Proteomics of small extracellular vesicle from human plasma for hepatocellular carcinoma

Wei Dong
Xiangya Hospital, Central South University, Changsha
Zeyu Xia
Guangxi Key Laboratory of Molecular Medicine in Liver Injury and Repair, the Affiliated Hospital of Guilin Medical University
Zehua Chai
Guangxi Key Laboratory of Molecular Medicine in Liver Injury and Repair, the Affiliated Hospital of Guilin Medical University
Zhidong Qiu
Guangxi Health Commission Key Laboratory of Basic Research in Sphingolipid Metabolism Related Diseases, the Affiliated Hospital of Guilin Medical University
Xuehong Wang
Guangxi Health Commission Key Laboratory of Basic Research in Sphingolipid Metabolism Related Diseases, the Affiliated Hospital of Guilin Medical University
Zebin Yang
China-USA Lipids in Health and Disease Research Center, Guilin Medical University
Junnan Wang
China-USA Lipids in Health and Disease Research Center, Guilin Medical University
Tingrui Zhang
Guangxi Key Laboratory of Molecular Medicine in Liver Injury and Repair, the Affiliated Hospital of Guilin Medical University
Qinqin Zhang
Department of Thyroid and Breast Surgery, Nanxishan Hospital of Guangxi Zhuang Autonomous Region
Junfei Jin (junfeijin@glmc.edu.cn)
Guangxi Key Laboratory of Molecular Medicine in Liver Injury and Repair, the Affiliated Hospital of Guilin Medical University

---

Research Article

Keywords: Proteomics, extracellular vesicles, hepatocellular carcinoma, complement

Posted Date: May 24th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1652245/v1

Abstract

Purpose

Liver cancer is one of the most common tumors with the seventh-highest incidence and the third-highest mortality. Many studies have shown that small extracellular vesicles (sEVs) play an important role in liver cancer. Here, we report comprehensive signatures for sEV proteins from hepatocellular carcinoma (HCC) patient plasma, which might be valuable for the evaluation and diagnosis of HCC.

Methods

We extracted sEVs from the plasma of controls and HCC patients. Differentially expressed proteins in the sEVs were analyzed via label-free quantification and bioinformatic analysis. Western blotting (WB) was used to validate the abovementioned sEV proteins that differentially accumulated in the validation cohorts.

Results

Protein expression profiles were performed for plasma sEVs from 21 HCC patients and 15 controls. Among 335 identified proteins in our study, 27 were significantly dysregulated, including 13 increased proteins that were involved predominantly in the complement cascade (C1QB, C1QC, C4BPA, and C4BPB) and the coagulation cascade (F13B, FGA, FGB, and FGG). We verified that the protein levels of C1QB, C1QC, C4BPA, and C4BPB were increased in the plasma sEVs of HCC patients in both the discovery cohort and validation cohort.

Conclusion

The complement cascade in sEVs was significantly involved in HCC progression. C1QB, C1QC, C4BPA, and C4BPB were highly abundant in sEVs of HCC patients' plasma, which might be potential molecular signatures.

Introduction

Liver cancer is one of the most common tumors, as liver cancer has the seventh-highest incidence and the third-highest mortality according to the IARC Global Cancer Observatory 2018(1). For liver cancer, there were more than 0.8 million new cases in 2018, of which 0.4 million were f rom China, and 0.78 million deaths, of which 0.39 million were from China. The Global Burden of Disease Study 2019 revealed that China had the highest burden of liver cancer in the world (2). Preliminary screening and liquid biopsy of HCC have relied on serum alpha-fetoprotein (AFP). However, due to the low sensitivity of

AFP, it is frequently normal in HCC and limited to clinical diagnosis, which leads to the delay of early treatment(3). Thus, new liquid biopsy markers for detection and prognosis are urgently needed.

The small extracellular vesicle (sEV) is a single-bilayer vesicle with a diameter ranging from 30 to 200 nm that originally buds at the plasma and endosomal membranes(4, 5). The sEV acts as not only ‘garbage bags’ but also novel forms of intercellular communication to regulate receiver cells(6, 7). Although body fluids contain a diversity of sEVs from multiple cell types, there are abundant and specific compounds, such as lipids, nucleic acids, and proteins, closely related to physiological diseases(3). As a common clinical sample, body fluids are tested for progression, diagnosis, and therapy assessment; however, many potential molecules can be hard to enrich and preserve in body fluids, such as plasma, sera, urine, and saliva. Owing to their microstructure, sEVs are stable in body fluids to protect the contents from degradation or destruction. Moreover, the sEVs could be enriched for trace quantity analysis. Therefore, sEVs could be very attractive for liquid biopsy(8).

In recent years, there has been much evidence suggesting that sEVs play a significant role in chronic hepatitis, cirrhosis, and liver cancer(9–11). The sEV from HCC transport LOXL4 and activate the FAK/Src pathway to facilitate the migration of HCC(12). In addition, sEVs from HCC cells deliver LOXL4 to HUVECs to promote angiogenesis by a paracrine mechanism. The sEVs originating from the plasma of HCC patients contain abundant SMAD3. Moreover, the SMAD3 content in sEVs is positively correlated with the stage of HCC and the expression of primary foci(13).

Owing to the close relationship between sEVs and HCC, we aimed to explore potential biomarkers from plasma sEVs for the evaluation of HCC. In this study, by comparing the protein expression among the groups of control and HCC patients, we comprehensively identified variable proteins from sEVs by the label-free quantification method. Then, we verified the diagnostic candidates for HCC in the independent validation cohort by western blotting.

Materials And Methods

Plasma sample collection

Plasma samples of controls and HCC patients were collected at the Affiliated Hospital of Guilin Medical University (Guangxi, China) after written informed consent was provided for patients and their families, according to protocols (#YJSL2021129) by the Institutional Review Board of the Affiliated Hospital of Guilin Medical University (Guangxi, China). All blood samples from 28 patients whose confirmatory pathological diagnosis was HCC and 15 control patients without liver cancer were harvested into EDTA anticoagulation tubes. Patients with any history of tumors or other organ diseases were excluded. Briefly, the initial blood samples were centrifuged at 2000 × g for 20 min at 4°C to remove the cells and cell debris. The supernatant was then aliquoted into a 2 ml centrifuge tube and stored at -80°C.

Plasma sEV isolation

The clinical blood samples were collected and centrifuged at 2000 × g for 20 min at 4° to remove the cells and cell debris. Then, the supernatants were harvested and centrifuged at 10000 × g for 20 min and subsequently at 12000 × g for 15 min to harvest the pellets of crude EVs. The moderate exosome isolation reagents (C10110-2, RiboBio, China) by Polyethylene glycol (PEG) precipitation were added to the pellets. The mixtures were thoroughly shaken and incubated for 30 min at 4°C. Finally, the mixtures were centrifuged at 15000 × g for 2 min at 4°C, after which the supernatants were discarded. The pellets were resuspended in 200 μl of PBS and stored at -80°C.

Nanoparticle tracking analysis (NTA)

The size of resuspended sEVs was measured by a NanoSight NS300 instrument (Malvern PANalytical, UK). The distribution and concentration of acquired data were performed and analyzed by NTA software (version 3.1, NanoSight).

WB analysis

The extracted sEVs were lysed in RIPA buffer (R0010, Solarbio, China), boiled, and loaded onto a 10% Tris–gly gel. The separated proteins were transferred to polyvinylidene difluoride membranes, which were subsequently blocked with 5% skimmed milk in TBST for 60 min at room temperature. After washing three times, the membrane was incubated overnight at 4°C with primary antibodies against CD63 (ab134045, Abcam), CD81(ab109201,abcam), TSG101 (DF8427, Affinity Biosciences), Calnexin (610523, Becton, Dickinson and Company), C1QB (16919-1-AP, Protein Tech), C1QC (16889-1-AP, Protein Tech), C4BPA(11819-1-AP, Protein Tech), and C4BPB (15837-1-AP, Protein Tech). After washing, the membrane was incubated for 60 min at room temperature together with an HRP-conjugated secondary antibody (RM3002, Ray, China) and then detected by Super ECL detection reagent (36208ES60, Yeasen, China). The immunoreactive bands were visualized by a Tanon Automatic Chemiluminescence and Fluorescence Image Analysis System (Tanon, China).

Transmission electron microscopy (TEM)

In brief, 5 μl of each sample was placed on Formvar/carbon-coated TEM grids for 3 min at room temperature. The excess fluid at the edge was blotted with filter paper. After rinsing with PBS, the samples were negatively stained with phosphotungstic acid and allowed to dry naturally for 5 min. TEM images for sEVs were taken using a JEM-1200EX Transmission electron microscope (JEOL, Tokyo, Japan) at 120 kV.

Liquid chromatography-tandem mass spectrometry (LC–MS/MS) preparation

The proteins of sEVs were lysed by SDT (4% SDS, 100 Mm Tris-HCl, 1 Mm DTT, pH 7.6) buffer. The protein was quantified using a BCA Protein Assay Kit (P0012, Beyotime, China). Two hundred micrograms of protein was incorporated into 30 μl of SDT buffer (4% SDS, 100 mM DTT, 150 mM Tris-HCl pH 8.0). Then, samples were ultrafiltrated to remove the detergent, DTT, and the other low-molecular-weight components by UA buffer (8 M urea, 150 mM Tris-HCl (pH 8.0)). The samples were incubated with

100 µl of iodoacetamide (100 Mm IAA in UA buffer) for 30 min in darkness to block cysteine residues and subsequently digested with 4 µg of trypsin (Promega) in 25 mM NH_4HCO_3 buffer overnight at 37°C. The digested peptides were filtered and collected after washing with 100 µl of UA buffer and 100 µl of 25 mM NH_4HCO_3 buffer. The peptides of each sample were desalted on C18 Cartridges (Empore SPE Cartridges C18, bed I.D. 7 mm, volume 3 ml, Sigma), concentrated, and reconstituted in 40 µl of 0.1% formic acid (FA). The peptide content was estimated by spectral density with ultraviolet light at 280 nm.

LC–MS/MS

LC–MS/MS analysis was performed using a Q Exactive mass spectrometer (Thermo Scientific) coupled to an Easy-Nlc system (Thermo Fisher Scientific) for 90 min. The peptide mixtures in 0.1% FA were loaded onto a reverse-phase trap column (Thermo Scientific Acclaim PepMap100, 100 µm *2 cm, nanoViper C18), separated by running a linear gradient of buffer B (84% acetonitrile and 0.1% FA) at a 300 nL/min flow rate controlled by IntelliFlow technology through a 10 cm long reversed-phase analytical column (ID 75 µm, 3 µm particle size, C18; Thermo Scientific Easy Column). The mass spectrometer was operated in positive ion mode. Survey MS spectra ranging from 300 to 1800 m/z were acquired at a resolution of 70,000 at m/z 200 using the data-dependent top 10 methods to dynamically choose the most abundant precursor ions for HCD fragmentation (17,500 at m/z 200) with automatic gain control (AGC) target of 3×10^6 and dynamic exclusion duration of 40.0 s, where the normalized collision energy was set to 30 eV and the underfill ratio was 0.1%. The instrument was operated with peptide recognition mode enabled.

The MS raw data for each sample were combined and searched using MaxQuant 1.5.3.17 software for identification and quantitation analysis against the public database UniProt (http://www.uniprot.org/). The indexes of MaxQuant identification and quantitation were as follows: two max missed cleavages of trypsin enzyme, the fixed modification of carbamidomethyl of cysteine, and the variable modification of oxidation of methionine. To acquire credible identification for proteins, we enrolled the included contaminants and used the razor and unique peptides for protein quantification with a false discovery rate of peptide and protein less than 1%. The quantification of protein expression was revealed by the LFQ algorithm.

The quantitive control of LC–MS/MS analysis contains the andromeda score distribution, mass error distribution, and the ratio distribution. The criteria of protein identification were used by the unique peptide that is unique to a group of highly similar protein sequences (protein group). The criteria of protein contaminant were used by the peptide contamination library which contained the common contaminating protein sequences such as keratins, BSA, and trypsin. These detected contaminations were excluded from further analysis.

Bioinformatics analysis

Subcellular localization

Subcellular localization of sEV proteins was predicted using the CELLO database (http://cello.life.nctu.edu.tw/), the contents of which are part of a multiclass SVM classification system.

Domain annotation

InterProScan software was applied in the search of protein sequences and the alignment of identified protein domain signatures from the InterPro member database Pfam (http://www.pfam.org/).

Gene Ontology (GO) annotation

NCBI BLAST + client software and InterProScan were applied to query the selected differentially accumulated protein sequences and identify homologous sequences. The software program Blast2GO was applied to map GO terms and annotate sequences. The GO annotation results were plotted by R scripts.

KEGG functional annotation

Following GO annotation steps, sequence alignment was performed against the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://geneontology.org/) to retrieve their KEGG orthologous pathways. Based on their orthologue information, the related KEGG pathway map was constructed.

Enrichment analysis

Based on the background dataset of the whole quantified proteins, enrichment analysis was applied using Fisher's exact test for the *p* value, which were successively adjusted by Benjamini–Hochberg correction for multiple testing. The functional categories and pathways were significantly enriched and evaluated only if the *p* value was less than 0.05.

Statistics

All data were analyzed with GraphPad Prism version 8, SPSS 19.0, ImageJ, or R. Depending on the data, Student's t-test or Fisher's exact probability test was used for comparing differences, and $p < 0.05$ was considered significant. The statistical tests and *p* values of each experiment are shown in the legends of the figures.

Results

Characterization of plasma sEVs

We extracted plasma sEVs from 21 patients and 15 controls with an sEV extraction kit (C10120, Ribo Bio, China) by Polyethylene glycol (PEG) precipitation, according to the workflow shown in Fig. 1. The extracted sEVs were confirmed by TEM, which detected mostly round and cup-like spheres (Fig. 2A), and by NTA, which analyzed the diameter and distribution of the particles. The size of the sEVs was 68.8 nm on average, ranging from 18.5 to 300 nm (Fig. 2B). Compare to the cell lysates of the HepG2 cell line, the sEV samples were analyzed by WB for the presence of three positive markers, CD63, CD81 and TSG101

(Fig. 2C), and the absence of one negative marker, Calnexin (Fig. 2C). As shown by our WB, TEM, NTA results, the sEV extraction kit we used could isolate the sEVs efficiently and accurately.

Proteomic profiling of plasma sEvs

After the identification of extracted sEVs, label-free nano-LC–MS/MS analysis was performed to comprehensively analyze the proteins from sEVs between the normal and HCC groups. To control for the interindividual variability, the sEVs from 15 controls were randomly but evenly divided into three control groups, and the sEVs from 21 HCC patients were randomly but evenly divided into three HCC groups. The results showed that there were 850,808 secondary spectral maps, of which 77,850 were matched to the UniProt public database. With the criteria of protein identification and contaminant described in the methods, 3,872 peptide segments were unique in a total of 4,167 peptide segments, and 281 proteins could be quantified in a total of 335 identified proteins(Fig. 3A).

To analyze the reproducibility of proteins identified among the groups, the overlap of proteins between groups was visualized by a Venn diagram. There were 293 and 296 sEV proteins identified in the control and HCC groups. A total of 272 proteins were expressed commonly among the control and HCC groups (Fig. 3B). A total of 229 (78.2%) and 236 (79.7%) proteins were present in all three replicates (Fig. 3C, 3D). The high overlap among the groups indicated that the results were highly reproducible and reliable.

Among these overlapping proteins, we identified significantly differentially accumulated proteins. Our quantification results revealed that a total of 54 proteins (fold change ≥ 1.5 and < 0.65) were differentially expressed in the HCC group compared to the control group. Among these proteins, 13 increased proteins and 14 decreased proteins were statistically significant ($p < 0.05$) in the HCC groups, as shown in the volcano plot in Fig. 4A. Hierarchical clustering was used to analyze the patterns of both proteins and samples, the results of which are shown in Fig. 4B.

Subcellular localization analysis

A subcellular organelle is a microorganism (such as mitochondria, the endoplasmic reticulum, etc.) with a certain shape and function in the cytoplasm. Different subcellular organelles tend to perform different cellular functions. We used Subcellular Structure Prediction Software (CELLO) to analyze the subcellular location of all differentially accumulated proteins. Among them, 30 (68.2%) proteins were extracellular, 7 (15.9%) were located in the cytoplasm, 4 (9.1%) were in the nucleus, 2 (4.5%) were in the plasma membrane, and 1 (2.3%) was in the mitochondrion (Fig. 5A).

GO and KEGG functional annotation

To understand the function, location, and biological pathways of proteins in the sEVs of HCC patients, we annotated proteins through Gene Ontology (GO). GO is a standardized functional classification system that provides a dynamically updated standardized glossary describing the properties of genes and their products in living organisms. The GO functional annotation was divided into three categories: biological processes (BPs), molecular functions (MFs), and cellular components (CCs).

We analyzed the top 20 differentially accumulated proteins in our result by the GO functional annotation with the total protein of reference species and obtained the signature of the difference by Fisher's exact test. Bubble graphs were constructed to show the concentration of GO entries under the three categories to describe the specific categories (Fig. 5B-D). Among the cellular components, compared to those of the NC group, the sEV proteins of the HCC group were significantly located in the insulin-like growth factor ternary complex, insulin-like growth factor binding protein complex, specific granule lumen, specific granule, and fibrinogen complex (Fig. 5B). Among the biological processes, the sEV proteins of HCC were significantly involved in the positive regulation of hepatocyte-cell adhesion, the positive regulation of the MAPK cascade, the epithelial cell apoptotic process, and the regulation of the extrinsic apoptotic signaling pathway via death domain receptors (Fig. 5C). Among the molecular functions, the sEV proteins of HCC were significantly involved in insulin-like growth factor binding, hydrolyzing O-glycosyl compounds, hydrolase acting on glycosyl bonds, growth factor binding, and extracellular matrix structural constituent (Fig. 5D). Differentially accumulated proteins in the HCC and NC groups were analyzed and annotated using the KEGG pathway database to predict the related pathways (the top 10 KEGG pathways) and involved proteins. We found that differentially accumulated proteins were contained mostly in the complement and coagulation cascades (Fig. 5E). Further significance was determined by Fisher's exact test, which found that complement and coagulation cascades, growth hormone synthesis, secretion and action, and platelet activation were significantly related to the progression of HCC ($p < 0.05$) (Fig. 5F). As the complement and coagulation cascades pathway is mainly involved in HCC sEVs, the detailed change in each protein in the complement and coagulation cascades pathway was visualized (as shown in Fig. 6) to understand the role of sEV proteins and the upstream and downstream pathways. The results indicate that the complement and coagulation pathway is the major pathway involved in sEV proteins.

Protein network analysis

Based on the STRING database, we analyzed the network of differentially accumulated proteins within our study. Clusters of well-known protein interactions were constructed, and the relationships among the clusters were associated under a confidence view, where more links represent closer associations, as shown in Fig. 7. The main clusters involved in the plasma sEVs of HCC patients were the complement cascade (C1QB, C1QC, C4BPA, and C4BPB) and coagulation cascade (F13B, FGA, FGB, and FGG). These data were highly correlated with the results of the hierarchical clustering and KEGG pathway analysis, where overexpression of C1QB, C1QC, C4BPA, C4BPB, F13B, FGA, FGB, and FGG in plasma sEVs was the major factor in the classification of HCC.

Validation of complement proteins in sEVs

To verify the findings in the proteomics, the protein expressions of the complement cascade were monitored in the above discovery cohort which samples had been performed as above. All the proteins of sEVs were quantified using the BCA method to correct for the difference among sEV samples without appropriate reference. The proteins significantly increased in the complement cascade were detected by WB. The results showed that C1QB, C1QC, C4BPA, and C4BPBP were increased in the plasma sEVs of

patients (Fig. 8A). The grey densities of C1QB, C1QC, and C4BPA in the HCC groups were significantly higher than in the control groups ($p < 0.05$) in Fig. 8B-D. The grey density of C4BPB in the HCC groups was significantly higher than in the control groups ($p < 0.01$) in Fig. 8E.The above results were consistent with our proteomic analysis.

To further explicit these differential expression proteins of not only mixed samples but also individual patients, we examined these differential expression proteins in an independent validation cohort with HCC patients (n = 7) and normal controls (n = 7). The proteins significantly increased in the complement cascade were detected by WB. The results showed that C1QB, C1QC, C4BPA, and C4BPBP were increased in the plasma sEVs of patients (Fig. 8F). The grey densities of C1QB, C1QC, and C4BPA in the HCC groups were significantly higher than in the control groups ($p < 0.0001$) in Fig. 8B-D. The grey density of C4BPB in the HCC groups was significantly higher than in the control groups ($p < 0.001$) in Fig. 8E.

Discussion

With in-depth research on sEVs, the compositions and functions of sEVs have gradually received attention in tumor biology and developmental biology. Among the numerous sEV omics studies, most of them have focused on the contents of miRNAs, lncRNAs, and circRNAs, but few have focused on proteins. Proteomics research on sEVs falls far behind transcriptomics research. Therefore, applying proteomics analysis, we aimed to comprehensively detect the differentially accumulated proteins in the sEVs to discover potential biomarkers for HCC evaluation. In our study, we launched a systematic approach for biomarkers in plasma sEVs of HCC, and the following components were included: (1) isolation and identification of sEVs, (2) determination of the profiles of differentially accumulated proteins with label-free quantification, (3) performance of domain annotations, enrichment analysis, and network analysis of the differentially accumulated proteins, and (4) validation of the significantly accumulated proteins in the sEVs of HCC patients.

Our results described the protein signature of plasma sEVs of HCC, the upregulated pathway of complement cascades (C1QB, C1QC, C4BPA, and C4BPB) and the coagulation cascade (F13B, FGA, FGB, and FGG). The coagulation factor XIII A chain is the last enzyme in the coagulation cascade, while the coagulation factor XIII B chain protects F13A from removal. It is clear that FGG fibrinogen and F13A are highly expressed in the plasma of HCC patients, but the relationship between F13B and HCC is unknown. The mRNA level of FGG and the elevated level of plasma fibrinogen are related to clinical stage, tumor thrombosis, and prognosis of HCC(14, 15). Ander Arbelaiz(16) found that compared to the primary sclerosing cholangitis (PSC) patients, fibrinogen gamma chain (FGG), alpha-1-acid glycoprotein 1 (A1AG1), and S100A8 proteins showed the best differential diagnostic capacity in the sEVs of Cholangiocarcinoma(CCA) patients. Most notably, ficolin-2(FCN2), inter-alpha-trypsin inhibitor heavy chain H4, and FGG showed higher diagnostic value than CA19-9 in CCA I-II versus PSC patients. In addition, Asad Uzzaman et al. found similar results that FGA, FGB, and FGG were dysregulated in liver cancer sEVs and verified that the expression of fibrinogen, fibulin-1(FCN1), and thrombospondin-1 could differentiate controls from patients with liver cancer or cirrhosis (17). However, they focused on the

fibrinogen and didn't verify the fibrinogen alpha, beta, and gamma chains in more detail. Besides, Sen Wang analyzed the sEV proteins between the MHCC97-H and MHCC97-L cell lines to find that the filamin A, talin 1 and fibulin 1 in the sEVs were strongly related to HCC metastasis(18). Their results were in strong agreement with our results, which supported the above perspective of fibrinogen in HCC

As the fundamental defense system, the complement cascade is central to immunological networks which tightly regulate humoral and cellular responses to noxious stimuli. Complement components have been linked with several cancers, which indicates that complement regulation might be a factor for oncogenesis(19). Importantly, the liver is the major source for the biosynthesis of complement components and expresses a variety of complement receptors. (20). Therefore, the progression of HCC might be largely influenced by the complement system. As previous studies have shown, upregulation of C1QB or C1QC is related to prostate cancer(21), melanoma(22), renal cell carcinomas(23), and HCC(24, 25). The upregulated C1qTNF6 activated the Akt pathway to promote HCC angiogenesis(26). Upregulation of C4BP is related to colorectal cancer(27, 28), epithelial ovarian cancer(29), non-small cell lung cancer(30), and HCC(31, 32). Interestingly, HCC cells can be protected from complement attack by upregulating C4b-binding protein alpha through binding to the transcription factor Sp1(31). Ardakani et al. found that C4b-binding protein was significantly related to HCC and cirrhosis. (32). Thus, the ability of C4BP to regulate tumorigenesis in multiple organs and the liver as a primary source strongly suggests a role for C4BP in HCC. However, the relationship between sEVs and the complement system in HCC has yet to be clarified. Our results showed that the components of the complement system (C1QB, C1QC, C4BPA, and C4BPB) were more highly expressed in the HCC group sEVs than in the control group sEVs, which might be significant for the evaluation of HCC.

In summary, our results highlighted the protein signatures of plasma sEVs from HCC patients, the upregulated pathway of complement cascades and the coagulation cascade. These factors might be used for noninvasive diagnosis. This hypothesis has been confirmed by previous studies and our experiments. Our data provide useful insight into the cargos from plasma sEVs, which might be candidates for evaluating and diagnosing liver cancer.

Conclusion

Differential and multivariate analysis of proteomics in the plasma sEVs of HCC indicated that sEV proteins were significantly related to HCC. Our results described the signature of plasma sEVs of HCC, the upregulated pathway of complement cascades (C1QB, C1QC, C4BPA, and C4BPB) and the coagulation cascade (F13B, FGA, FGB, and FGG), which might be the major factors in the classification of HCC. We also verified the upregulated expression of C1QB, C1QC, C4BPA, and C4BPB. Taken together, these data suggested that sEV analysis is a valid approach for the evaluation of HCC and that C1QB, C1QC, C4BPA, and C4BPB might be potential molecular signatures.

Abbreviations

HCC: Hepatocellular carcinoma; AFP: Alpha-fetoprotein; sEV: Small extracellular vesicle; C1QB: Complement C1Q subcomponent subunit B; C1QC: Complement C1Q subcomponent subunit C; C4BPA : BP: C4b-binding protein alpha chain; C4BPB: C4b-binding protein beta chain; FGA: Fibrinogen alpha chain; FGB: Fibrinogen beta chain; FGG: Fibrinogen gamma chain; F13A: Coagulation factor XIII A chain; F13B: Coagulation factor XIII B chain; LC–MS/MS: Liquid chromatography-tandem mass spectrometry; Biological process; CC: Cellular component; MF: Molecular function; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes.

Declarations

Acknowledgement

We thank Dr. Juan Xiao, Guangxi Key Laboratory of Molecular Medicine in Liver Injury and Repair, the Affiliated Hospital of Guilin Medical University, for her valuable suggestions for our subject.

Authors' contributions

Wei Dong performed the experiments, analyzed the results, and wrote the manuscript. Zeyu Xia, Zehua Chai, Xuehong Wang, Zebin Yang, Tingrui Zhang, and Junnan Wang collected the clinical specimens and helped in performing experiments. Qinqin Zhang and Zhidong Qiu were involved in the manuscript revision and discussion. Junfei Jin designed and supervised the study, reviewed the paper, and provided technical support.

Funding

The present study was supported in part by the National Natural Science Foundation of China (81960520), the Science and Technology Planned Project in Guilin (20190206-1, 20210102-1), the Guangxi Distinguished Experts Special Fund (2019B12), the Natural Science Foundation of Guangxi (2020GXNSFDA238006), the Medical High-Level Talents Training Plan in Guangxi (G202002005), the Construction Fund of Guangxi Health Commission Key Laboratory of Basic Research in Sphingolipid Metabolism Related Diseases (the Affiliated Hospital of Guilin Medical University, ZJC2020005), the Construction Fund of Key Disciplines of Medical and Health in Guangxi (2021-8-4-3), and the Special Fund of the Central Government Guiding Local Scientific and Technological Development by Guangxi Science and Technology Department (ZY21195024).

Availability of Data and Materials

All authors will have full access to the final trial dataset. The full protocol and participant-level dataset will be made available upon request if agreed upon by the corresponding author.

Ethics approval and consent to participate

The research involving human plasma sEVs have been performed in accordance with the Declaration of Helsinki and plasma samples of controls and HCC patients were collected at the Affiliated Hospital of Guilin Medical University (Guangxi, China) after written informed consent was provided for patients and their families, according to protocols (#YJSL2021129) approved by the Institutional Review Board of the Affiliated Hospital of Guilin Medical University (Guangxi, China).

Consent for publication

Not applicable.

Competing interests

The authors have declared that no conflict of interest exists.

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 Cancer J Clin. 2018;68(6):394–424.
2. Diseases GBD, Injuries C. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020;396(10258):1204–22.
3. Benson AB, D'Angelica MI, Abbott DE, Anaya DA, Anders R, Are C, et al. Hepatobiliary Cancers, Version 2.2021, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2021;19(5):541–65.
4. Pegtel DM, Gould SJ. Exosomes. Annu Rev Biochem. 2019;88:487–514.
5. Thery C, Witwer KW, Aikawa E, Alcaraz MJ, Anderson JD, Andriantsitohaina R, et al. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. J Extracell Vesicles. 2018;7(1):1535750.
6. LeBleu VS, Kalluri R. Exosomes as a Multicomponent Biomarker Platform in Cancer. Trends Cancer. 2020;6(9):767–74.
7. Wortzel I, Dror S, Kenific CM, Lyden D. Exosome-Mediated Metastasis: Communication from a Distance. Dev Cell. 2019;49(3):347–60.
8. Mori MA, Ludwig RG, Garcia-Martin R, Brandao BB, Kahn CR. Extracellular miRNAs: From Biomarkers to Mediators of Physiology and Disease. Cell Metab. 2019;30(4):656–73.
9. Thietart S, Rautou PE. Extracellular vesicles as biomarkers in liver diseases: A clinician's point of view. J Hepatol. 2020;73(6):1507–25.
10. Wu Q, Zhou L, Lv D, Zhu X, Tang H. Exosome-mediated communication in the tumor microenvironment contributes to hepatocellular carcinoma development and progression. J Hematol Oncol. 2019;12(1):53.

11. Sirica AE, Gores GJ, Groopman JD, Selaru FM, Strazzabosco M, Wei Wang X, et al. Intrahepatic Cholangiocarcinoma: Continuing Challenges and Translational Advances. Hepatology. 2019;69(4):1803–15.
12. Li R, Wang Y, Zhang X, Feng M, Ma J, Li J, et al. Exosome-mediated secretion of LOXL4 promotes hepatocellular carcinoma cell invasion and metastasis. Mol Cancer. 2019;18(1):18.
13. Fu Q, Zhang Q, Lou Y, Yang J, Nie G, Chen Q, et al. Primary tumor-derived exosomes facilitate metastasis by regulating adhesion of circulating tumor cells via SMAD3 in liver cancer. Oncogene. 2018;37(47):6105–18.
14. Kinoshita A, Onoda H, Imai N, Iwaku A, Oishi M, Tanaka K, et al. Elevated plasma fibrinogen levels are associated with a poor prognosis in patients with hepatocellular carcinoma. Oncology. 2013;85(5):269–77.
15. Zhu WL, Fan BL, Liu DL, Zhu WX. Abnormal expression of fibrinogen gamma (FGG) and plasma level of fibrinogen in patients with hepatocellular carcinoma. Anticancer Res. 2009;29(7):2531–4.
16. Arbelaiz A, Azkargorta M, Krawczyk M, Santos-Laso A, Lapitz A, Perugorria MJ, et al. Serum extracellular vesicles contain protein biomarkers for primary sclerosing cholangitis and cholangiocarcinoma. Hepatology. 2017;66(4):1125–43.
17. Uzzaman A, Zhang X, Qiao Z, Zhan H, Sohail A, Wahid A, et al. Discovery of small extracellular vesicle proteins from human serum for liver cirrhosis and liver cancer. Biochimie. 2020;177:132–41.
18. Wang S, Chen G, Lin X, Xing X, Cai Z, Liu X, et al. Role of exosomes in hepatocellular carcinoma cell mobility alteration. Oncol Lett. 2017;14(6):8122–31.
19. Laskowski J, Renner B, Pickering MC, Serkova NJ, Smith-Jones PM, Clambey ET, et al. Complement factor H-deficient mice develop spontaneous hepatic tumors. J Clin Invest. 2020;130(8):4039–54.
20. Malik A, Thanekar U, Amarachintha S, Mourya R, Nalluri S, Bondoc A, et al. "Complimenting the Complement": Mechanistic Insights and Opportunities for Therapeutics in Hepatocellular Carcinoma. Front Oncol. 2020;10:627701.
21. Hong Q, Sze CI, Lin SR, Lee MH, He RY, Schultz L, et al. Complement C1q activates tumor suppressor WWOX to induce apoptosis in prostate cancer cells. PLoS One. 2009;4(6):e5755.
22. Luo Y, Robinson S, Fujita J, Siconolfi L, Magidson J, Edwards CK, et al. Transcriptome profiling of whole blood cells identifies PLEK2 and C1QB in human melanoma. PLoS One. 2011;6(6):e20971.
23. Zhang L, Jiang H, Xu G, Chu N, Xu N, Wen H, et al. iTRAQ-based quantitative proteomic analysis reveals potential early diagnostic markers of clear-cell Renal cell carcinoma. Biosci Trends. 2016;10(3):210–9.
24. Bulla R, Tripodo C, Rami D, Ling GS, Agostinis C, Guarnotta C, et al. C1q acts in the tumour microenvironment as a cancer-promoting factor independently of complement activation. Nat Commun. 2016;7:10346.
25. Zhao X, Yang W, Pei F, Ma W, Wang Y. Downregulation of matrix metalloproteinases contributes to the inhibition of cell migration and invasion in HepG2 cells by sodium valproate. Oncol Lett. 2015;10(1):531–5.

26. Takeuchi T, Adachi Y, Nagayama T. Expression of a secretory protein C1qTNF6, a C1qTNF family member, in hepatocellular carcinoma. Anal Cell Pathol (Amst). 2011;34(3):113–21.
27. Kopylov AT, Stepanov AA, Malsagova KA, Soni D, Kushlinsky NE, Enikeev DV, et al. Revelation of Proteomic Indicators for Colorectal Cancer in Initial Stages of Development. Molecules. 2020;25(3).
28. Battistelli S, Stefanoni M, Lorenzi B, Dell'Avanzato R, Varrone F, Pascucci A, et al. Coagulation factor levels in non-metastatic colorectal cancer patients. Int J Biol Markers. 2008;23(1):36–41.
29. Matsuo K, Tanabe K, Ikeda M, Shibata T, Kajiwara H, Miyazawa M, et al. Fully sialylated alpha-chain of complement 4-binding protein (A2160): a novel prognostic marker for epithelial ovarian carcinoma. Arch Gynecol Obstet. 2018;297(3):749–56.
30. Liu YS, Luo XY, Li QR, Li H, Li C, Ni H, et al. Shotgun and targeted proteomics reveal that pre-surgery serum levels of LRG1, SAA, and C4BP may refine prognosis of resected squamous cell lung cancer. J Mol Cell Biol. 2012;4(5):344–7.
31. Feng G, Li J, Zheng M, Yang Z, Liu Y, Zhang S, et al. Hepatitis B virus X protein up-regulates C4b-binding protein alpha through activating transcription factor Sp1 in protection of hepatoma cells from complement attack. Oncotarget. 2016;7(19):28013–26.
32. Ehsani Ardakani MJ, Safaei A, Arefi Oskouie A, Haghparast H, Haghazali M, Mohaghegh Shalmani H, et al. Evaluation of liver cirrhosis and hepatocellular carcinoma using Protein-Protein Interaction Networks. Gastroenterol Hepatol Bed Bench. 2016;9(Suppl1):S14-S22.

Figures

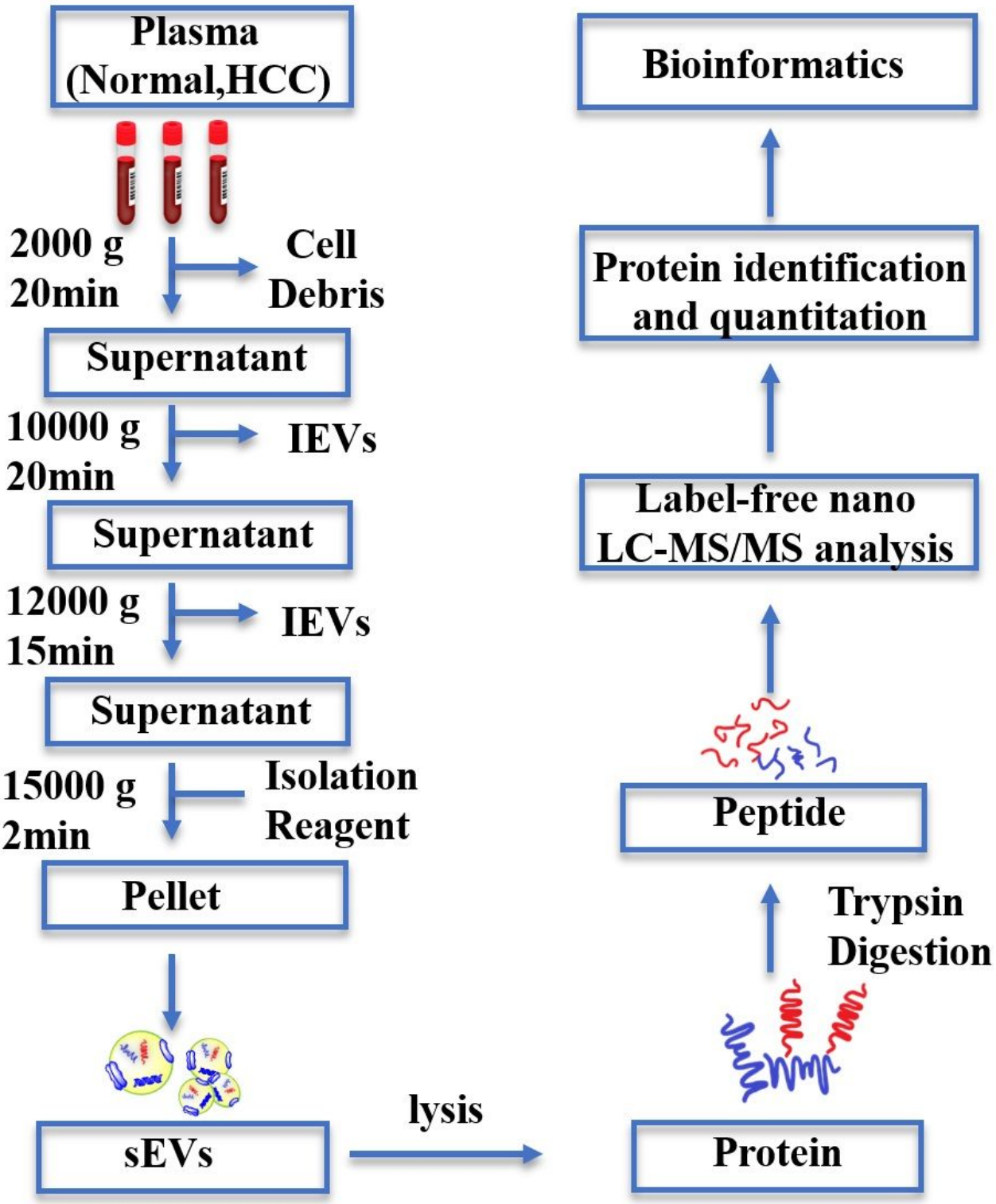

Figure 1

The strategy for sEV isolation by PEG precipitation and biomarker selection.

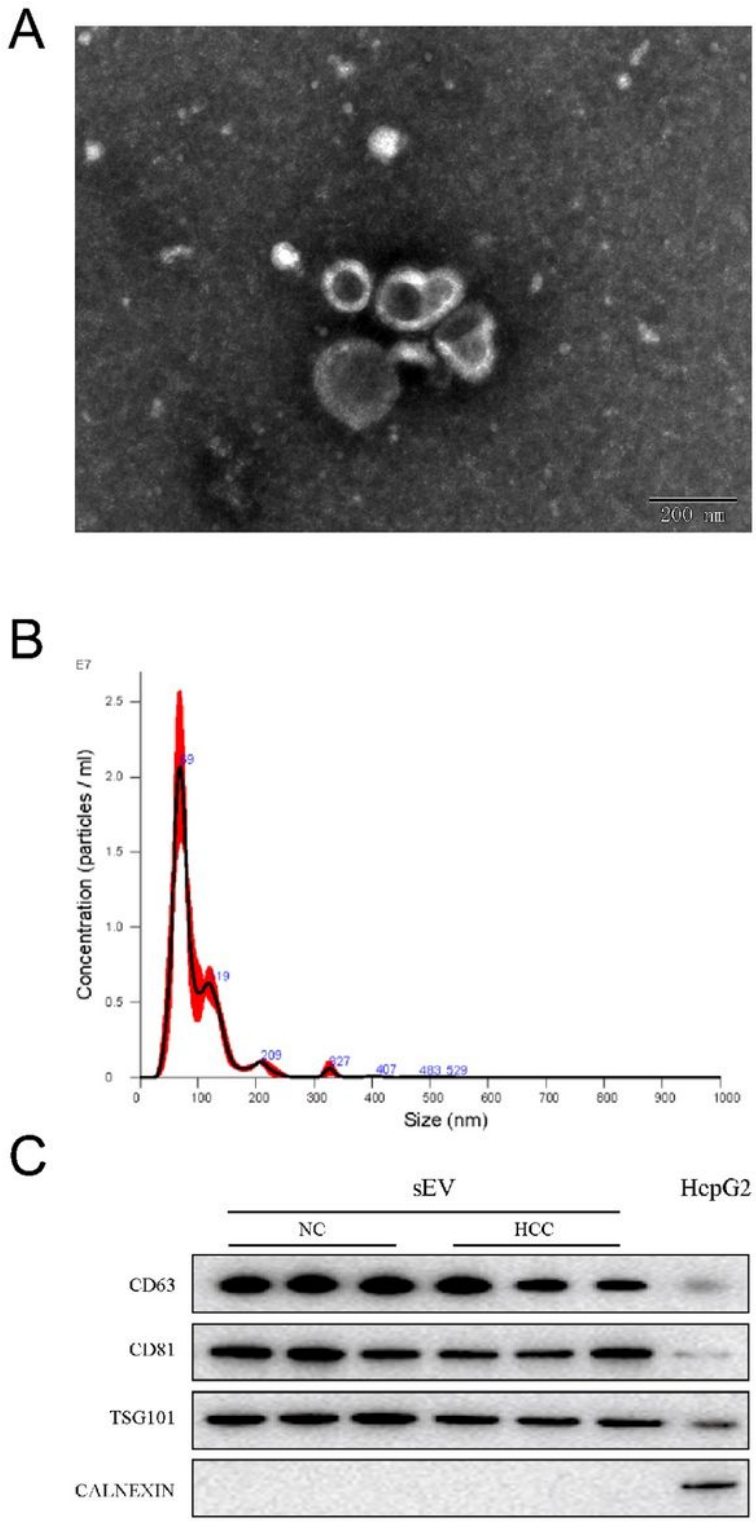

Figure 2

sEV characterization and verification with TEM, WB, and NTA.

(A) Electron microscopy of the morphology of plasma sEVs from plasma.

(B) NTA results of the sEVs from plasma.

(C) WB results of three positive markers and one negative marker of sEVs.

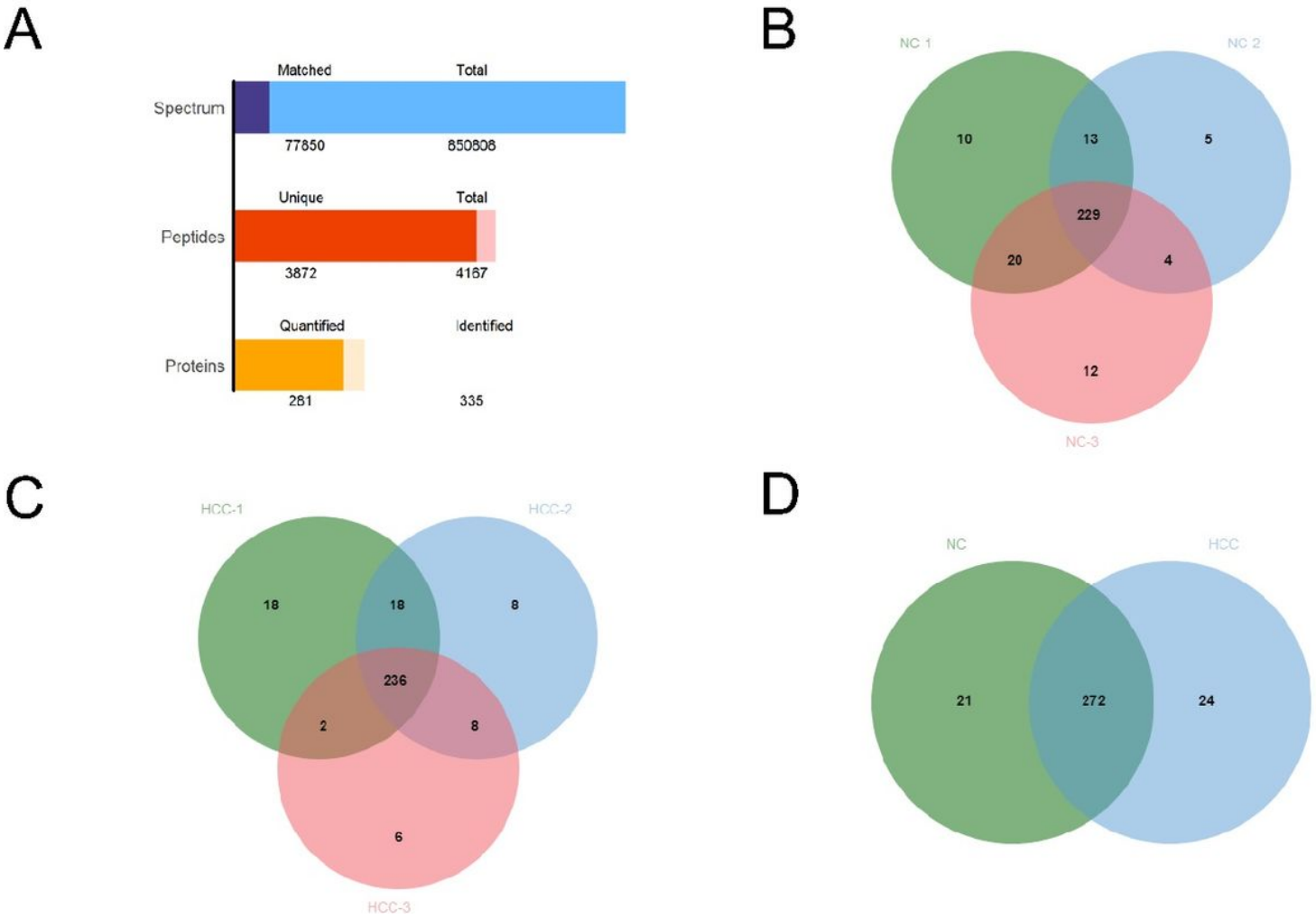

Figure 3

Proteomic profiling of plasma sEVs.

(A) Histogram of the statistics of the identification and quantitative results. (B) Proteins of sEVs identified from 3 independent control cohorts as Venn diagrams. (C) Proteins of sEVs identified from 3 independent HCC cohorts as Venn diagrams. (D) Proteins of sEVs identified between the normal and HCC cohorts as Venn diagrams.

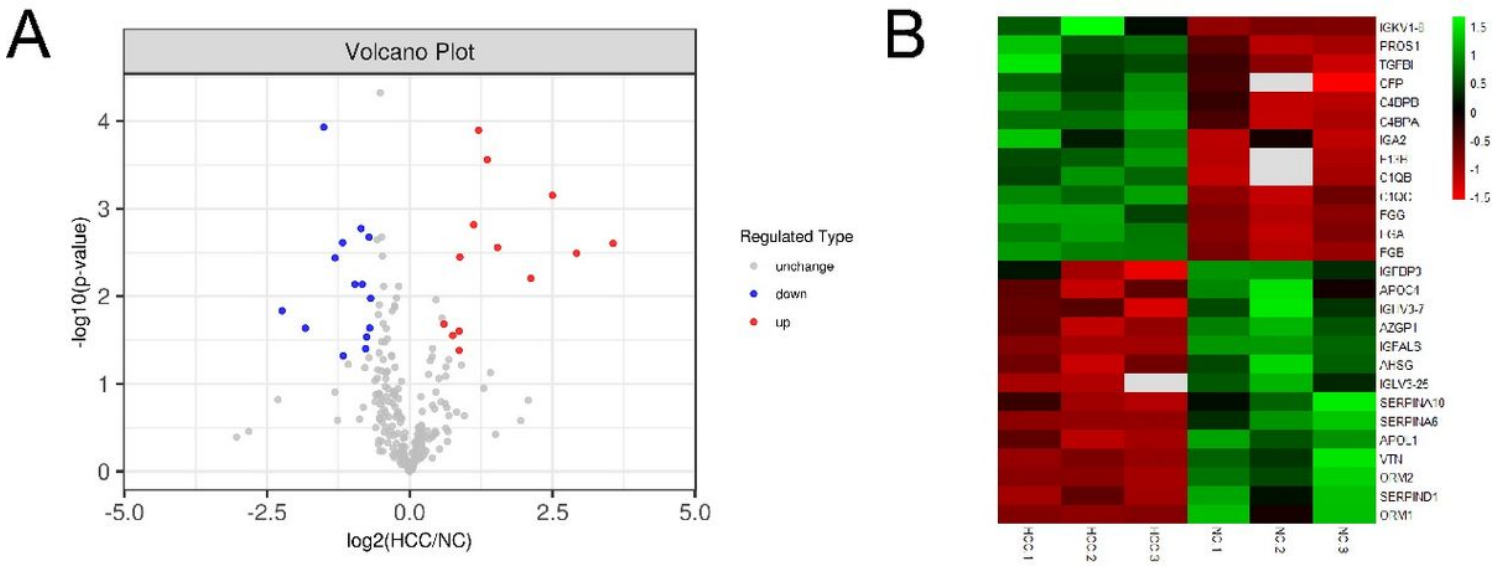

Figure 4

Differences in protein expression.

(A) Volcano plot of protein expression from HCC patients' sEVs compared to the normal group.

(B) Horizontal hierarchical clustering results expressed in tree heatmaps. Each column represents a set of samples (abscissa for sample information), and each row represents a protein (that is, the ordinate for significantly differentially accumulated protein). The colour scale shown in the map illustrates the relative protein expression: red represents increased proteins, green represents decreased proteins, and the blanks represent samples for which there was no quantitative information.

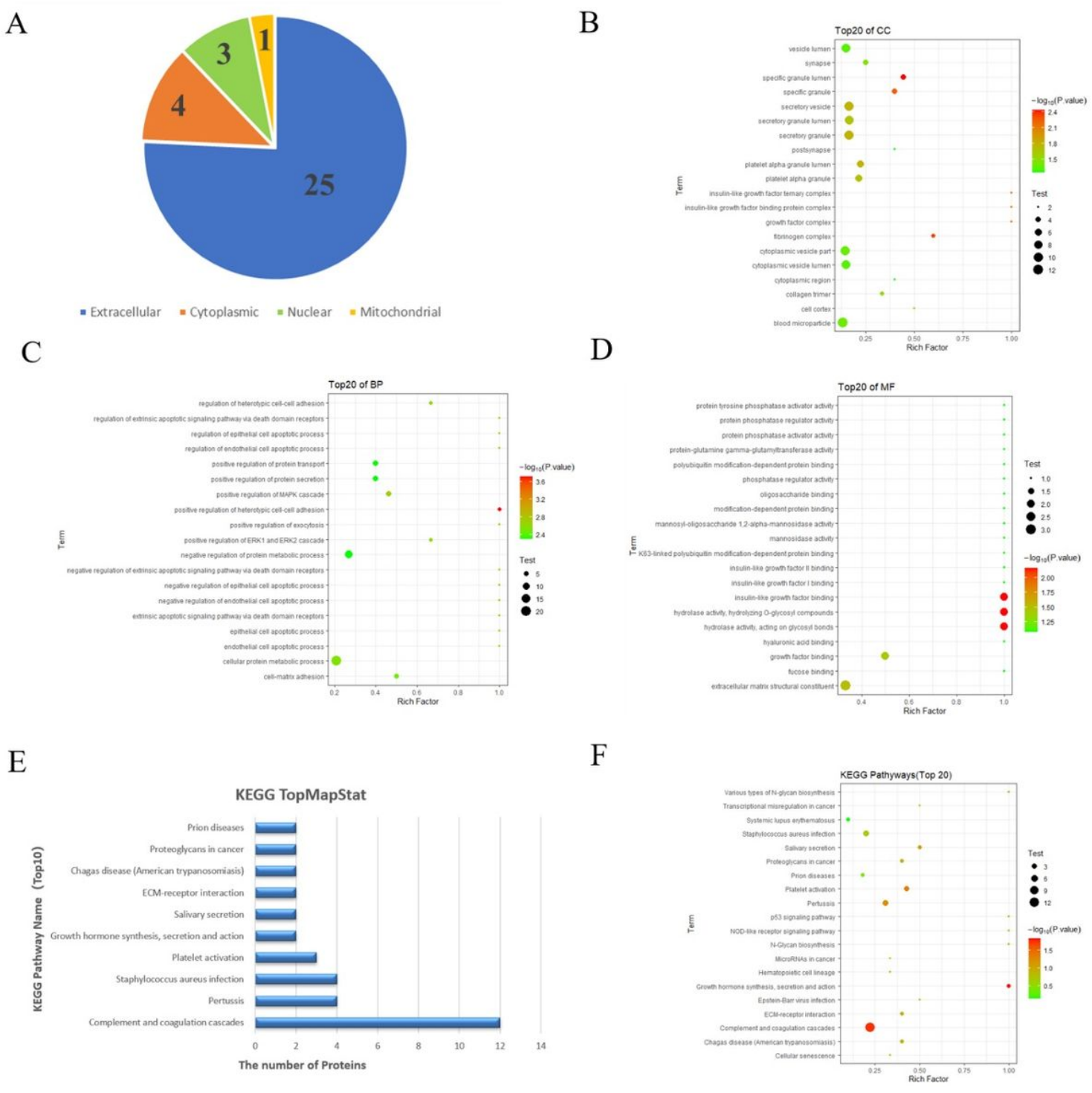

Figure 5

Localization analysis and functional annotation of the HCC and control groups.

(A) Proportion of subcellular locations among differentially accumulated proteins. (B) Bubble diagram of GO functional enrichment in the biological processes (BPs). (C) Bubble diagram of GO functional enrichment in the cell components (CCs). (D) Bubble diagram of GO functional enrichment in molecular functions (MFs). (E) Bar graph of the number of differentially accumulated proteins assigned to KEGG pathways. (F) Bubble diagram of the KEGG pathway enrichment analysis. The horizontal axis in the

figure is the enrichment factor (richness factor ≤ 1). The color of bubbles represents the significance of the enriched GO functional classification. The color gradient represents the *p* value ($-\log_{10}$), and the size of bubbles represents the number of enriched GO functional classifications.

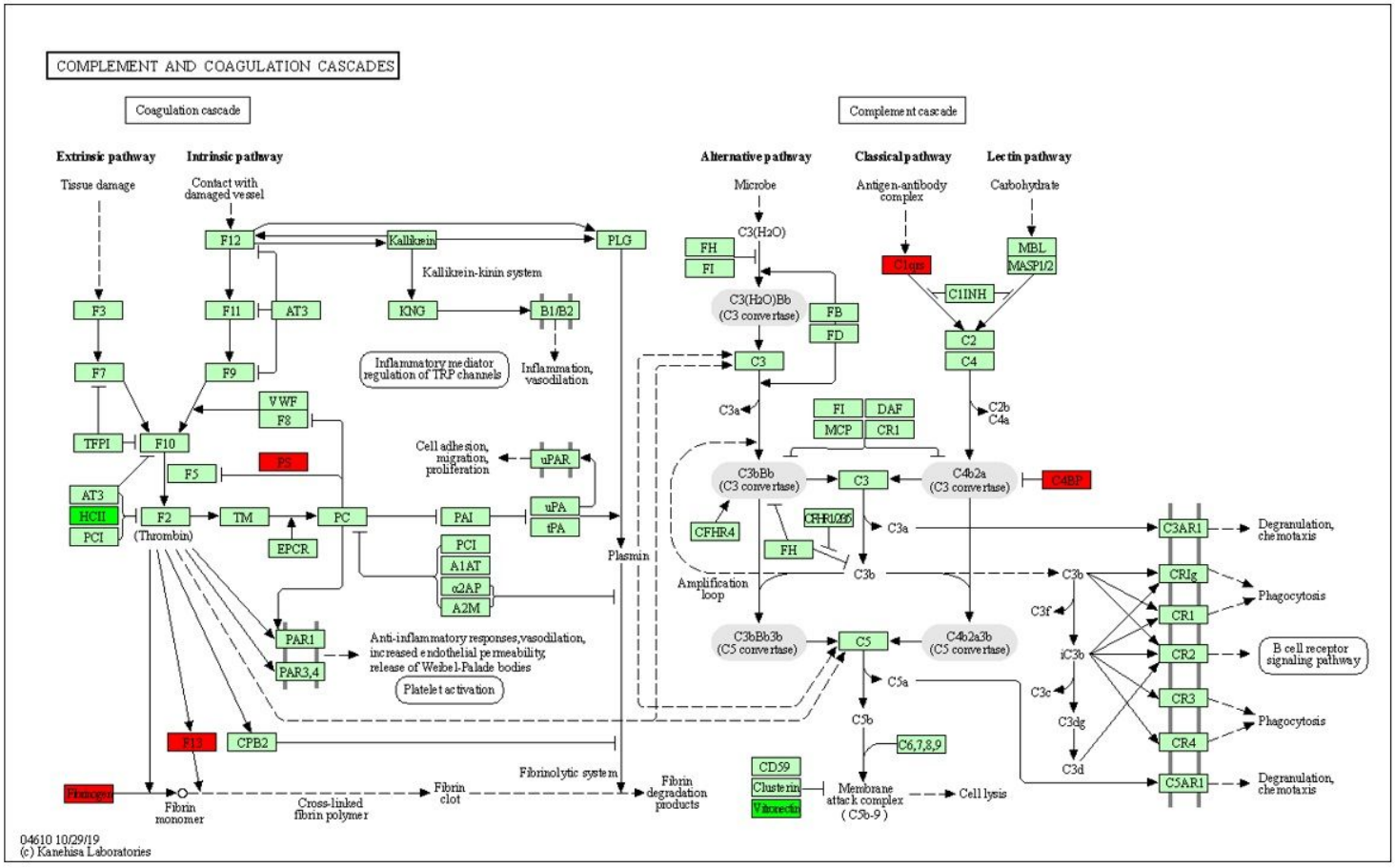

Figure 6

KEGG pathway map of the complement and coagulation cascades.

The red boxes indicate increased proteins, and the green boxes indicate decreased proteins.

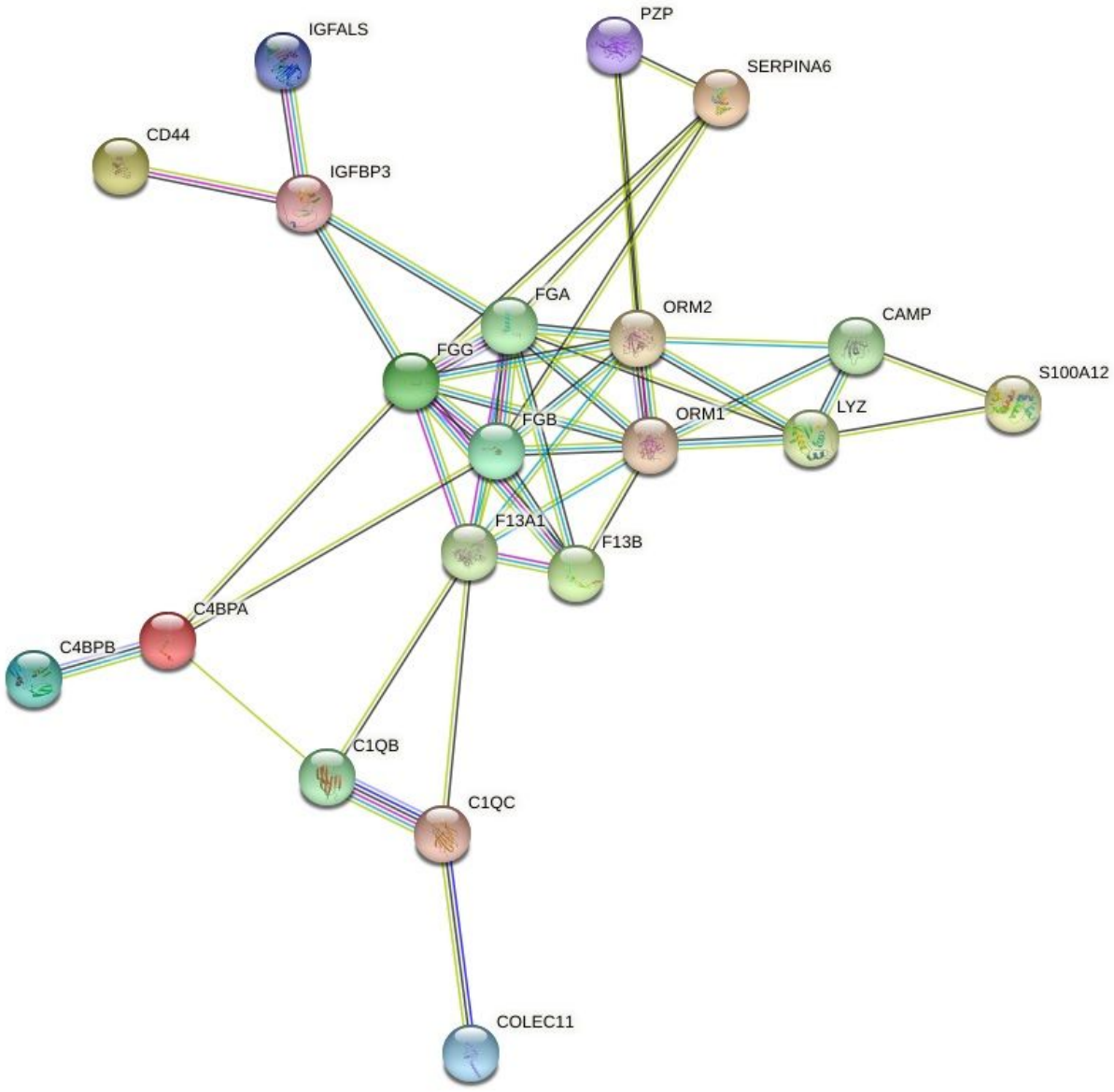

Figure 7

Protein-protein interactions of dysregulated proteins identified in HCC patients and controls.

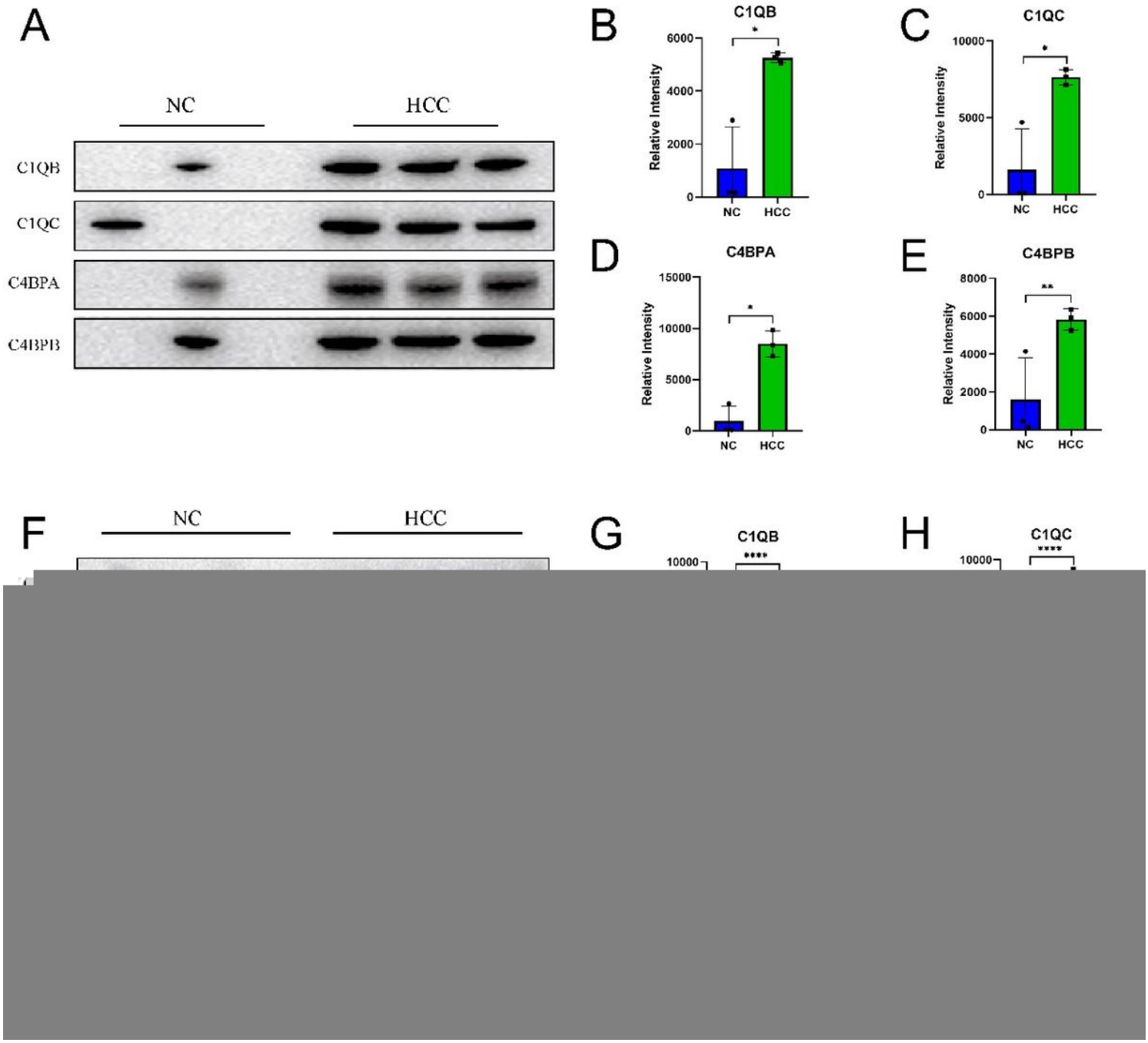

Figure 8

Validation of complement proteins. (A) WB shows the high expression levels of C1QB, C1QC, C4BPA, and C4BPBP in the plasma sEVs of HCC patients in the discovery cohort. (B) Column chart showing the grey value of the C1QB protein band between the control and HCC groups in the discovery cohort. (C) Column chart showing the grey value of the C1QC protein band between the control and HCC groups in the discovery cohort. (D) Column chart showing the grey value of the C4BPA protein band between the control and HCC groups in the discovery cohort. (E) Column chart showing the grey value of the C4BPB protein band between the control and HCC groups in the discovery cohort. (F) WB shows the high expression levels of C1QB, C1QC, C4BPA, and C4BPBP in the plasma sEVs of HCC patients in the validation cohort. (G) Column chart showing the grey value of the C1QB protein band between the control

and HCC groups in the validation cohort. (H) Column chart showing the grey value of the C1QC protein band between the control and HCC groups in the validation cohort. (I) Column chart showing the grey value of the C4BPA protein band between the control and HCC groups in the validation cohort. (J) Column chart showing the grey value of the C4BPB protein band between the control and HCC groups in the validation cohort. The left ordinate shows the grey value measured by ImageJ software. *, p <0.05; **, p <0.01; ***, p <0.001, and ****, p <0.0001.