Noninvasive, Accurate Diagnosis of Chronic Liver Diseases Using Whole-Transcriptome Profiling of Platelets

Huikun Wu
Hubei Provincial Hospital of Traditional Chinese Medicine
Gang Chen
Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology
Tianyuan Zhang
Wuhan Benagen Tech Solutions Company Limited
Mingzhong Xiao
Hubei Provincial Hospital of Traditional Chinese Medicine
Ye Xia
Wuhan Benagen Tech Solutions Company Limited
Miaojun Han
Wuhan Benagen Tech Solutions Company Limited
Liam C. Lee
Wuhan Benagen Tech Solutions Company Limited
Hao Su
Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology
Xin Qiu
Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology
Ying Wang
Wuhan Benagen Tech Solutions Company Limited
Junxiu Tao
Hubei Provincial Hospital of Traditonal Chinese Medicine
Zuoyu Shao
Hubei Provincial Hospital of Traditional Chinese Medicine
Xudong Zhang
Wuhan Benagen Tech Solutions Company Limited
Min Chen
Wuhan Benagen Tech Solutions Company Limited
Yan Niu
Wuhan Benagen Tech Solutions Company Limited
Chenxia Lu

Hubei University of Chinese Medicine

Hongli Teng

Guangxi International Zhuang Medicine Hospital

Xiaodong Li

Hubei Provincial Hospital of Traditional Chinese Medicine

Chi Song (songchi@benagen.com)

Wuhan Benagen Tech Solusions Company Limited

Research

Keywords: chronic liver disease, chronic liver disease-educated blood platelet, RNA-Seq, support vector machine, liquid biopsy

Posted Date: November 12th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-103725/v1

Abstract

Background: Hepatocellular carcinoma (HCC) is the most serious tumor in the world. It generally undergoes a series of processes from HBV infection, chronic hepatitis, cirrhosis, and HCC from early to late stages. Patients could benefit from early detection of chronic liver diseases (CLD). Tumor-Educated Platelets play an important role in tumor progression, which maybe a potential biomarker for CLD early diagnosis. Here, we developed a noninvasive liquid biopsy technique using platelet RNA for the early screening of patients with liver diseases.

Methods: This study included a total of 163 individuals, including 50 healthy individuals, 39 chronic hepatitis B (CHB) patients, 40 liver cirrhosis (LC) and 34 patients with HCC. Blood was collected before initiation of treatment. Platelet RNA-Seq combined with Support Vector Machine (SVM), was used for the first time to distinguish the different stages of CLD in Asian patients.

Results: Developed diagnostic model could distinguished with 92.4% accuracy between 34 HCC and 50 healthy, 89.92% accuracy between 34 patients HCC and 129 non-cancer individuals, and 83.67% between 50 healthy and 113 CLD. Across four different individual types, the accuracy of distinction (healthy/chronic hepatitis B/liver cirrhosis/hepatocellular carcinoma) was 65.31%. This model was internally validated, resulting in optimism-corrected AUC's of 86.8%.

Conclusions: Our data indicate that the developed platelet RNA-Seq is a valuable platform for the diagnosis of CLD, providing an effective solution for its diagnosis.

Background

Hepatitis B virus (HBV) infection causes asymptomatic chronic hepatitis B (CHB), liver cirrhosis (LC), and hepatocellular carcinoma (HCC), all with high mortality rates; it poses a massive healthcare burden on the community worldwide. HCC is a malignant tumor of hepatocytes and the second leading cause of cancer-related death worldwide. In 2018, an estimated 841,080 new liver cancer cases and 781,631 deaths occurred worldwide, and China is the most high-risk HCC area(1). Globally, HBV infection contributes to approximately 50% of all HCC cases, with the majority (70–80%) of patients with HBV-related HCC having LC(2). In China, HCC is the third leading type of cancer in adults, with the total incidence and mortality of new cases being 466,100 and 422,100, respectively, in 2015(3). Approximately 90% of Chinese patients with HCC report HBV infection, and the general pathogenesis involves the following stages: HBV infection–CHB–LC–HCC. Among these, LC is a crucial factor in the development of HCC, making it an important stage of diagnostic monitoring.Typically, advanced chronic liver diseases (CLD), particularly HCC, are diagnosed by imageological examination, α-fetoprotein (AFP) test, and liver biopsy. In a previous study, in the diagnosis of early-stage HCC, AFP exhibited a sensitivity of 66%, a specificity of 81%, and a cut-off value of 10.9 ng/mL(4). Although liver biopsy is the gold standard diagnostic method to determine the typical characteristics of LC and HCC, invasive strategies are more

harmful to the patients, so it cannot be widely applied to early-detected CLD. Therefore, it is urgent to find advanced non-invasive marker for early detection of CLD.

The use of disease-related circulatory biomarkers is a rapid and effective way of diagnosing any disease. Recent studies have introduced the concept of liquid biopsy(5, 6), which has been shown to satisfactorily detect components including cell-free DNA(7–9), circulating tumor cells(10–12), and extracellular vesicles(13, 14). A recent study explicated that mRNA sequencing of tumor-educated platelets (TEPs) differentiates pan-cancer patients, including those with HCC, from healthy individuals with an accuracy of 96%(15). Platelets can also be used as biomarkers in the diagnosis of sarcomas and non-small cell lung cancer (NSCLC).(16, 17) Platelets are produced from megakaryocytes, which are large bone marrow cells. As vital biologically potent molecules(18), platelets contribute to bacterial infection resistance, growth factor release, inflammation response, and liver regeneration(19) as well as adversely affect cancer metastasis(20). In addition, previous studies have determined diagnostic platelet RNA signatures for cardiovascular abnormalities, inflammatory conditions, sickle cell disease, essential thrombocytosis, and cancer(15, 21, 22), but those for CLD remain poorly reported.

Hence, in this study, we characterized the platelet RNA profiles of patients with different types of CLD and healthy volunteers and subsequently developed the first application technique to diagnose CLD.

Material And Methods

Study cohort

In this study, we recruited 113 patients with CLD, including CHB (n = 39), LC (n = 40), and HCC (n = 34) between December 2016 and February 2017 based on defined criteria. 50 healthy donors. 50 sex- and age-matched healthy control was included as well. The study protocol was approved by the Ethics Committee of Hubei Provincial Hospital of Traditional Chinese Medicine (HBZY2014-C047-01) and Tongji Medical College of Huazhong University of Science & Technology (No. JDZX2015178). The inclusion criterion was adult patients with confirmed diagnosis of HBV-related CLD. However, other factors leading to CLD and its acute complications were excluded. The disease diagnostic standard was based on the guideline of prevention and treatment of CHB(23), management of clinical diagnosis, evaluation, and antiviral therapy for HBV-related cirrhosis(24), and standardization of diagnosis and treatment for HCC(25, 26).

Platelet separation

From all donors, 10 mL whole blood samples were collected in a sodium citrate coagulation test tube (blue cap), which were transported to a laboratory at room temperature without vigorous shaking for subsequent experiments within 2 h. After centrifuging at 150 × g for 15 min, the samples were stratified; the upper layer constituted platelet-rich plasma (PRP) with a straw color. Then, we connected a 5-mL syringe, leukocyte-depleted filter, and 15-mL centrifuge tube sequentially from the top to bottom. Next, PRP was added to the syringe along the tube wall and filtered under gravity. Further, we flushed the tube

wall with a buffer (134 mM NaCl, 12 mM $NaHCO_3$, 2.9 mM KCl, 0.34 mM Na_2HPO_4, 1 mM $MgCl_2$, and 10 mM HEPES; pH 7.4). PRP and buffer were mixed by repeatedly reversing the collection tube. Then, 100 µL of the mixed solution was collected for platelet and white blood cell counts using a blood cell analyzer. When the white blood cell count was 0, the sample was determined to be qualified. Unqualified samples were not used for subsequent experiments. Finally, the qualified mixed solution was centrifuged at room temperature for 10 min at 2000 × g, and the supernatant was removed.

Platelet RNA extraction and preparation

The collected precipitate was transferred to ice, and total RNA was isolated using TRIzol® Reagent (Invitrogen) according to the manufacturer's protocol. Total RNA was dissolved with 30 µL RNase-free water. The solution was stored directly at − 80 °C.

We monitored RNA degradation and contamination on 1% agarose gels and checked RNA purity using the NanoPhotometer® spectrophotometer (Implen, CA, USA). RNA concentration was measured using Qubit® RNA Assay Kit in Qubit® 3.0 Flurometer (Life Technologies, CA, USA). In addition, RNA integrity was evaluated using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA). Subsequently, 20–100 ng platelet total RNA was used for sequencing libraries generated with NEBNext® UltraTMRNA Library Prep Kit for Illumina® (NEB, USA) according to the manufacturer's instructions. We then sequenced library preparations on an Illumina HiSeq X-ten platform and generated 150-bp paired-end reads.

Processing of raw mRNA sequencing data to read count matrix

For each sample, the obtained sequencing reads were cleaned by 5'-end quality trimming and clipping of the sequencing adapters by SOPAnuke (http://soap.genomics.org.cn/). Then, we performed prealignment quality control of the cleaned sequencing reads using FastQC (v 0.11.5) (27). Using STAR (version 2.5.0a) (28), we aligned the RNA-Seq reads against the University of California Santa Cruz hg19 genome sequences (http://www.genome.ucsc.edu/index.html) under default parameters, allowing one or two base mismatches.

Furthermore, read summarization of only reads spanning introns (intron spanning) was performed with HTSeq (version 0.6.0, https://htseq.readthedocs.io/). We included protein-coding and noncoding RNAs during read mapping, summarization, and subsequent analyses. Notably, we excluded the genes encoded on mitochondrial DNA and the Y chromosome from the analyses. In addition, the genes not expressed in 5% of the samples were excluded. Moreover, gene expression clustering analysis using k-mean clustering, principal component analysis, and subsequent statistical analyses were performed using R (version 3.0.3) and R-studio (version 0.98.1091). The accession number for the raw sequencing data reported in this paper is GEO: GSExxxxx.

Gene functional classification and differential expression of genes

For functional annotation, we assessed ensemble gene IDs for enrichment in Gene Ontology (GO) databases (www.geneontology.org/). RNA profile was normalized using counts per million. Using the edgeR package(29), we performed differential gene expression analysis for the paired samples. False discovery rate (FDR) < 0.001 was used to identify the significance of differentially expressed genes (DEGs). Subsequently, we performed GO functional enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of DEGs using ClusterProfiler(30).

Support vector machine (SVM) classifier

In this study, we used the SVM classifier for 50 healthy samples, 39 CHB samples, 40 LC samples, and 34 HCC samples. We used this method for comparison-based diagnosis: (1) between patients with cancer (n = 34) and healthy donors (n = 50); (2) between patients with cancer (n = 34) and non-cancer individuals (n = 129); (3) between healthy donors (n = 50) and unhealthy individuals (n = 113); and (4) across four different types (Healthy/CHB/LC/HCC). Overall, 70% of samples were selected as a training set, and the selected DEGs were used for SVM classification. Each procedure was randomly repeated 100 times to ascertain the accuracy of the method. Furthermore, we used predictive strength as an input to generate a receiver-operating curve (ROC) as implemented in ggplot2 (version 1.7.3).

Marker selection and verification

We identified DEGs, which occurred simultaneously at HBV–LC, LC–HCC, and HBV–HCC stages, and then, we used the SVM classifier to distinguish the three types of CLD based on these DEGs, determining the accuracy of distinguishing the different types of CLD.

Results

Transcriptome profiling of chronic liver disease-educated platelets (CLDEPs) reveals unique mRNA signatures in affected patients compared with those in healthy volunteers

Peripheral blood samples and isolated circulating platelets from healthy donors (n = 50), patients with CHB (n = 39), patients with LC (n = 40), and patients with HCC (n = 34) were collected, all diagnosed based on clinical presentations and conventional pathological analysis of liver tissues (Fig. 1A-B). 16,968 genes were identified for subsequent analyses. Of 16,968 genes, 4700 exhibited difference in terms of expression levels among the four groups (FDR < 0.001; Table 2, Fig. 1C). The GO functional enrichment analysis revealed that transcriptome data enriched for transcripts correlated with platelet functions (false discovery rate [FDR] < 0.05; Table S1; Fig. 1D). Furthermore, the KEGG enrichment analysis of the transcriptome data determined endocytosis as the most enriched signatures (FDR < 0.05, Table S2, Fig. 1E).

Table 1
Summary of study participants

Group	Total (n)	Training	Validation	Male	Female	Age (SD)
Healthy	50	35	15	22	28	34(10)
HBV	39	27	12	29	10	39(13)
LC	40	28	12	35	4	44(9)
HCC	34	24	10	29	4	52(11)

CHB, chronic hepatitis B; LC, liver cirrhosis; HCC, hepatocellular carcinoma.

Table 2
Differentially expressed genes among the study groups

	Healthy	CHB	LC	HCC
Healthy (n = 50)	0	144	3582	2676
CHB (n = 39)		0	103	530
LC (n = 40)			0	362
HCC (n = 34)				0

CHB, chronic hepatitis B; LC, liver cirrhosis; HCC, hepatocellular carcinoma.

Signature identification and diagnosis between HCC and healthy or non-HCC group

Comparison of the transcriptome profile of patients with HCC with that of healthy volunteers revealed 2,676 DEGs, whereas that of patients with HCC and CHB revealed 530 DEGs and those of patients with HCC and LC revealed 362 DEGs. Remarkably, the number of DEGs between HCC and other groups declined with disease progression, suggesting the underlying biological rationale of platelet transcriptome profiling as a surrogate marker for CLD (Table 2). GO and KEGG enrichment analyses between the two groups identified specific pathways and functional groups (Fig. 2A-B; Table S1–S2).

For HCC diagnosis-specific SVM algorithm, we selected DEGs from the healthy and HCC groups that were incorporated in the training cohort (n = 59), yielding the ideal sensitivity, specificity, and accuracy for tests performed within the training cohort (100%; Fig. 3A). Subsequent validation (n = 25) of the DEG-trained SVM algorithm yielded 90% sensitivity, 93% specificity, and 92% accuracy with an area under the curve (AUC) of 0.993, illustrating high predictive strength of the algorithm in correctly differentiating patients with HCC from healthy donors (Fig. 3A). Unfortunately, one sample of the validation cohort was misdiagnosed (6.6%). A total of 100 random class-proportional subsampling processes of the entire transcriptome profile data, combining training and validation cohorts, produced similar accuracy rates

with a mean overall accuracy of 92.4% (SD: ±4.95%), establishing reproducible classification algorithm within this dataset.

In addition, we used the SVM classifier for cancer and non-cancer classification. In the training cohort (n = 114), an accuracy of 100% and AUC of 0.99 were obtained. The subsequent validation cohort (n = 49) yielded a sensitivity of 80% and specificity of 92.3%, with localized disease correctly classified in 44/49 patients (89.80%), and an AUC of 0.94 to detect the disease and a high predictive strength. A total of 100 random proportional subsampling processes of the entire dataset in a training and validation set (ratio: 70:30) yielded similar accuracy rates (mean overall accuracy: 89.92% ± 3.46%), confirming the reproducible classification accuracy in this dataset (Fig. 3B).

Signature identification and diagnosis between CLD and Heathy group

248 DEGs were identified between CLD (CHB, LC, and HCC; all collectively considered CLD) and healthy honor. The most significant biological process among the genes was Ribosome (25 genes), with adjusted P value = 1.11E-24, and blood microparticle (33 genes), with adjusted P value = 3.97E-24. In the KEGG pathway analysis, focal adhesion, was most significant pathways, with 16 genes enriched (adjusted P value = 0.001) (Fig. 2C-H).

Next, we established the diagnostic accuracy of the CLDEP-based broad classification algorithm, whereby each participant would be diagnosed with a CLD (CHB, LC, and HCC; all collectively considered CLD) or classified as a healthy donor. For the SVM algorithm for the training cohort (n = 114), we optimized it to again yield the ideal sensitivity, specificity, and accuracy (100%). The subsequent validation (n = 49) of this SVM algorithm yielded 73% sensitivity, 88% specificity, and 84% accuracy with an AUC of 0.900, illustrating high predictive strength in correctly differentiating patients with a liver disease from healthy volunteers (Fig. 3C); however, the model's misdiagnosis rate was 11.8%. After a total of 100 random repetitions, it produced similar accuracy rates with a mean overall accuracy of 83.2% ± 4.30%, establishing a reproducible classification algorithm within this dataset and indicating that CLDEPs is a potential surrogate marker for screening CLD.

Diagnostic signature in four group

The same dataset is used to provide an all-in-one biosource for blood-based liquid biopsies in patients with CLD. All samples were categorized into four groups. The training set demonstrated an excellent distinction of patients. In addition, the classification capacity of the multiclass SVM-based classifier was established in the validation cohort of 49 samples. In this classifier, the sensitivity of a healthy donor was 73.33%; the probability of patients with CHB, LC, and HCC being correctly diagnosed was 58.33%, 58.33%, and 70%, respectively, with an AUC of 0.8588. The multiclass CLD diagnostic test resulted in an average accuracy of 65.31% (mean overall accuracy random classifiers: 66.35% ± 6.19%; $P < 0.01$), demonstrating the significant discriminative power of the multiclass CLD diagnostic test of platelet mRNA profiles (Fig. 3D).

Notably, 14 DEGs occurred in all pairs of CLD groups (TGM2, EPAS1, HAPA12B, H19, DOCK6, CARF10, KANK3, CASKIN2, RELN, IGFBP4, SLC9A3R2, LIMS2, PPFIBP1, and A2M; Fig. 3E). Then, we constructed an SVM/LOOCV discriminator algorithm based on 14 DEGs in all CHB, LC, and HCC cases to test the feasibility of a targeted panel-based diagnostic assay for liver diseases. The small targeted panel NGS assay depicts an attractive alternative to comprehensive assays (e.g., whole-genome and whole-transcriptome sequencing) in most clinical laboratory settings owing to their increased throughput (i.e., more patient samples per flow cell) and enhanced cost-effectiveness, especially for disease-specific assays such as the CLDEP diagnostic assay used herein. In the training cohort (n = 80), this simplified approach yielded sensitivities of 92% for patients with HCC, 81% for patients with CHB, and 75% for patients with LC (Fig. 3F). In addition, subsequent tests of the validation cohort yielded sensitivities of 80% for patients with HCC, 92% for patients with CHB, and 42% for patients with LC. For a total of 100 random class-proportional subsampling processes, the overall accuracy rate average was 68.2% ± 6.89%.

Discussion

The study showed that platelets can be used as an all-in-one biosource to broadly scan the molecular traces of CLD and offer a strong indication on disease type and molecular subclass. Recently, platelet RNA was used as a biomarker for diagnostic diseases. RNA-Seq gene expression characterization offers a detailed, unbiased overview of the platelet RNA content. In this study, we reported the molecular integration of blood platelets, which are involved in almost every stage of CLD. We recruited HCC patients and CLD patients at high risk of developing HCC. The total RNA of peripheral blood of 163 patients was purified using standard procedures and sequenced using high-throughput. After the DEGs were screened, the diagnosis was made by SVM method, and a good result was obtained to distinguish between HCC, CLD and Healthy. The AUC reached 0.993, 0.99 and 0.900 in the diagnosis of the respective samples, and the AUC reached 0.8588 in all samples. Additonally, we reported the molecular integration of blood platelets, which are involved in almost every stage of CLD. Perhaps, mRNA will offer valuable diagnostic information about all patients with CLD.

Blood platelets are widely involved in disease and cancer progression. In this study, we found enriched function, including platelet degranulation, activation, aggregation, and focal adhesion. Also, we observed a significant number of genes associated with Endocytosis, Cell cycle,Protein processing in endoplasmic reticulum, Apoptosis, cytoskeleton, and Focal adhesion pathway, which may refelcting the “alert” and pro-tuorigenic stats of CLD. A crucial mechanism in the development and progression of LC is hepatocyte death induced directly by hepatocyte damage or triggered by immune responses and activation of Kupffer cells and the activation of hepatic stellate cells and resistance of these cells to apoptosis(31). All these processes might sustain further inflammation and fibrosis development, leading to LC. The gene enrichment analysis in the patients at CHB–LC stages correlated with keratin filaments, structural constituent of the cytoskeleton, epidermal development, and intermediate filament. Keratin is the largest member of the intermediate filament protein family. Some animal model experiments have established that keratin protects hepatocytes from apoptosis and necrosis; natural human keratin variants are generally silent but become pathogenic by predisposing their carriers to apoptosis during acute or chronic

toxin-mediated liver injury, viral infection, or metabolic stress(32). In addition, K8/K18 is considered as a biomarker in CLD(33). In this study, platelet RNA detection also demonstrated that K4, K5, K6, K14, K15, and K16 are markedly expressed in the progression of healthy CHB–LC, possibly enlightening the subsequent research on the mechanism of liver fibrosis. The DEGs at the LC–HCC stage are enriched in a metabolism-related pathway. Most clinically used drugs are metabolized by cytochrome P450, which has been best characterized in the liver(34) and plays a major role in the metabolism of several chemical carcinogens involved in the development of HCC(35). A previous study reported a substantial change in cytochrome P450 activities in patients with HCC(36). In addition, CYP3A7 is overexpressed in HCC where it contributes to the elimination of drugs (37).

Interestingly, 14 genes were differentially expressed in patients with different CLD. TGM2 is involved in the key bioprocess of cell apoptosis, cell transmembrane signaling, cell adhesion, and extracellular matrix formation. TGM2 is overexpressed in liver fibrosis, and its deletion in mouse resulted in collagen crosslinking but could not contribute to inhibiting liver fibrosis(38). In addition, TGM2 is positively correlated with tumor diameter, vascular invasion, and TNM staging(39), suggesting its crucial role in the development, progression, and metastasis of HCC. HSPA12B is a heat-shock protein family A (Hsp70) member 12B. HSPs play a pivotal role in regulating apoptosis, therapeutic resistance, and invasion and metastasis in HCC(40). In fact, HSP70 is the most abundantly upregulated gene in early HCC components(41). The combined use of glutamine synthetase and HSP70 could be useful in the diagnosis of well-differentiated HCC(42). Reportedly, HIF-2α/EPAS1 exerts a substantial impact on cell proliferation, tumor angiogenesis, metastasis, and resistance to chemotherapy and radiation(43, 44). Overexpression of HIF-2α induced apoptosis in HCC cells and increased the levels of proapoptotic proteins, Bak, ZBP-89, and PDCD4, which was also confirmed in another study(45). In addition, HIF-2α acts as a safeguard to initiate sinusoidal reconstruction only upon successful hepatocyte mitosis, thereby enforcing a timely order onto cell type-specific regeneration patterns. These findings suggest the hypoxia-driven HIF-2α–VEGF axis to be a prime node in coordinating SEC–hepatocyte crosstalk during liver regeneration(46). The DOCK6 is involved in small GTPase-mediated signal transduction and blood coagulation-positive regulation of GTPase activity. However, limited reports exist on DOCK6 in liver diseases.

Conclusions

In conclusion, this study provides robust evidence on the clinical relevance of blood platelets for liquid biopsy-based molecular diagnostics in patients with several types of CLD. Nevertheless, further validation is warranted to establish the potential of surrogate CLDEP profiles for blood-based companion diagnostics, therapy selection, longitudinal monitoring, and disease recurrence monitoring.

Abbreviations

AFP, α-fetoprotein; AUC, area under the curve; CHB, chronic hepatitis B; GO, Gene Ontology; HBV, hepatitis B virus; CLD, chronic liver disease; KEGG, Kyoto Encyclopedia of Genes and Genomes; LC, liver cirrhosis;

HCC, hepatocellular carcinoma; CLDEPs, chronic liver disease-educated blood platelets; ROC, receiver-operating curve; SVM, support vector machine; TEP, tumor-educated platelets

Declarations

Authors' Contributions

Conceptualization: C.S. and X.L;data curation: H.W, G.C., M.X., H.S., X.Q., J.T., Z.S., C.L., H.T. and X.L.;formal analysis: H.W, G.C., M.H., L.C.L., H.S., X.Q., J.T., Z.S., C.L., H.T. and X.L.;funding acquisition: G.C., M.X. and X.L.;investigation: H.W., G.C., H.S., X.Q., J.T., C.L., H.T. and X.L.; methodology: H.W., G.C., T.Z., M.X., Y.X., Y.W., C.L., H.T., X.L. and C.S.; project administration: C.S. and X.L;resources: C.S. and X.L;Software: T.Z., Y.X., Y.W., X.Z., M.C. and Y.N.;supervision: X.L. and C.S.;validation: H.W, G.C., M.X., H.S., X.Q., J.T., Z.S., C.L., H.T. and X.L;Visualization:Y.W., X.Z., M.C., Y.N. and Y.N; writing—original draft preparation: H.W., G.C., T.Z., M.X., Y.X.;writing—review and editing: H.W., G.C., T.Z., M.X., Y.X., M.H., L.C.L., X.L. and C.S. All co-authors have read and edited the manuscript.

Funding

This study was supported by (1) the Major Special Project of Ministry of Science and Technology supported by China's "Twelfth Five—Year Plan (NO.2014ZX10005001), (2)
The Key Research Project of National Chinese Medicine Research Base (Hubei),
(3) National Natural Science Foundation Project (No.81874423), and (4) Special Project for Business Construction of National Chinese Medicine Clinical Research Base (No. JDZX2015178).

Availability of data and materials

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

Ethics approval and consent to participate

The study protocol was approved by the Ethics Committee of Hubei Provincial Hospital of Traditional Chinese Medicine (HBZY2014-C047-01) and Tongji Medical College of Huazhong University of Science & Technology (No. JDZX2015178). This study conformed to the guidelines proposed in the Helsinki Convention. After a full explanation of the entire study process, each subject signed an informed consent form before inclusion in the study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2018;68(6):394-424.
2. El-Serag HB. CURRENT CONCEPTS Hepatocellular Carcinoma. New England Journal of Medicine. 2011;365(12):1118-27.
3. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. Ca A Cancer Journal for Clinicians. 2016;66(2):115.
4. Marrero, A J, Feng, Ziding, Wang, Yinghui, et al. α-Fetoprotein, Des-γ Carboxyprothrombin, and Lectin-Bound α-Fetoprotein in Early Hepatocellular Carcinoma. Gastroenterology. 2009;137(1):110-8.
5. Joosse SA, Pantel K. Tumor-Educated Platelets as Liquid Biopsy in Cancer Patients. Cancer Cell. 2015;28(5):552.
6. Siravegna G, Marsoni S, Siena S, Bardelli A. Integrating liquid biopsies into the management of cancer. Nature Reviews Clinical Oncology. 2017;14(9):531.
7. Burnham P, Khush K, De VI. Myriad Applications of Circulating Cell-Free DNA in Precision Organ Transplant Monitoring. Annals of the American Thoracic Society. 2017;14(Supplement_3):S237.
8. Adams NR, Vasquez YM, Mo Q, Gibbons W, Kovanci E, DeMayo FJ. WNK lysine deficient protein kinase 1 regulates human endometrial stromal cell decidualization, proliferation, and migration in part through mitogen-activated protein kinase 7. Biology of reproduction. 2017;97(3):400-12.
9. Liu MC, Oxnard GR, Klein EA, Swanton C, Oncology DABJAo. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. 2020.
10. Wan JC, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. Nature Reviews Cancer. 2017;17(4):223.
11. McDonald BR, Contente-Cuomo T, Sammut S-J, Odenheimer-Bergman A, medicine MMJet. Personalized circulating tumor DNA analysis to detect residual disease after neoadjuvant therapy in breast cancer. 2019;11(504):eaax7392.
12. Sun, Meng, Tang, Wang, Guo, Ma, et al. Design of a liver cancerspecific selector for the analysis of circulating tumor DNA. 2019.
13. O'Driscoll L. Expanding on exosomes and ectosomes in cancer. New England Journal of Medicine. 2015;372(24):2359.
14. Melo, Sonia, nbsp, Sugimoto, Hikaru, O'Connell, et al. Cancer Exosomes Perform Cell-Independent MicroRNA Biogenesis and Promote Tumorigenesis. Cancer Cell. 2014;26(5):707.
15. Best MG, Sol N, Kooi I, Tannous J, Westerman BA, Rustenburg F, et al. RNA-Seq of Tumor-Educated Platelets Enables Blood-Based Pan-Cancer, Multiclass, and Molecular Pathway Cancer Diagnostics. Cancer Cell. 2015;28(5):666.
16. Zhang Q, Hu H, Liu H, Jin J, Zhu P, Wang S, et al. RNA sequencing enables systematic identification of platelet transcriptomic alterations in NSCLC patients. Biomedicine & Pharmacotherapy.

2018;105:204-14.
17. Heinhuis K, Veld SIt, Dwarshuis G, Broek Dvd, Oncology WvHJEJoS. RNA-Sequencing of Tumor-Educated Platelets, A Novel Biomarker for Blood Based Sarcoma Diagnostics. 2020;46(2):e7.
18. Leslie M. Beyond Clotting: The Powers of Platelets. Science. 2010;328(5978):562-4.
19. T L, RJ P. The role of platelets in liver inflammation and regeneration. Seminars in thrombosis and hemostasis. 2010;36(2):170.
20. Labelle M, Begum S, Hynes RO. Platelets guide the formation of early metastatic niches. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(30):3053-61.
21. Best MG, Sol N, Veld SGJGIt, Vancura A, Muller M, Niemeijer ALN, et al. Swarm Intelligence-Enhanced Detection of Non-Small-Cell Lung Cancer Using Tumor-Educated Platelets. Cancer Cell. 2017;32(2):238.
22. Best MG, Vancura A, Wurdinger T. Platelet RNA as a circulating biomarker trove for cancer diagnostics. Journal of Thrombosis & Haemostasis. 2017;15(7):1295.
23. Chinese Society of Hepatology CMA, Hou JL, Lai W. [The guideline of prevention and treatment for chronic hepatitis B: a 2015 update]. Zhonghua gan zang bing za zhi = Zhonghua ganzangbing zazhi = Chinese journal of hepatology. 2015;23(12):888.
24. Management of clinical diagnosis, evaluation, and antiviral therapy for HBV - related cirrhosis2014.
25. Wang J, Dong W, Liu FC, Guo XG, Liu H. Interpretation of 2017 updated consensus on Standardization of Diagnosis and Treatment for Hepatocellular Carcinoma. Academic Journal of Second Military Medical University. 2017.
26. Zhou J, Sun HC, Wang Z, Cong WM, Wang JH, Zeng MS, et al. Guidelines for Diagnosis and Treatment of Primary Liver Cancer in China (2017 Edition). Liver Cancer. 2018:1-26.
27. Andrews S. FastQC A Quality Control tool for High Throughput Sequence Data. 2015.
28. Valencia A. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2014.
29. Robinson M, Mccarthy D, Smyth GK. edgeR: differential expression analysis of digital gene expression data. Journal of Hospice & Palliative Nursing. 2010;4(4):206-7.
30. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. Omics A Journal of Integrative Biology. 2012;16(5):284.
31. Guicciardi ME, Gores GJ. Apoptosis as a Mechanism for Liver Disease Progression. Seminars in Liver Disease. 2010;30(4):402-10.
32. Ku NO, Strnad P, Bantel H, Omary MB. Keratins: Biomarkers and modulators of apoptotic and necrotic cell death in the liver. Hepatology. 2016;64(3).
33. Waidmann O, Brunner F, Herrmann E, Zeuzem S, Piiper A, † BK. Cytokeratin 18-based cell death markers indicate severity of liver disease and prognosis of cirrhotic patients. Liver International. 2016;36(10).
34. Watkins PB. Drug metabolism by cytochromes P450 in the liver and small bowel. Gastroenterology Clinics of North America. 1992;21(3):511-26.

35. Zhang YJ, Chen S, Tsai WY, Ahsan H, Lunn RM, Wang L, et al. Expression of cytochrome P450 1A1/2 and 3A4 in liver tissues of hepatocellular carcinoma cases and controls from Taiwan and their relationship to hepatitis B virus and aflatoxin B1-and 4-aminobiphenyl-DNA adducts. Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals. 2000;5(4):295.
36. Zhou J, Wen Q, Li SF, Zhang YF, Gao N, Tian X, et al. Significant change of cytochrome P450s activities in patients with hepatocellular carcinoma. Oncotarget. 2016;7(31):50612-23.
37. Neunzig I, Drăgan CA, Widjaja M, Schwaninger AE, Peters FT, Maurer HH, et al. Whole-cell biotransformation assay for investigation of the human drug metabolizing enzyme CYP3A7. Biochimica Et Biophysica Acta. 2011;1814(1):161-7.
38. Popov Y, Sverdlov DY, Sharma AK, Bhaskar KR, Li S, Freitag TL, et al. Tissue transglutaminase does not affect fibrotic matrix stability or regression of liver fibrosis in mice. Gastroenterology. 2011;140(5):1642.
39. Olsen KC, Sapinoro R, Kottmann RM, Kulkarni AA, Iismaa SE, Johnson GV, et al. Transglutaminase 2 and its role in pulmonary fibrosis. American Journal of Respiratory and Critical Care Medicine. 2011;184(6):699-707.
40. Wang C, Zhang Y, Guo K, Wang N, Jin H, Liu Y, et al. Heat shock proteins in hepatocellular carcinoma: Molecular mechanism and therapeutic potential. International Journal of Cancer. 2016;138(8):1824.
41. Chuma M, Sakamoto M, Yamazaki K, Ohta T, Ohki M, Asaka M, et al. Expression profiling in multistage hepatocarcinogenesis: Identification of HSP70 as a molecular marker of early hepatocellular carcinoma. Hepatology. 2003;37(1):198.
42. Nguyen T, Roncalli M, Tommaso LD, Kakar S. Combined use of heat-shock protein 70 and glutamine synthetase is useful in the distinction of typical hepatocellular adenoma from atypical hepatocellular neoplasms and well-differentiated hepatocellular carcinoma. Modern Pathology. 2016;29(3):283-92.
43. Yang Y, Sun M, Wang L, Jiao B. HIFs, angiogenesis, and cancer. Journal of Cellular Biochemistry. 2013;114(5):967-74.
44. Semenza GL. Hypoxia-inducible factors: mediators of cancer progression and targets for cancer therapy. Trends in Pharmacological Sciences. 2012;33(4):207.
45. Yang SL, Liu LP, Niu L, Sun YF, Yang XR, Fan J, et al. Downregulation and pro-apoptotic effect of hypoxia-inducible factor 2 alpha in hepatocellular carcinoma. Oncotarget. 2016;7(23):34571.
46. Kron P, Linecker M, Limani P, Schlegel A, Kambakamba P, Lehn JM, et al. Hypoxia-driven Hif2a coordinates mouse liver regeneration by coupling parenchymal growth to vascular expansion. Hepatology. 2016;64(6):2198.

Figures

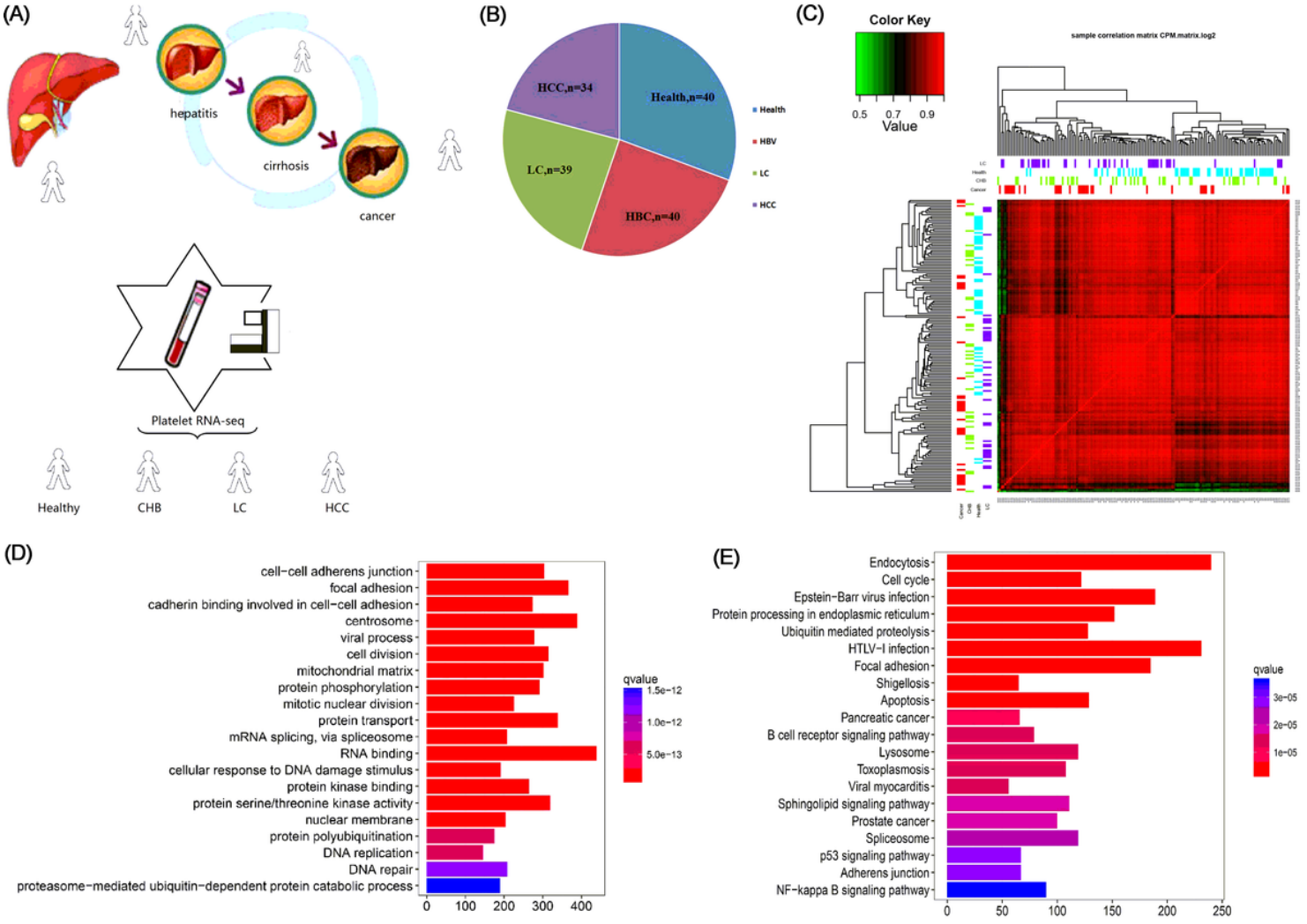

Figure 1

The information of CLDEP mRNA profiling. (A) Schematic overview of CLDEP as a biosource for liquid biopsies. (B) Number of platelet samples of healthy donors and patients with different types of CLD. (C) Functional enrichment maps of differentially expressed genes by GO. (D) Functional enrichment maps of differentially expressed genes by KEGG. (E) Heatmaps of unsupervised clustering of platelet mRNA profiles of healthy donors (n = 50) and patients with CHB (n = 30), patients with LC (n = 40), and patients with HCC (n = 34). CHB, chronic hepatitis B; LC, liver cirrhosis; HCC, hepatocellular carcinoma.

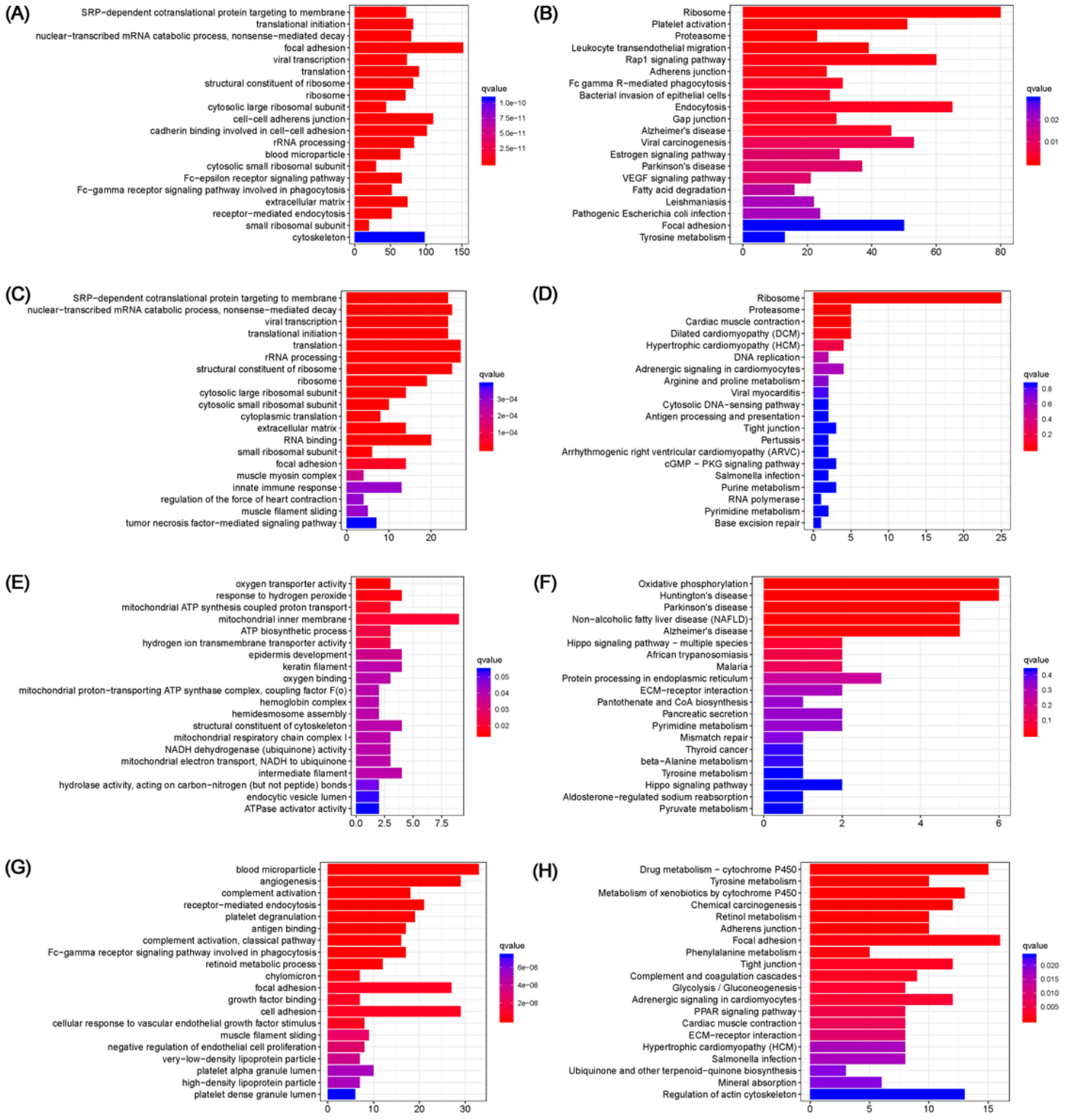

Figure 2

GO and KEGG enrichment in (A-B) healthy vs. HCC; (C-D) healthy vs. CHB; (E-F) CHD vs. LC; (G-H) LC vs. HCC. CHB, chronic hepatitis B; LC, liver cirrhosis; HCC, hepatocellular carcinoma.

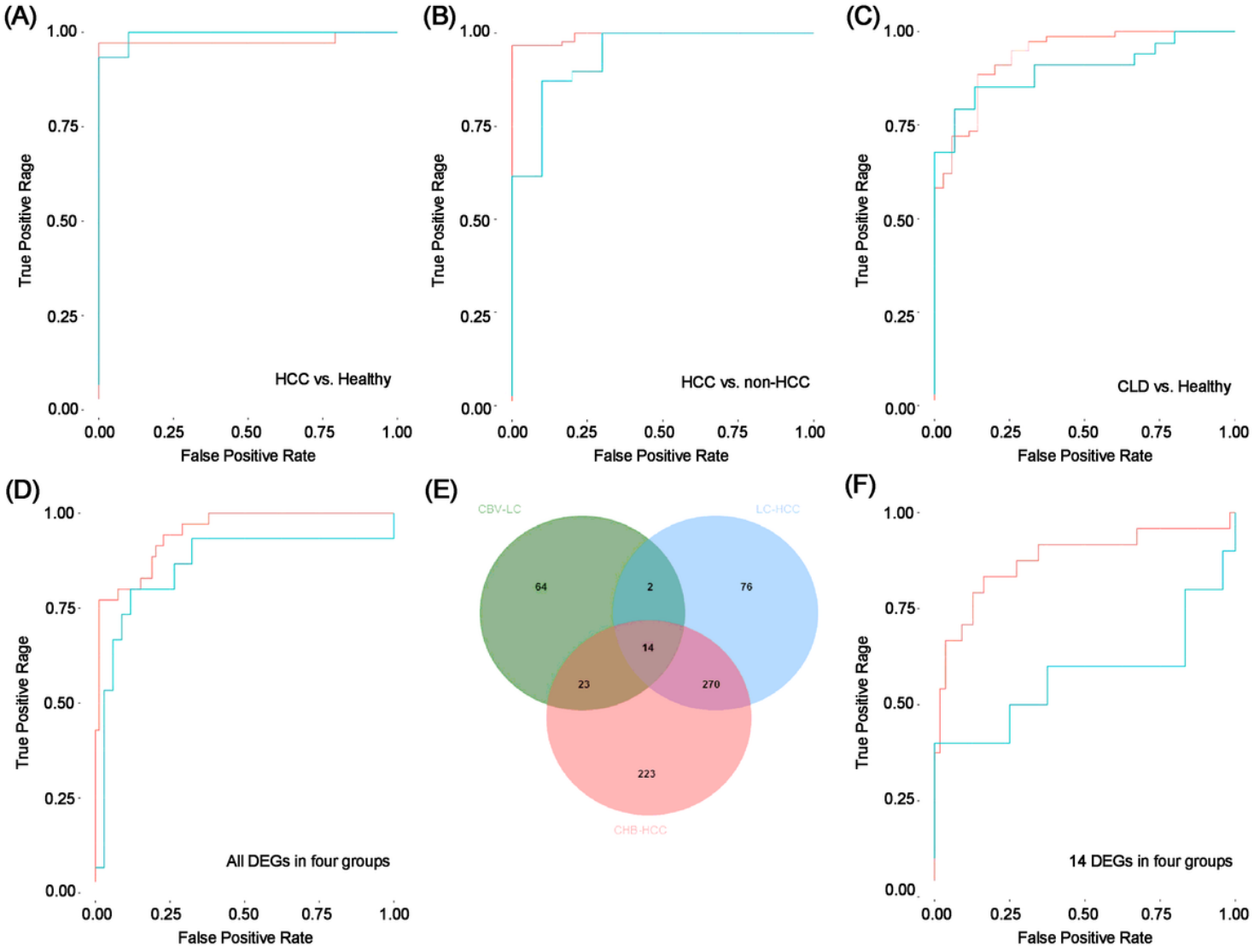

Figure 3

CLDEP mRNA profiles for diagnostics. (A) The receiver-operating curve (ROC) curve of SVM diagnostics of training (red line) and validation (green line) cohort classifiers between healthy and HCC. (B) The ROC curve of the SVM diagnostics of the training and validation cohorts between HCC and non-HCC. (C) ROC curve of healthy/unhealthy SVM diagnostics of training (red line) and validation (green line) cohort classifiers between healthy and CLD. (D) ROC curve of healthy/unhealthy SVM diagnostics of training (red line) and validation (green line) cohort classifiers among four groups. (E) The Venn diagram of differentially expressed genes by CLD. (F) The ROC curve of multiclass SVM diagnostics of the training (red line) and validation (green line) cohorts using the intersection gene. CHB, chronic hepatitis B; LC, liver cirrhosis; HCC, hepatocellular carcinoma; CLD, chronic liver diseases.