients with underlying liver diseases (such as Hepatitis B virus (HBV) infection and cirrhosis) [5]. More than half of HCC sick persons are diagnosed as advanced, restricting treatment of cancer. Imaging diagnoses, like positron emissions tomography (PET, for example), is a very specific tool for diagnosing hepatocellular carcinoma, although it lacks typical small to micrometastase imaging features. Alpha-fetoprotein (AFP) and alkaline phosphatase (ALP or AKP) are common at the moment, but the biomarkers are unsuitable in clinical practice for the early diagnosis of hepatocellular carcinoma. AFP (threshold of 20 μg/mL) is reported to have a small sensitivity of 40% to 60%, and a specificities of 80% to 90% [6-8]. False negatives (e.g., small HCCs with ordinary AFP standard), false positive positive ones (e.g., liver injury and certain gastrointestinal tumors) and the low sensitivity might decrease the likelihood of early diagnosis which could lead to clinical outcomes. Deterioration that also emphasizes that HCC detection techniques need to be more contributing.

Here, we investigated peripheral blood mononuclear cells (PBMCs) transcriptomes in HCC and evaluated the diagnostic value of PBMCs transcripts. PBMCs, a sample simple to access and least invasive, could be isolated from patients with hepatocellular carcinoma and healthy control in this present study [9]. Using simple elements to the platform will improve the standard of the solution and give more reliable results. Besides, using RNA-seq for initial genetic screen of dys-regulated mRNAs [10]. Then, researchers classified PCA based on the similarities of gene expression, and reduced the heterogeneity between tumors by minimizing the size of the effector. Eventually, choosing the path or processes involved in the disorder, and confirms the new potential transcript by qRT-PCR. The design of the study could be seen in **Figure.1**. The project will enable us to know the progression of tumors and lay the foundation for clinical diagnosis in the long run.

Transcriptomes of PBMCs from HCC and controls were profiled by RNA-Seq and then analyzed by bioinformatics methods. Subsequently, the screened genes were validated by qRT-PCR in the additional validation cohort with 33 hepatocellular carcinoma patients and 32 healthy controls.

2. Materials And Methods

2.1. Separation of PBMCs and RNA extraction

Researchers gathered 34 people with peripheral blood (about 2 milliliters) (17 specimen of HC patients and the 17 from the healthy control people who have similar gender and age with these cancer patients, additional 33 patient samples and 32 healthy controls were also collected for validation by qRT-PCR, the clinical characteristics of samples collected shown in Figure.2), For this demonstration, Ficoll-Paque^{TM} PREMIUM was utilized to separate PBMC from 2 ml of EDTA blood of the donor according to business plan. Subsequently, the PBMCs specimens were processed with Trizol reagent (Invitrogen, Carlsbad, CA) according to the previously noted standard procedure (Marioni et al., 2008). NanoDrop ND-1000 (Thermo Fisher Scientific, Waltham, MA) is used by researchers to obtain quality of RNA through 260 nm (A260) and 280 nm (A280) of absorption and to evaluate integrity of RNA by RNA integrity number (RIN; Agilent 2100 RIN Beta Version Software).

Researchers received ethical clearance from the Ethical Committee of Shenzhen Traditional Chinese Medicine Hospital. In compliance with the principles and regulations laid down by the ethics committee, all experiments were conducted.

2.2 RNA-seq

34 RNA samples collected are sent to the sequencing researcher. 17 of these were samples extracted and matched by healthy controls from patients with hepatocellular carcinoma. Based on the previous studies, the sequencing library was formulated after the removal of rRNA as per the Illumina®TruSeq®RNA sample preparation guide (Illumina, San Diego, California, USA). The index adaptor should be ligated

once the double-stranded cDNA had been synthesized. After size selection by using Agencourt AMPure XP (Beckman), Qubit ® 2.0 Fluorometer and Qubit ® dsDNA HS Analyse Kit (Invitrogen™, Eugene, OR, USA) and Agilent Bioanalyzer Quantitative and qualitative library (Agilent Technologies, Santa Clara, CA, USA), respectively, and it was submitted to Illumina HiSeq X-10 (Illumina, San Diego, CA, USA) for pair-end innovation sequences of 2×150 bp.

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 supplied by Illumina iGenomes (Download source: 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) [11]. When the results were analyzed, DEG demonstrated a major change if the value of p is less than 0,05 and the folding change is more than 2 times.

The researchers utilized the PANTHER (protein annotation through evolutionary relationship) classification system for GO assessment in the functional analysis section (http://www.pantherdb.org/) [12]. Formed on the Mann-Whitney test, the researchers performed a statistically significant overrepresentation test, which can be used to evaluate if any ontology class or path has a non-randomly distributed significance relative to the overall set of attributes.

Then, researchers study the co-expression modules for DEGs in liver cancer using Weighted Gene Correlation Network (WGCNA) analysis and examined the link between various modules and clinical properties (such as age, gender, tumor size, and classification). Ultimately, the Protein-Protein Interaction (PPI) framework of DEGs was visualized with Cytoscape software, with the purpose of gaining insights into the involvement of DEGs (version 3.7.2).

2.4 qRT-PCR

For the qRT-PCR verification, 65 additional RNA samples should be submitted. Of these, 33 cases and 32 cases were extracted from hepatocellular carcinoma patients and matched healthy samples, respectively. Using the PrimeScript^{TM} RT Master Mix (Perfect Real Time)(Takara Bio, Inc.) for 500 ng of total RNA for the reverse transcription reaction, at 37˚C and at 85˚C for 15 minutes and 5 seconds, ultimately with a final volume of 20 µL. The following qPCR was conducted 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. As shown in Table 1, Sangon Biotech, Inc. synthesized the primers. The PCR reaction took three minutes at 95°C, then 5 seconds at 95°C for fourty cycles and then 40 sec. at 60°C. Every reaction was repeated three times at that time, with an average Ct value calculated for each triple. The ΔCt target cDNA was the difference between Ct target gene and Ct of reference gene (GAPDH). To assess the gene expression fold changes, the ΔΔCt value between ΔCt of

HCC and ΔCt of NC was adopted (ΔΔCt = ΔCt HC -ΔCt NC; ΔCt = Ct target – Ct reference). By using 2% agarose gel and a dissociation curve, the quality of the amplified product could be achieved.

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

2.5 Statistical analysis:

All of data analysis was performed with R (version 4.0.1) and SPSS (version 22.0). To evaluate the similitude of the samples, researchers use PCA and Pearson correlation analysis. To evaluate the correlations, Spearman correlation analysis was conducted. A t-test was examined for statistical differences. The two-sided inspections were carried out in all mathematical inspections. If the p value was less than 0.05, it could be seen as statistically significant.

3. Results

3.1 Baseline characteristics and RNA-seq information of samples

Researchers used RNA-Seq pair-end for displaying profiles of gene expression in PBMCs from 17 hepatocellular patients and 17 appropriate-age healthy individuals (as a control group study). The raw reads RNA-Seq from 58,012,158 to 83,083,036 is in line with the human reference hg19, which represented readings mapped to exons from 22,894,689 to 42,821,652 (37,879% -57.238%). The baseline features of the seventeen patients with hepatocellular carcinoma could be seen in Appendix Figure 2. These include age, gender, BCLC and grade information, the size of the tumors, and whether the therapy was done on sampling.

3.2. Characterizing the gene expression profiles of HCC

The researchers analyzed the correlation coefficients between PCA and pearson's correlation to assess the similarity or dissimilarity of the RNA-seq output for an insight into the characteristics of profiles of the expression of liver cancer generation. PCA is a linear projection methodology that enables researchers to visualize large data in a smaller space. The findings suggested that 26% of the total data variance was represented by the first main component (PC1), while 9% of the figures represent the second main component (PC2).As Figure 3A demonstrates, in PC1, the control samples and the hepatocellular carcinoma samples were gathered respectively. Each one of them had similar effects and showed that they are mutually similar while PC1 showed the basic features of the expression pattern of hepatocellular carcinoma. In addition, it had been shown that there are certain distinctions in gene expression profiles between the two groups of PC2 in the collection of control samples into one group and in cancer samples simultaneously.

The heatmap of inter-sample correlation showed the Pearson correlation factor for the mRNA expression that was considerably dysregulated. It can be seen from the Figure 3B that in terms of the hepatocellular carcinoma group and the control group, there were obvious differences in the level of transcription expression between the two groups, but the differences between the groups are small. In addition, differentially expressed genes (DEGs) were categorized by the researchers depending on multiple chromosome position (Figure.3C). All chromosomes might show the expression of the related genes, 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. These molecules are indicated to be able to have potential roles for hepatocellular carcinoma incidence and growth.

3.3. Different expressed genes in HCC

RNA-seq data revealed that 6,0006 genes in these 34 PBMC samples were expressed from RNA-seq data. After the researchers filtered, on the basis of the volcano plot and split plot, 1578 genes were regulated by the difference of fold change greater than or equal to 2.0 between HC and an ordinary control group and by p value of less than 0.05 (Figure 4A). What's more, most DEGs are genes that encode proteins. In

PBMCs of patients with hepatic carcinoma 1334 genes were up-regulated and 244 genes were down-regulated, the top 25 dysregulated expressed genes were shown in supplemental Table S1. The hierarchical clustering of these dysregulated genes had shown that gene expression profiles are distinguishable between hepatic carcinoma and control group (Figure 4B). The research results showed that PBMC RNA is distinctive from that of healthy controls for hepatocellular carcinoma.

3.4 GO and KEGG pathway analysis of differentially expressed genes

These molecules are indicated to have possible roles for hepatocellular carcinoma incidence and advancement. From Figure.5A and 5B, researchers discovered that in biological processes, important terms rich in GO are neutrophil degradation, activation of the Neutrophil, neutraphil-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 adhesion, actin filament binding, SH2 domain binding, phosphatidylinositol binding, Toll-like receptor binding, phospholipid binding, GTPase activator activity, ad GTPase regulator activity. Among them, the GO terms of up-regulated genes are significantly rich, mainly related to immune reaction, secretive granule lumen, toll-like receptor binding and body fluid level regulation, which have been confirmed to be cancer-related. 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.

On the basis of KEGG annotations, the pathway study revealed the top twenty signal pathways, along withg 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. For the purpose of more effectively linking information with HCC, WGCNA used DEG derived from 17 patients with HCC to formulate the co-expression module. The researcher defined the amount of genes in each module, the number is at least 10, and the depth of the cutting is 0.8. Finally, 14 gene modules in hepatocellular carcinoma had been defined and shown in multiple colors (Figure 6A), genes not allocated to either of the modules are returned to the gray module. Figure 6B demonstrated the various DEG numbers in the 14 gene modules. 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 suggests that variously expressed genes in liver cancer have nothing to do with the influence of 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, the researchers found that the genes in the 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 a great part in the generation, development and molecular regulation of HCC. It was worth noting that the researchers found that the analysis of the module-trait showed that the salmon module had a certain correlation with several clinical features, 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. The result of the genetic functional enrichment research indicate that the great mass of the genes associated in neutrophil mediated immunity, leukocyte, mediated immunity, leukocyte activation, immune system process and cell migration, adhesion and motility.The protein-protein interaction (PPI) system of DEG is formulated from the first 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

In order to verify the sequence performance, the researchers adopted qRT-PCR to verify the expression of DEG in PBMCs of these other HCC patients (n=33) and in healthy control group (n=32). 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 HCC 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 Figure 8A is shown, the qRT-PCR outcomes revealed similarly changes, and significant differences found between group HCC and normal control in all the 6 genes.

The ROC curve assessment conducted by the researchers may evaluate the diagnostic value of these DEGs for the presence of HCC patients. As shown 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 findings demonstrate that these genes have prospective diagnostic value in patients with liver cancer.

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 it is strongly linked to the occurring, growth, metastases and prognosis of tumors caused by irregular 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 the suitable choice for studying gene expression and tumor monitoring.

Using RNA-seq, the researchers examined 1,334 up-regulated and 244 down-regulated genes in PBMCs of liver cancer, with a fold change greater than or equal to 2.0 and a p value below 0.05. The DEGs defined might be generated from tumor cells circulating in the vasculature and therefore circulating throughout the blood [13].

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 [14-16].

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 [17-19].Calmodulin related gene has been showed in many cancers and CALML3 was reported to be a potential biomarkers for pulmonary metastasis of hepatocellular carcinoma [20], while the dysregulation of CALM1 in hepatocellular carcinoma was firstly found in our study[21]. The increasing expression of HSPA8 was reported in hepatocellular carcinoma and depressive disorder [22], 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 [23-25].

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 [26]. 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. Functional analysis of gene expression of hepatocellular carcinoma revealed that the majority of the genes in the hepatocellular carcinoma samples were related to immune responses. 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.

Declarations

Acknowledgments:

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

Funding

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

Conflicts of interest/Competing interests

The writers have declared that they do not have conflicts of interest.

Availability of data and material

All test statistic produced or evaluated in the course of this research is disclosed in the paper as well as its supplementary information files. According to reasonable requirements, the original sequencing data from this research can be obtained from the corresponding author.

Code availability

Not applicable

Authors' contributions:

Due to the cooperation in the definition, design, research and analysis of the data, both Zhou XZ and Han ZY also contributed to the work. Feng WX; Hu R; Ma WF; Zhang W; Xu SM and Zhan BL contributed to sample collection. Han ZY and Ge QY played a part to the bio-informatic study. Feng WX and Hu R made contribution to the data gathering. Sun XF and Zhang L made a contribution to specialized instructions. All the writers have given their consent for publishing the manuscript.

Ethics approval

All patients participating in the study provided their informed understanding in writing. Ethics approval was obtained from the Ethics Committee of Shenzhen Traditional Chinese Medicine Hospital. All tests were conducted according to the pertinent guidelines and regulations implemented by this Ethics Committee.

Consent to participate

All volunteers signed an informed consent to participate in the experimental project.

Consent for publication

Not applicable

References

1. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology.* 2007;132(7):2557-2576.
2. Befeler AS, Di Bisceglie AM. Hepatocellular carcinoma: diagnosis and treatment. *Gastroenterology.* 2002;122(6):1609-1619.
3. Budhu A, Forgues M, Ye QH, et al. Prediction of venous metastases, recurrence, and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment. *Cancer cell.* 2006;10(2):99-111.
4. Attwa MH, El-Etreby SA. Guide for diagnosis and treatment of hepatocellular carcinoma. *World journal of hepatology.* 2015;7(12):1632-1651.
5. Hu L, Zhu Y, Zhang J, et al. 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 infectious diseases.* 2019;19(1):523.
6. Edoo MIA, Chutturghoon VK, Wusu-Ansah GK, et al. Serum Biomarkers AFP, CEA and CA19-9 Combined Detection for Early Diagnosis of Hepatocellular Carcinoma. *Iranian journal of public health.* 2019;48(2):314-322.
7. Zhu K, Dai Z, Zhou J. Biomarkers for hepatocellular carcinoma: progression in early diagnosis, prognosis, and personalized therapy. *Biomarker research.* 2013;1(1):10.
8. Forner A, Bruix J. Biomarkers for early diagnosis of hepatocellular carcinoma. *The Lancet Oncology.* 2012;13(8):750-751.
9. Varela-Martinez E, Abendano N, Asin J, et al. Molecular Signature of Aluminum Hydroxide Adjuvant in Ovine PBMCs by Integrated mRNA and microRNA Transcriptome Sequencing. *Frontiers in immunology.* 2018;9:2406.
10. Shen Y, Bu L, Li R, Chen Z, Tian F, Ge Q. Expression And Biological Interaction Network Of RHOC For Hepatic Carcinoma With Metastasis In PBMC Samples. *Onco Targets Ther.* 2019;12:9117-9128.
11. Trapnell C, Roberts A, Goff L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols.* 2012;7(3):562-578.
12. Mi H, Muruganujan A, Huang X, et al. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nature protocols.* 2019;14(3):703-721.
13. Maheswaran S, Haber DA. Circulating tumor cells: a window into cancer biology and metastasis. *Current opinion in genetics & development.* 2010;20(1):96-99.
14. Greenhill C. New pathways in development of liver cancer. *Nature reviews Gastroenterology & hepatology.* 2018;15(12):718.

15. Greenhill C. New pathways in development of liver cancer. *Nature reviews Endocrinology.* 2018;15(1):2.
16. Yu QJ, Liang YZ, Mei XP, Fang TY. Tumor mutation burden associated with miRNA-gene interaction outcome mediates the survival of patients with liver hepatocellular carcinoma. *EXCLI journal.* 2020;19:861-871.
17. Yi T, Zhou X, Sang K, Huang X, Zhou J, Ge L. 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.* 2019;47(1):3765-3773.
18. Yu C, Hong H, Zhang S, et al. Identification of key genes and pathways involved in microsatellite instability in colorectal cancer. *Molecular medicine reports.* 2019;19(3):2065-2076.
19. Wadhwa R, Ryu J, Ahn HM, et al. Functional significance of point mutations in stress chaperone mortalin and their relevance to Parkinson disease. *The Journal of biological chemistry.* 2015;290(13):8447-8456.
20. Yang B, Li M, Tang W, et al. Dynamic network biomarker indicates pulmonary metastasis at the tipping point of hepatocellular carcinoma. *Nature communications.* 2018;9(1):678.
21. Bhagwan JR, Mosqueira D, Chairez-Cantu K, et al. Isogenic models of hypertrophic cardiomyopathy unveil differential phenotypes and mechanism-driven therapeutics. *Journal of molecular and cellular cardiology.* 2020;145:43-53.
22. Xiang X, You XM, Li LQ. Expression of HSP90AA1/HSPA8 in hepatocellular carcinoma patients with depression. *OncoTargets and therapy.* 2018;11:3013-3023.
23. Yu L, Yin B, Qu K, et al. Screening for susceptibility genes in hereditary non-polyposis colorectal cancer. *Oncology letters.* 2018;15(6):9413-9419.
24. Dirami T, Rode B, Jollivet M, et al. Missense mutations in SLC26A8, encoding a sperm-specific activator of CFTR, are associated with human asthenozoospermia. *American journal of human genetics.* 2013;92(5):760-766.
25. El Khouri E, Toure A. 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. *The international journal of biochemistry & cell biology.* 2014;52:58-67.
26. Lee YM, Kim S, Park RY, Kim YS. Hepatitis B Virus-X Downregulates Expression of Selenium Binding Protein 1. *Viruses.* 2020;12(5).