

Quantitative proteomic profiling of hepatocellular carcinoma at different serum alpha-fetoprotein level

Xuyong Wei^{a,d,1}, Renyi Su^{c,1}, Mengfan Yang^c, Binhua Pan^c, Jun Lu^a, Hanchao Lin^a,
Wenzhi Shu^a, Rui Wang^a, Xiao Xu^{a,b,d,*}

^a Key Laboratory of Integrated Oncology and Intelligent Medicine of Zhejiang Province, Department of Hepatobiliary and Pancreatic Surgery, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou 310006, China

^b Institute of Organ Transplantation, Zhejiang University, Hangzhou 310003, China

^c Department of Hepatobiliary and Pancreatic Surgery, The First Affiliated Hospital, Zhejiang University School of Medicine, 79 Qingchun Road, Hangzhou 310003, China

^d Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou 310024, China

ARTICLE INFO

Keywords:

Adjacent noncancerous tissue
Alpha-fetoprotein
HCC
Quantitative proteomics
Prognosis biomarker

ABSTRACT

Purpose: Hepatocellular carcinoma (HCC) is characterized by a poor long-term prognosis and high mortality rate. Serum alpha-fetoprotein (AFP) levels show great prognostic value in patients undergoing hepatectomy. This study aims to explore proteomic profiling in HCC samples based on AFP subgroups and identify potential key targets involved in HCC progression.

Methods: Twelve paired tumor and adjacent noncancerous tissue samples were collected from patients with HCC who underwent primary curative resection from January 2012 to December 2013. Clinical information was curated from four tissue microarrays to conduct survival analysis based on serum AFP levels. TMT-based quantitative proteomic analyses and bioinformatics analyses were performed to comprehensively profile molecular features. Immunohistochemistry was carried out to validate protein expression of identified targets. Kaplan-Meier survival analysis was performed to assess the overall survival and recurrence-free survival based on protein expressions.

Results: AFP (400 ng/mL) was a turning point in prognosis, metabolic- and invasion-associated pathways. The mass spectrometry analysis yielded a total of 5573 identified proteins. Annotations of 151 differentially expressed proteins in tumors and 95 proteins in paracancerous tissues (1.2-fold) showed similarities in biological processes, cellular components, molecular functions. Furthermore, differentially expressed hub proteins with five innovatively nominated druggable targets (C1QBP, HSPE1, GLUD2 for tumors and CHDH, ITGAL for paracancerous tissues), of which four (C1QBP, HSPE1, CHDH, ITGAL) targets were associated with poor overall survival (all Log-rank $P < 0.05$).

Conclusions: Our quantitative proteomics analyses identified four key prognostic biomarkers in HCC and provide opportunities for translational medicine and new treatment.

Introduction

Hepatocellular carcinoma (HCC) is one of the prevailing tumors featuring a poor long-term prognosis and high mortality rate [1]. Despite recent progress in treatment options and regimens, HCC is still a disease with a 5-year survival rate of only 20% [2]. Hepatitis B virus (HBV) infection is the main etiology of HCC with a significant global

health burden, and approximately 257 million people are chronically infected with HBV, as estimated by the World Health Organization [3].

Alpha fetoprotein (AFP), a secreted 70 kD glycoprotein, is mainly produced in the fetal yolk sac and liver during pregnancy. Since the discovery of tumor-associated biological characteristics in the mid-1960s, AFP has become the most widely tested biomarker for HCC in the clinical setting [4]. Clinically, AFP has been identified as a promising

* Corresponding author at: Key Laboratory of Integrated Oncology and Intelligent Medicine of Zhejiang Province, Department of Hepatobiliary and Pancreatic Surgery, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou 310006, China.

E-mail address: zjxu@zju.edu.cn (X. Xu).

¹ These authors contributed equally to this work.

<https://doi.org/10.1016/j.tranon.2022.101422>

Received 1 December 2021; Received in revised form 30 March 2022; Accepted 30 March 2022

1936-5233/© 2022 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

biomarker for screening and surveillance, prognosis stratification, predicting the therapy response and selecting an ideal beneficiary population [5]. Based on completed phase 3 trials, including SHARP, REFLECT, and RESORCE, baseline serum AFP levels may provide clear prognostic information to guide therapy [6–8]. Mechanistically, increasing evidence shows the complexity of AFP-mediated mechanisms, including PI3K/AKT-mediated proliferation and invasion, Fas/FADD-mediated apoptosis, and modulation of the antitumor immune response [9–12]. However, the precise functions of AFP in carcinogenesis and progression and the associated mechanisms underlying differences in the prognosis of patients stratified into AFP subgroups remain incompletely understood.

Mass spectrometry-based quantitative proteomic profiling has been an effective approach for elucidating the molecular subtypes, interactive networks and associated pathway features that might be missed when analyzing other omics data [13]. Therefore, in-depth mining of proteins, the main executors of physiological and pathological functions, will facilitate a comprehensive understanding of the underlying mechanisms of HBV-related HCC and the identification of new therapeutic targets.

In this study, the focus was to explore the molecular mechanisms underlying differences in the prognosis of AFP subgroups. Hence, we performed a TMT-based quantitative proteomic analyses, and comprehensive bioinformatics analyses were performed on 12 patients with hepatitis B virus-related HCC, and tumor and paracancerous tissues were analyzed separately after stratification into AFP subgroups. The prognostic analysis revealed a shorter survival for patients with a serum AFP level > 400 ng/mL than in those with an AFP level < 400 ng/mL. Gene set variation analysis (GSVA) and gene set enrichment analysis (GSEA) revealed that tumors and adjacent noncancerous tissues presented distinct patterns in pathways, both with a cutoff value of 400 ng/mL for serum AFP levels. Integrative annotation analyses indicated similarities in the molecular features of tumor and adjacent noncancerous tissues in an AFP-dependent manner. Moreover, we constructed protein–protein interaction networks and innovatively identified three druggable targets in tumors and two in adjacent noncancerous tissues, of which four (C1QBP, HSPE1, CHDH, ITGAL) out of five (C1QBP, HSPE1, GLUD2, CHDH, ITGAL) were validated as prognostic markers and potential therapeutic targets, suggesting tailored treatment strategies for different tissue types in patients with HBV-related HCC.

Materials and methods

Clinical sample and information acquisition

Paired tumor and adjacent non-cancerous tissues were collected from 12 HBV-related HCC patients, resulting in a total of 24 samples. All of patients underwent primary curative resection from January 2012 to December 2013 at the First Affiliated Hospital of Zhejiang University without previous anticancer treatments. Tissue specimens were snap-frozen using liquid nitrogen within 30 min after surgery. The median age of the patients was 56, ranging from 43 to 68, with 10 males and 2 females. Histologic grades were defined as well differentiated, moderately differentiated and poorly. Detailed clinical information were summarized in Table S1. The tissue microarrays (Microarray Number: ZL-LVC1606; ZL-LVC1607) were purchased from Shanghai Zhuolibiotech company Co., Ltd. (Shanghai, China). Each tissue microarray contained 80 paired HCC samples and adjacent noncancerous tissues. Another tissue microarray (Microarray Number: HLivH180Su17) was purchased from Shanghai Outdo Biotech Co., Ltd (Shanghai, China) containing 88 paired tumor and adjacent noncancerous tissues and 4 more tumor tissues. Additional information of tissue microarray (Microarray Number: HLivH180Su06; HLivH180Su17) were downloaded from the website of Shanghai Outdo Biotech Co., Ltd (<http://www.superchip.com.cn/biology/tissue.html>). The merged clinical information was summarized in Table S2. The transcriptional expression matrix of five identified proteins and information of serum AFP level were curated from the

UCSC Xena (<https://xenabrowser.net/>) using the following steps: (I) Select “GDC TCGA Liver Cancer (LIHC)” study (II) Choose “fetoprotein_outcome_value” in “Phenotypic” module (III) Choose gene expression RNA sequencing data of five identified proteins in FPKM format using “Genomic” module. Data were processed using GDC pipeline with details deposited in website (https://docs.gdc.cancer.gov/Data/Bioinformatics_Pipelines/Expression_mRNA_Pipeline). The study was approved by the Research Ethics Committee of The First Affiliated Hospital, Zhejiang University School of Medicine, and written informed consent was obtained from each patient. The study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki (6th revision, 2008) as reflected in a priori approval by the institution’s human research committee.

Tissue microarray construction and immunohistochemistry

Tissue microarrays were constructed by Shanghai Zhuoli Biotechnology Co., Ltd (Zuoli Biotechnology Co, Shanghai, China). The construction workflow was mainly based on previously published methods [14,15]. In brief, hematoxylin–eosin was used to validate the pathologic diagnoses of tissue paraffin blocks collected from test and validation cohorts with fixed points marked under microscope to present typical histological characteristics. For each case, 1.5 mm diameter was adopted to divert cores in point-to-point arrangement from donor block into recipient block microarray and then 4 μm thick sections from recipient paraffin tissue blocks were transferred to glass slides by adhesive tape transfer system, aiming to perform ultraviolet cross linkage.

The immunohistochemistry (IHC) was performed according to a two-step protocol as previously described [16]. Briefly, tissue microarrays were firstly subjected to antigen retrieval using EDTA buffer (PH 6.0) at high temperature. Then, the sections were incubated with primary antibody and secondary antibody. The primary antibodies included C1QBP (1:1000, Cell Signaling TECHNOLOGY, 6502S), GLUD2 (1:600, Proteintech, 14462-1-AP), HSPE1 (1:10000, SANTA CRUZ BIOTECHNOLOGY, sc-376313), CHDH (1:900, Proteintech, 17356-1-AP) and ITGAL (1:200, Proteintech, 15574-1-AP). After incubated with primary antibody at 4 °C overnight, the microarrays were then incubated with an HRP-labeled anti-mouse secondary antibody (DAKO). The IHC staining process was performed under the manufacturer’s instructions. Followed by washing in PBS and visualization with diaminobenzidine and hematoxylin counterstain, the microarrays were observed and analyzed under microscope.

The protein levels of five identified molecules in tissue microarrays were semi-quantified by two experienced pathologists. When disagreements existed, a third reviewer would be required and made final decisions. The staining intensity scores were defined as 0 (no staining), 1 (weak positive), 2 (moderate positive), and 3 (strong positive). The immunoreactive scores were evaluated according to the percentage of positive cells: 0 (negative), 1 (1–25%), 2 (26–50%), 3 (51–75%), and 4 (76–100%). The final combined scores (0–12 point) for each patient were decided, where the scores < 8 point were defined as low expression group and scores ≥ 8 point were defined as high expression group, thus utilized for subsequent analyses.

TMT labeling and LC-MS/MS analysis

For TMT Labeling and LC-MS/MS analysis, the workflow was mainly based on previously published methods [17]. Briefly, the proteins were denatured and reduced at 95 °C for 5 min. After centrifugation, the remained supernatant was used for following experiment. For protein digestion, the workflow was mainly based on filter-aided sample preparation (FASP) procedure [18]. Briefly, proteins were alkylated and then digested by trypsin. The concentration of peptides was determined by BCA protein quantification kit. For each sample, 100 μg peptides were used for the following TMT labeling experiment. All tumor and adjacent noncancerous tissues were used and mixed to equal protein amount as

internal reference, and then divided into 100 µg per tube for labeling experiment. The TMT labeling experiment was performed by using Mass Tagging Kits and Reagents (90064, USA) according to the TMT kit instructions. The mixed peptides were labeled by TMT-126 as internal reference, and 12 pairs tumor and adjacent noncancerous tissues were labeled by other channels (Tumors labeled by 127T, 128T, 129T, 130T and 131T; adjacent noncancerous tissues labeled by 127C, 128C, 129C, 130C and 131C). Equal amount of Sample T and Sample C were combined, separated into 15 components by high-performance liquid chromatography, and subjected to following LC-MS/MS analysis.

All mass spectrometric data were analyzed using Proteome Discoverer software (Thermo Fisher Scientific) against the National Center for Biotechnology Information (NCBI) database. The false discovery rates (FDR) thresholds for both peptide and protein were set at 0.01. The quantitative proteomic data was summarized in Table S3

Data normalization

Proteomic data were normalized according to previous paper [17]. \log_2 (TMT-ratio value) should be centered at zero after normalization. The Refseq IDs were then converted to standard symbol IDs by R package clusterProfiler (v3.18.1) [19].

Missing value manipulation

To obtain robust molecule features of specimens, no imputation but deletion was used for any missing value, considering the relatively small sample size that may be disturbed by imputation.

Data quality analysis for proteomic data

The quality of proteomics data was evaluated by heatmap analysis, principal component analysis (PCA) analysis, relative standard deviation analysis and normalized expression analysis performed in R software (4.0.3).

Integrative bioinformatics analyses

The proteomic data with a minimal number of 3 identified peptides but without missing values ($n = 2264$ proteins) was used as input for following bioinformatics analysis. Molecular Signature Database (MsigDB, <http://software.broadinstitute.org/gsea/msigdb/>) website was used to download pathway related gene sets (c2.cp.v7.4.symbols.gmt) for GSEA and GSEA analysis [20,21]. Gene sets were filtered with a minimal size of 10 to consolidate results. Significant expression changes were determined for GSEA and GSEA at both P -value ($p < 0.05$) and fold change (FC) ($FC > 1.5$ -fold or $FC < 1/1.5$ -fold). GSEA analysis was based on GSEA R package (1.38.2) and GSEA was based on GSEABase R package (1.52.1). To meticulously identify proteins that were differentially expressed (limma R package, 3.46.0) within different sample types, enlarged thresholds were adopted at P -value ($p < 0.05$) and fold change (FC) ($FC > 1.2$ -fold or $FC < 1/1.2$ -fold). Gene Ontology (GO) annotation was performed using clusterProfiler R package (v3.18.1) and visualized by GOplot R package (1.0.2). The Clusters of orthologous groups for Eukaryotic Orthologous Groups (KOG) annotation data was downloaded from NCBI and visualized by ggplot2 R package (3.3.3). The wolfspout software (v.0.2) was used to predict subcellular localization of identified significant differential proteins. To further explore gene interactions and functional molecules at protein level, co-expression protein-protein interaction (PPI) networks were constructed by STRING website (<https://www.string-db.org/>) and all of settings were set as default. Interaction network from STRING was downloaded and visualized by R package "networkD3" (0.4). The downloaded STRING data were as input to extract hub proteins using Mcode plugin in Cytoscape software (3.8.2).

Drug target analysis

Stringent criteria, set to maximize prediction power of potential drug targets, were as follows: 1. Hub proteins without reported previously in HCC by searching PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) with combination of key words "HCC", "liver cancer" and hub proteins. 2. Drug targets either approved by US Food and Drug Administration (FDA) or under clinical trials were obtained from Drugbank Online (version 5.1.8, released 2021-01-03) (<https://www.drugbank.ca/>). The 3D structures of identified potential drug targets for tumor and adjacent noncancerous tissues were obtained from AlphaFold Protein Structure Database (<https://alphafold.ebi.ac.uk/>) and visualized in PyMOL (2.5.2) [22].

Quantification and statistical analysis

Statistical tests used in the study included Student's t test, Wilcoxon rank sum test and Log-rank test. Unpaired two-tailed Student's t tests were used to compare relative standard deviation. Two-tailed Wilcoxon rank sum tests were used to compare mRNA and protein expression levels of the five identified proteins. Kaplan–Meier plots were used to describe overall survival. All statistical tests were considered statistically significant when P -value < 0.05 . All of statistical analyses were performed using R language software (version 4.0.3) and Prism software (version 9.1.1, GraphPad Software, LLC, USA).

Results

Association between AFP levels and clinical prognosis

We meticulously explored the clinical prognostic value of serum AFP levels by classifying patients into 4 groups (Group A, AFP level < 20 ; Group B, $20 \leq$ AFP level < 100 ; Group C, $100 \leq$ AFP level < 400 ; Group D, AFP level ≥ 400 , ng/mL) [23–25]. A survival analysis was performed on 329 patients with HCC using clinical information from a previously prepared tissue microarray (Table S2). Significant differences were observed between Group A and Group D (P value = $9E-04$, Fig. S1A), as well as Group B and Group D (P value = 0.04, Fig. S1A). In contrast, no differences in survival were observed between Group C and Group A (P value = 0.3, Fig. S1A) or between Group C and Group B (P value = 0.7, Fig. S1A). A subsequent survival analysis showed that an AFP level equal to 400 ng/mL was an important turning point that was significantly correlated with overall survival (P value = 0.00051, Fig. S1B).

Discrepancy in the pattern between tumor and nontumor tissues from patients with HCC

A TMT-quantitative proteomics approach was employed using 12 paired tumor and nontumor tissues from patients with pathologically diagnosed HCC related to HBV infection to elucidate the discrepancy in the pattern between tumor and nontumor tissues from patients with HCC. The mass spectrometry analysis yielded 5573 identified proteins. The heatmap analysis showed a clear discrimination between tumor and adjacent noncancerous tissues at the protein level (Fig. 1A), which indicated a high discrepancy feature that underpinned our subsequent comparison analysis. The quality control of samples showed a similar significant discrepancy in the pattern that was corroborated by principal component analysis (Fig. 1B) and a comparison of the relative standard deviation value (P value < 0.0001 , Fig. 1C). A jittered boxplot showed the comparability of specimens at the protein level after normalization (Fig. 1D). Intriguingly, a paired review of samples showed opposite trends (AFP cutoff of 20 ng/mL), of which the number of identified proteins was higher in tumors from patients with an AFP level > 20 ng/mL than in the paired noncancerous samples but opposite in the patients with AFP levels < 20 ng/mL (Fig. 1E).

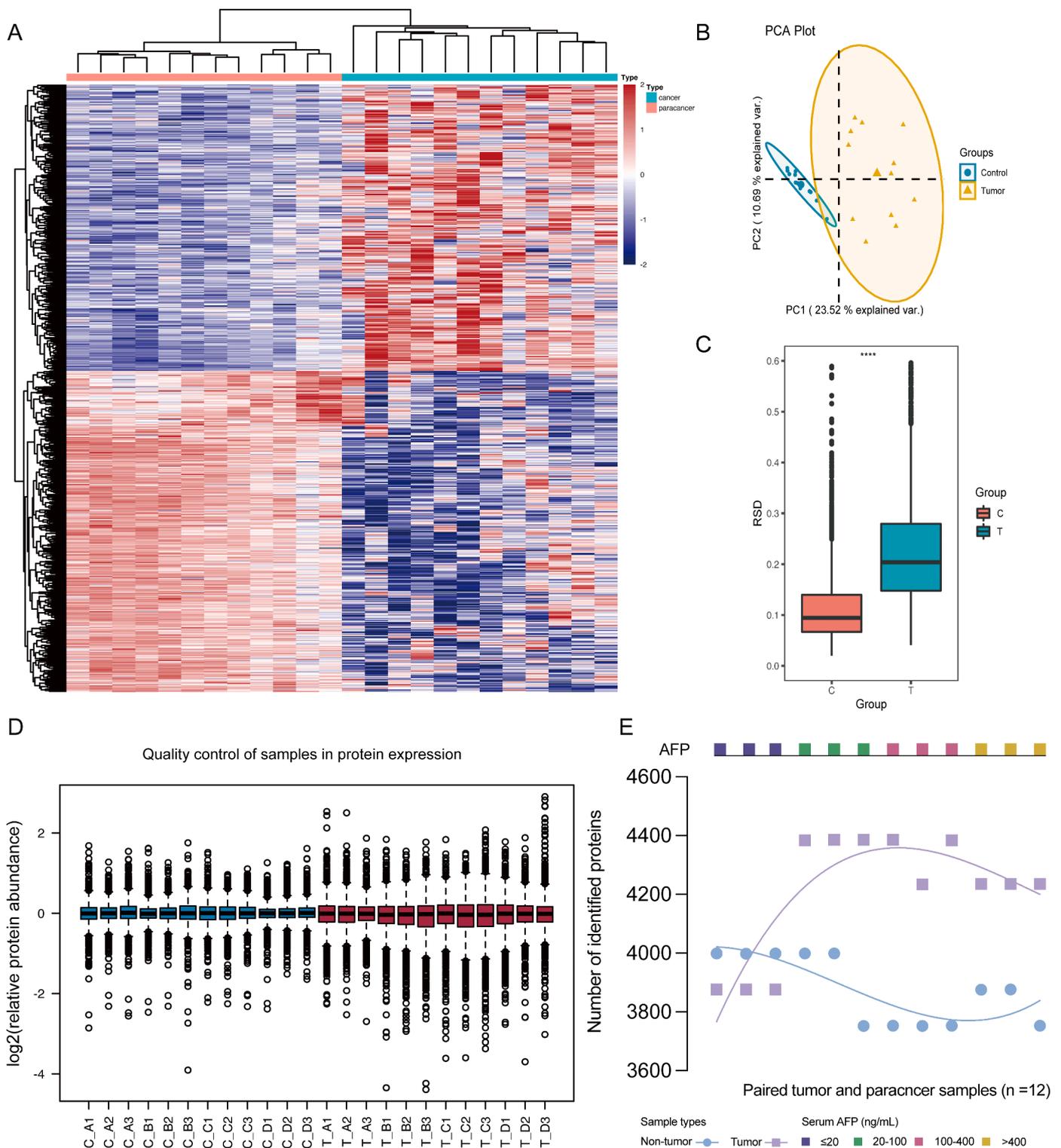


Fig. 1. Systemic quality control of TMT-based quantitative proteomic data. (A) Heatmap analysis of differential proteins. The rows indicate protein and columns indicate samples. (B) Principal components analysis of proteomic data across 12 paired tumor and adjacent noncancerous tissues. (C) Comparison of RSD between tumor and adjacent noncancerous tissues. *****p* value < 0.0001, ****p* value < 0.001, ***p* value < 0.01, **p* value < 0.05. *P* < 0.05 was considered statistical significant. (D) Boxplot of log₂ transformed normalized proteomic data. (E) Overview of number of quantified proteins in HBV-related HCC based on AFP subgroups. AFP, alpha-fetoprotein; HBV, hepatitis B virus; HCC, Hepatocellular carcinoma; RSD, relative standard deviation.

Proteomic pathway analysis of HCC samples in AFP subgroups

Analyses were performed independently for different sample types to obtain general insights into aberrant pathways altered in cancer and adjacent noncancerous tissues that accounted for differences in the

prognosis of diverse AFP groups. Intriguingly, GSVA revealed clearly separated patterns both for tumors and adjacent noncancerous tissues at a cutoff value of 400 ng/mL for serum AFP levels (Fig. 2A, C), indicating the biological activity-dependent behaviors of adjacent noncancerous tissues. However, relatively diminished activities of multiple

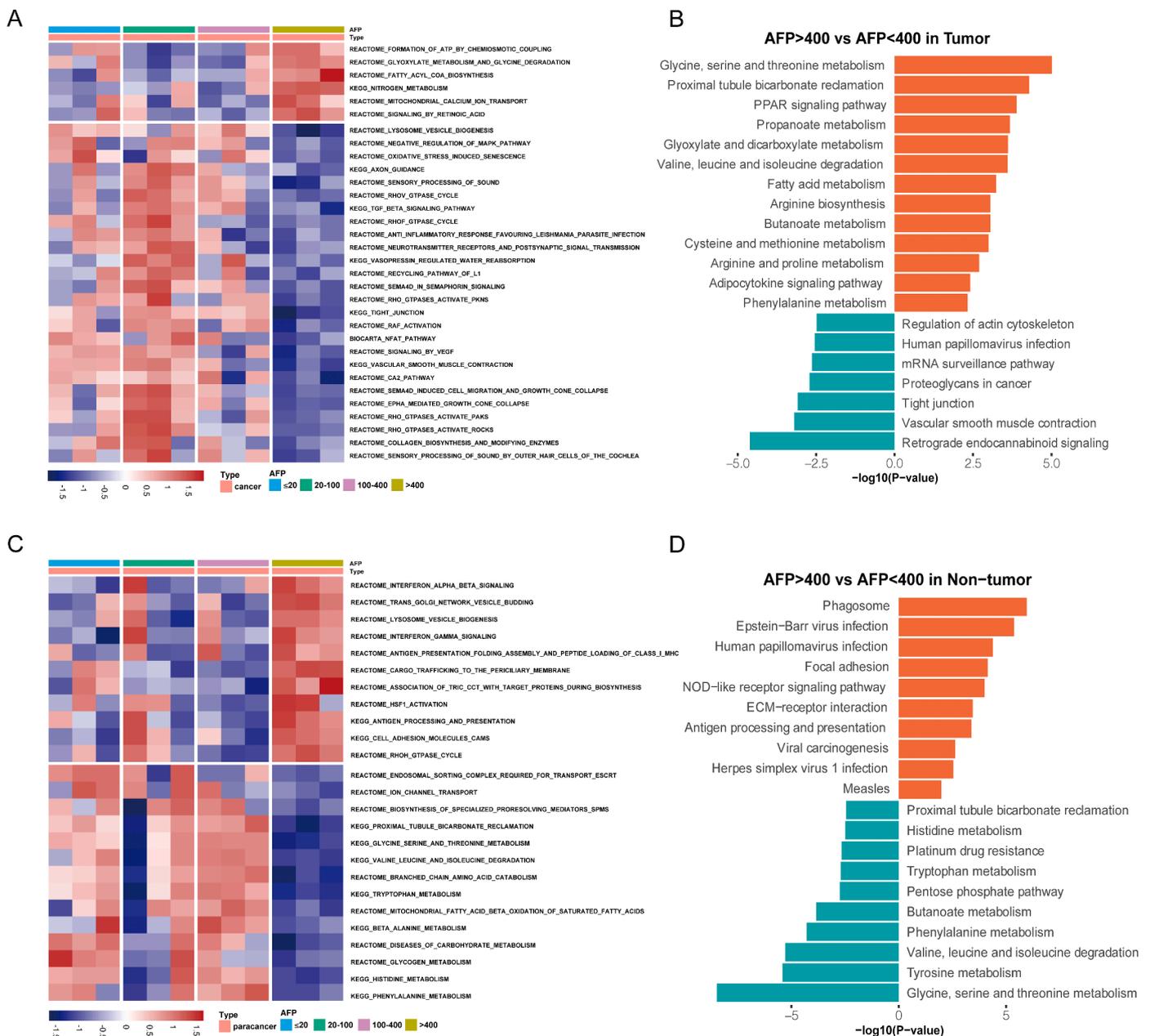


Fig. 2. Proteomic pathway analysis of HCC samples in AFP subgroups. Heatmaps of enriched pathways by GSVA analyses based on AFP subgroups in tumors (A) and adjacent noncancerous tissues (C). The rows indicate enriched pathways and columns indicate samples. Bar plots of pathway alterations by GSEA analyses at the cut-off value of 400 ng/mL for AFP in tumors (B) and adjacent noncancerous tissues (D). Up-regulated and down-regulated pathways are annotated in orange and green, separately. AFP, alpha-fetoprotein; HCC, Hepatocellular carcinoma; GSEA, gene set enrichment analysis; GSVA, Gene set variation analysis.

metabolism-related pathways were observed in adjacent noncancerous tissues, such as glycogen and histidine metabolism, whereas the opposite trend was observed in tumor tissues (Fig. 2A, C). Similarly, we found relatively increased activities of invasion-related pathways in tumors, such as tight junctions, vascular smooth muscle contraction, collagen biosynthesis and modifying enzymes and regulation of the MAPK pathway, but the opposite results were obtained for adjacent noncancerous tissues (Fig. 2A, C). Since an AFP level equal to 400 ng/mL was pivotal in distinguishing the clinical prognosis and pathway activities, further investigation was performed by stratifying patients into dichotomous categories based on the turning point. Likewise, GSEA revealed the activation of diverse metabolism-related pathways but decreased activity of invasion-related pathways in tumors from patients with a serum AFP level >400 ng/mL (Fig. 2B), whereas opposite patterns for metabolism- and invasion-related pathways were observed in adjacent

noncancerous tissues (Fig. 2D). Altogether, adjacent noncancerous tissues also presented with tumor-related biological activity-dependent behaviors in diverse metabolic and invasion-associated pathways.

Integrative analyses of HCC samples from AFP subgroups

Integrative bioinformatics analyses were performed to explore the molecular features of tumor and adjacent noncancerous tissues at the cut point value of 400 ng/mL for serum AFP levels. A volcano plot showed 151 significantly altered proteins, with 103 upregulated and 48 down-regulated proteins in tumors based on the cutoff point of 400 ng/mL (Fig. 3A), while 95 significantly altered proteins, with 41 upregulated and 54 downregulated proteins, were observed in adjacent noncancerous tissues at the cutoff point of 400 ng/mL (Fig. 3B). Detailed information is summarized in Tables S4 and S5. Notably, the GO analysis

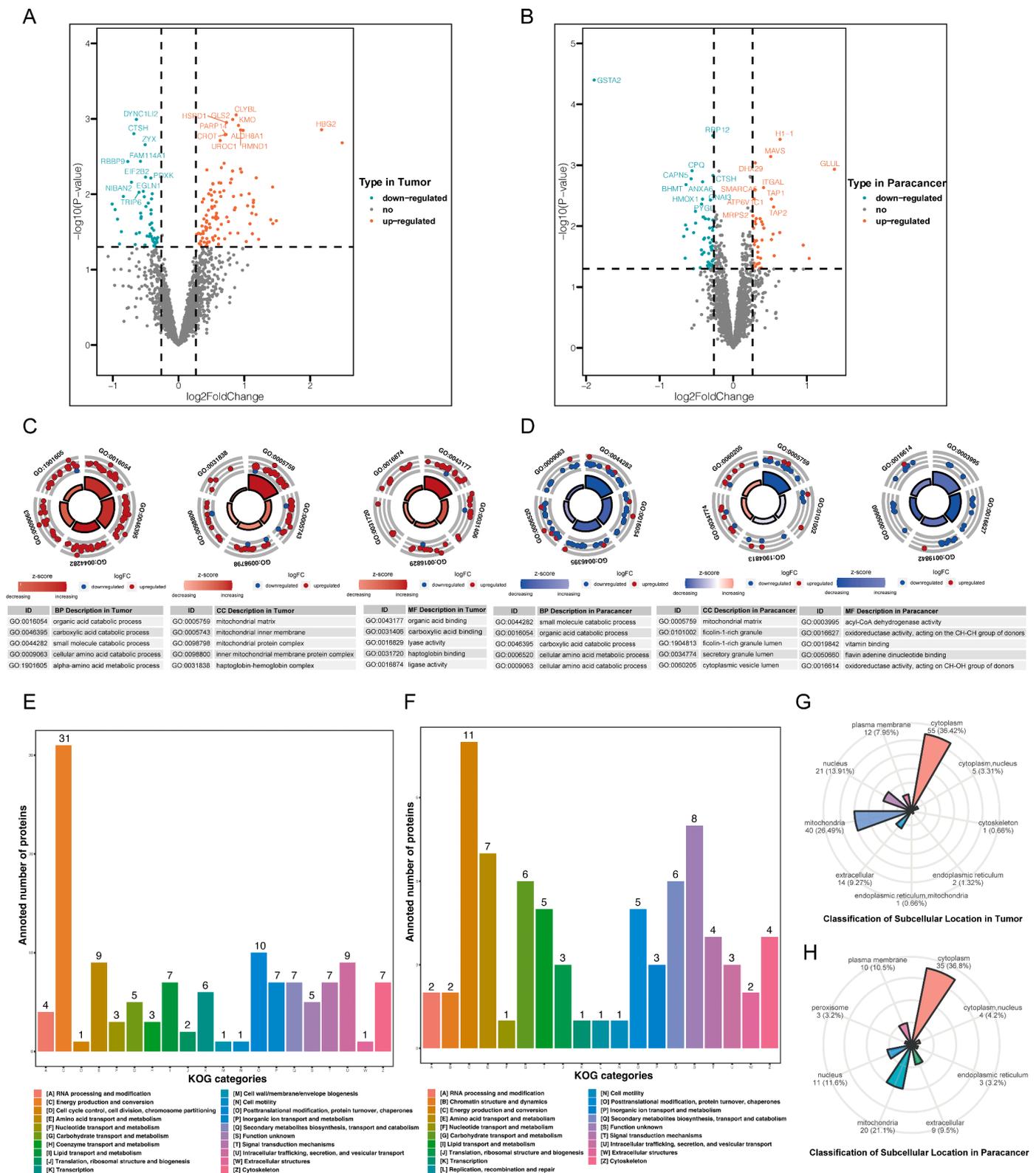


Fig. 3. Integrated functional annotations of differential proteins from AFP >400 and AFP <400 ng/mL groups. Volcano plots of proteomic data in tumor (A) and adjacent noncancerous tissues (B). The top 10 most significant up-regulated and down-regulated proteins were labeled. Circle plots of GO clusters significantly enriched in tumor (C) and adjacent noncancerous tissues (D). Bar plots of differential proteins mapped onto KOG categories in tumors (E) and adjacent noncancerous tissues (F). Windrose plotting of subcellular location in tumor (G) and adjacent noncancerous tissues (H). AFP, alpha-fetoprotein; GO, gene ontology; KOG, EuKaryotic Orthologous Groups.

showed a relatively consistent pattern of differentially expressed proteins in tumor and adjacent noncancerous tissues (Fig. 3C, D). Specifically, we observed univocal significantly enriched GO items, such as multiple common catabolic and metabolic processes in biological processes, congruent mitochondrial-related components in cellular components and coherent molecule binding- and activity-related functions in molecular function modules (Fig. 3C, D). Among the 25 KOG classifications, the differentially expressed proteins in tumors were mainly enriched in energy production and conversion, followed by post-translational modification, protein turnover, and chaperones (Fig. 3E), while the counterparts of adjacent noncancerous tissues were congruently enriched in energy production and conversion (Fig. 3F). Additionally, an analysis of the subcellular locations of differentially expressed proteins annotated by WoLF PSORT software revealed similar compositions but at different proportions for the top 3 subcellular locations “cytoplasm”, “mitochondria” and “nucleus” in tumor and adjacent noncancerous tissues (36.42% vs. 36.8%, 26.49% vs. 21.1%, 13.91% vs. 11.6%, respectively). Collectively, these data indicated similarities in the molecular features of tumor and adjacent noncancerous tissues in an AFP-dependent manner.

Identification of potential drug targets

We used differentially expressed proteins to obtain information on the PPI networks and better understand the regulatory networks responsible for the differences in survival at a cutoff point of 400 ng/mL for serum AFP levels (Fig. 4A, B). Three hundred sixty-five interactions were identified in tumors with the highest combined score residing in AMT-GLDC (0.999), while 104 interactions were identified in adjacent noncancerous tissues with the highest combined score residing in TAP1-TAP2 (0.998). Detailed PPI information is summarized in Tables S6 and S7. The hub protein module identified by the Mcode plugin in Cytoscape revealed 18 genes for tumors (Fig. 4C) and 8 genes for adjacent noncancerous tissues (Fig. 4D).

A literature search was performed on the hub proteins to search for potential drug targets in HCC that have not previously been reported in HCC, and we identified five candidates either approved by the US FDA or undergoing clinical trials, including C1QBP, HSPE1, and GLUD2 for tumors and CHDH and ITGAL for adjacent noncancerous tissues. We further illustrated the transcriptome features of the five candidates by curating and analyzing mRNA data from TCGA, resulting in four significant genes (Fig. S2A-C, E) with three (C1QBP, HSPE1, GLUD2) consistent with our proteome data (Fig. S2A-C, Table S4) but the opposite result was obtained for ITGAL expression (Fig. S2E, Table S5). The 3D structures of the five proteins were determined using AlphaFold to facilitate potential drug development, and the results may provide a deeper understanding of human health and tumor features (Fig. 4E-I).

Expressions of identified proteins were correlated with prognosis in HCC

To further study the clinical translational value of five identified hub proteins as potential treatment targets, we purchased another five tissue microarrays (comprising of 88 paired tumor and adjacent normal tissues and four additional tumor tissues) and compare the protein expression level between tumor and paired paracancerous tissues. IHC results showed C1QBP, HSPE1 and GLUD2 were significantly up-regulated in tumors (All P value < 0.05, Fig. 5A, 5C, 5E), while CHDH and ITGAL were significantly down-regulated in tumors at the protein level (All P value < 0.05, Fig. 5B, 5D), respectively.

To further explore whether these identified proteins could be utilized as potential prognosis biomarkers in HCC, we conducted survival analysis based on subgroups categorized by expression level (Figs. 6, S3). Largely consistent to our previous results, the results of IHC showed four (C1QBP, HSPE1, CHDH, ITGAL) out of five identified proteins (C1QBP, HSPE1, GLUD2, CHDH, ITGAL) could significantly distinguish the overall survival of HCC patients (Log-rank P value were 0.028, 0.043,

0.013 and 0.014, respectively; Fig. 6A, C-E). Intriguingly, despite none of these identified proteins reached statistical significance, there was a relatively clear trend that the high expressions of HSPE1 in tumor and CHDH in adjacent noncancerous tissues were presented with worse recurrence-free survival (Log-rank P value were 0.086 and 0.052, respectively; Fig. S3C-D).

Discussion

AFP levels showed great prognostic value in patients with HCC to guide therapy based on completed phase 3 trials [6–8]. However, the phenomenon of differential prognosis of AFP subgroups remains unclear at molecular level. In this study, TMT-based quantitative proteomic analyses were performed on 12 patients with hepatitis B virus-related HCC, and tumor and paracancerous tissues were analyzed separately after patient were stratified into AFP subgroups. We found that an AFP level equal to 400 ng/mL was a pivotal turning point with a distinct difference in survival, congruent with previous studies [24,26]. The pathway enrichment analyses revealed that metabolic and invasion pathways were mainly altered in tumors. Notably, multiple metabolic pathways were enhanced in tumors from patients with AFP levels > 400 ng/mL compared with those with AFP levels < 400 ng/mL, indicating compensatory increases in metabolism to meet energetic demand that were potentially associated with HCC aggressiveness. Recently, studies have documented aberrant expression patterns among patients with diverse gastrointestinal cancers, including gastric cancer and colorectal cancer [27–29]. Here, we showed that adjacent noncancerous tissues also presented with aberrations in patients with HCC. Specifically, diverse metabolism-associated pathways (e.g., glycine, serine and threonine metabolism and phenylalanine metabolism) presented the opposite pattern in activity at a cutoff point of an AFP level equal to 400 ng/mL. This result may be explained by the potential benefits of tumors in energy utilization to remodel the utilization modes of adjacent noncancerous tissues. Considering the tight correlations between cancer metabolism, tumor proliferation and survival, as well as the discrepancies in metabolic reprogramming patterns in tumor and adjacent noncancerous tissues, tailored treatment strategies may be useful for different tissue types in patients with HBV-related HCC. Integrative annotation analyses of differentially expressed proteins, GO terms, KOG classifications and subcellular locations further indicated similarities in the molecular features of tumor and adjacent noncancerous tissues in an AFP-dependent manner, indicating that adjacent noncancerous tissues per se possessed tumor features.

Our work identified three potential druggable targets in tumors and two in adjacent noncancerous tissues that were enriched in a metabolic-centered manner but were not previously reported in HCC. All of the targets were reported to be correlated with tumor formation and/or progression. Specifically, C1QBP, a ubiquitous evolutionarily conserved acidic protein, was initially identified as a complement component C1q-binding protein [30]. Recent studies have shown that C1QBP is pivotal for diverse phenotypes, including the immune response, apoptosis and DNA damage repair [31–33]. HSPE1 mainly functions as a chaperone mediating protein folding with HSPD1 [34]. Surprisingly, HSPE1 not only engages in modulating the proapoptotic phenotype and cell signaling network but also promotes human Treg cell differentiation in a Treg subtype-dependent manner, suggesting potential immune-oncology therapeutic efficacy [35–37]. GLUD2, which was originally identified as a glutamate dehydrogenase (GDH)-coding gene, was deemed an evolutionarily novel version of GLUD1 [38]. Generally, GLUD2 mainly modulates glucose homeostasis by tuning fasting serum insulin levels and acts on the tricarboxylic acid cycle, potentially facilitating lipid biosynthesis [39,40]. However, a recent study revealed an ammonia-centered recycling mode mediated by tuning GDH to promote proliferation in breast cancer, which indicated multiple potential tumor metabolic-modulating roles of the GLUD2 protein [41]. CHDH is known to catalyze the dehydrogenation of choline in mitochondria. Preclinical

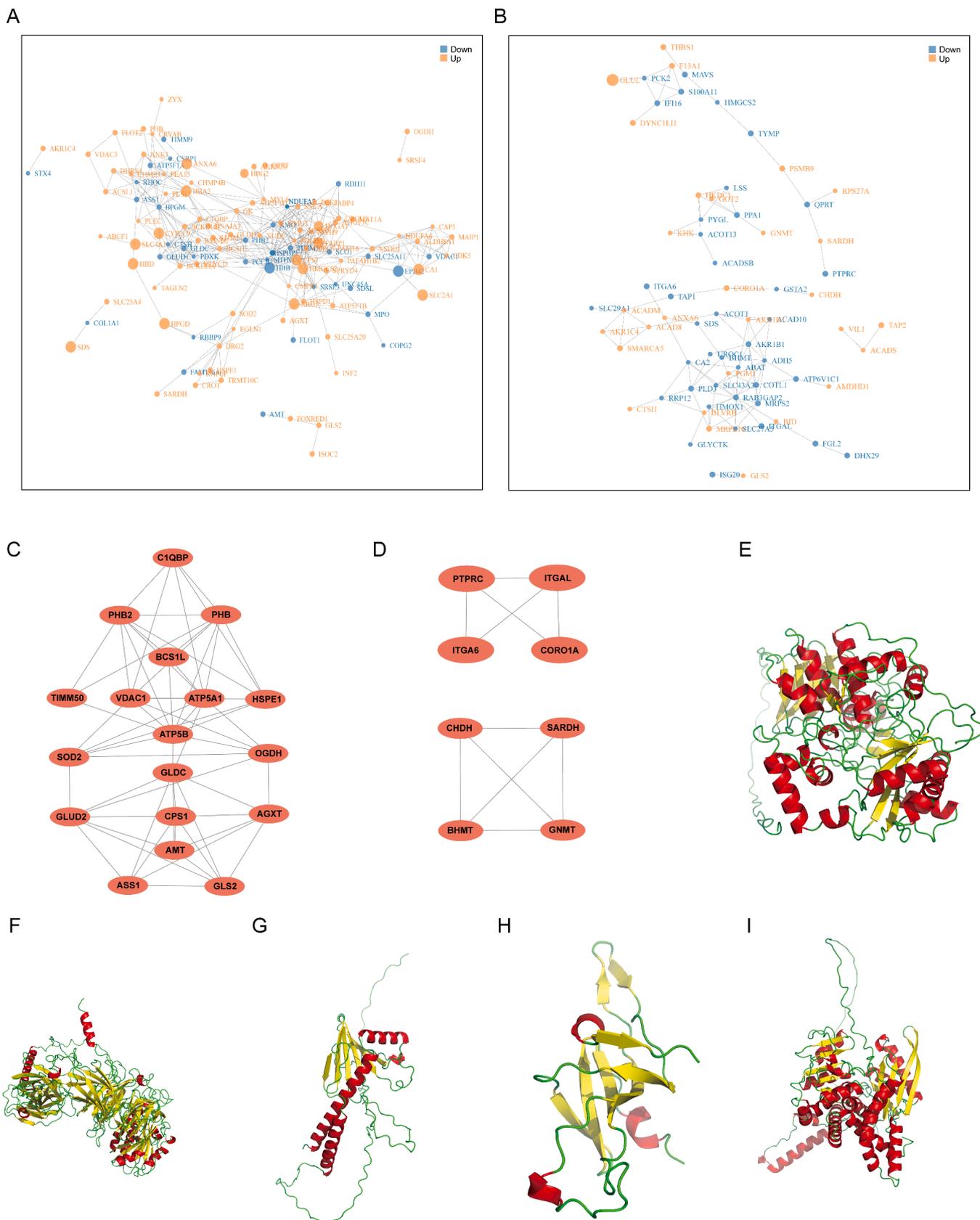


Fig. 4. Identification of potential drug targets in HBV-related HCC. PPI network of differentially expressed proteins in tumor (A) and adjacent noncancerous tissues (B) at cut-off value of 400 ng/mL for AFP with node size representing fold changes. Hub proteins of the PPI by Mcode plugin in tumor (C) and adjacent noncancerous tissues (D). Predicted three-dimensional structures of CHDH (E), ITGAL (F) in adjacent noncancerous tissues and C1QBP (G), HSPE1 (H), GLUD2 (I) in tumor tissues predicted by AlphaFold. AFP, alpha-fetoprotein; HBV, hepatitis B virus; HCC, Hepatocellular carcinoma; PPI, protein-protein interaction.

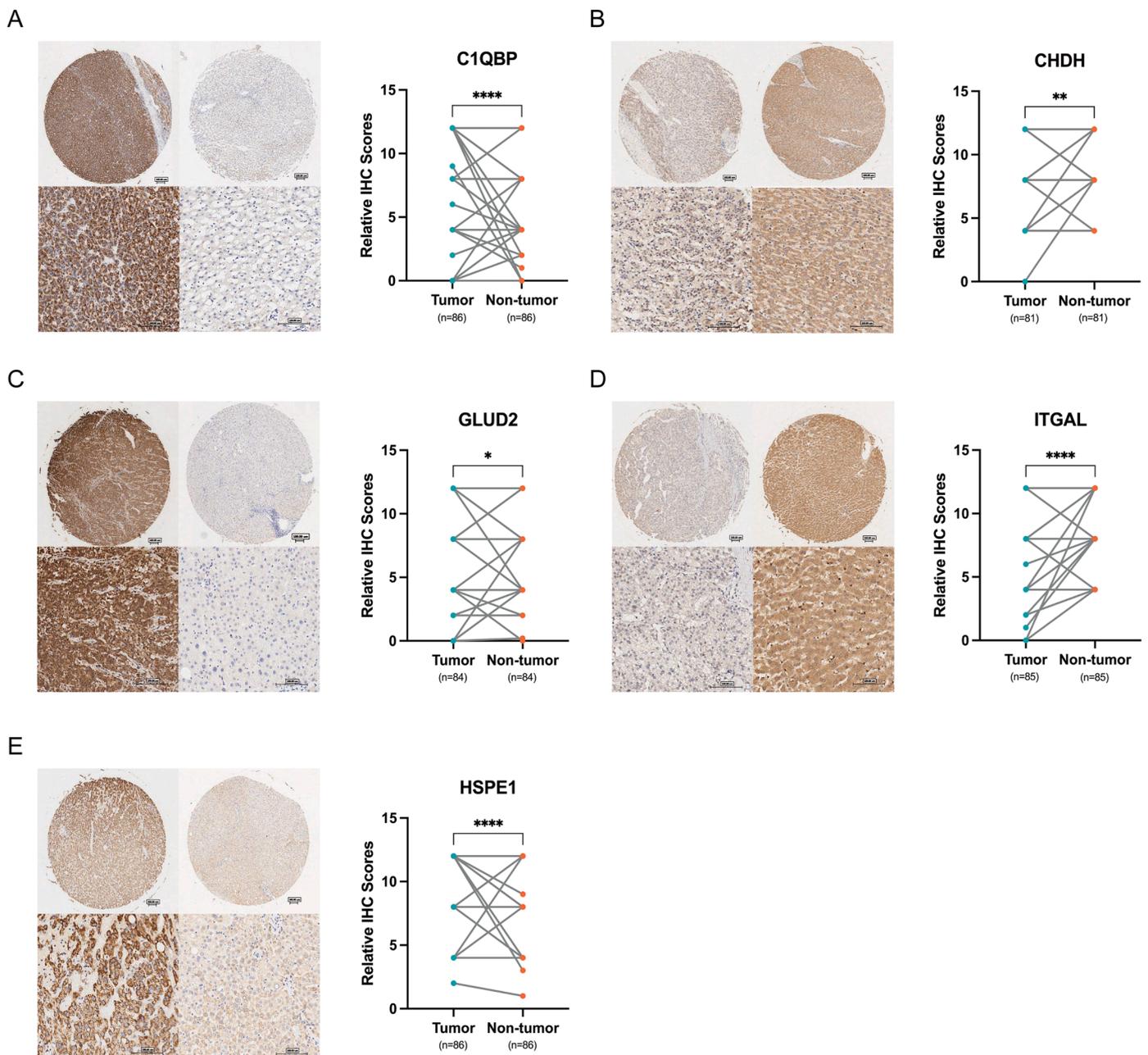


Fig. 5. Differential expression levels of five identified hub proteins by immunohistochemistry. The upper and lower in left panel were magnified by 5x and 20x, separately. The left and right in left panel were representative immunohistochemistry results of tumor and adjacent noncancerous tissues. The right panels displayed comparison results of relative immunohistochemistry scores between tumor and adjacent noncancerous tissues in HCC tissue microarrays at protein level for C1QBP (A), CHDH (B), GLUD2(C), ITGAL (D) and HSPE1 (E), separately. Scale bar, 100 μ m. Wilcoxon rank sum test was used for comparison statistics. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and **** $p < 0.0001$. $P < 0.05$ was considered statistical significant. Scale bar 100 μ m. HCC, Hepatocellular carcinoma; Immunohistochemistry, IHC.

research showed a crucial role for CHDH in the mitophagy phenotype, while clinical studies confirmed the predictive power of CHDH for the tamoxifen monotherapy response and recurrence in patients with breast cancer [42–44]. ITGAL, the gene encoding CD11a, interacts with the beta 2 chain (ITGB2) to form integrin lymphocyte function-associated antigen-1, which regulates the intercellular adhesion phenotype and lymphocyte costimulatory signaling [45]. According to mechanistic studies, T cells modulate the function of ITGAL by altering DNA methylation in the promoter region, which indicated a general function in cellular and humoral immunity [46]. Additionally, the ITGAL-based six-gene model showed a remarkable degree of discrimination in patients with castration-resistant prostate cancer [47].

Intriguingly, the profiles of five candidates at the transcriptional level showed largely consistent mRNA and protein expression (C1QBP, HSPE1, and GLUD2), but the opposite results were obtained for ITGAL expression, which might be partially explained by potential post-transcriptional modification mechanisms. In the background of precision and translational medicine, we further explored the prognostic value of five identified protein targets. Four (C1QBP, HSPE1, CHDH and ITGAL) out of five (C1QBP, HSPE1, GLUD2, CHDH and ITGAL) differentially expressed proteins were significantly and negatively correlated with overall survival rate, while high expressions of HSPE1 in tumor and CHDH in paracancerous tissues showed trend of shorter recurrence-free survival, indicating great potential as prognostic biomarkers. However,

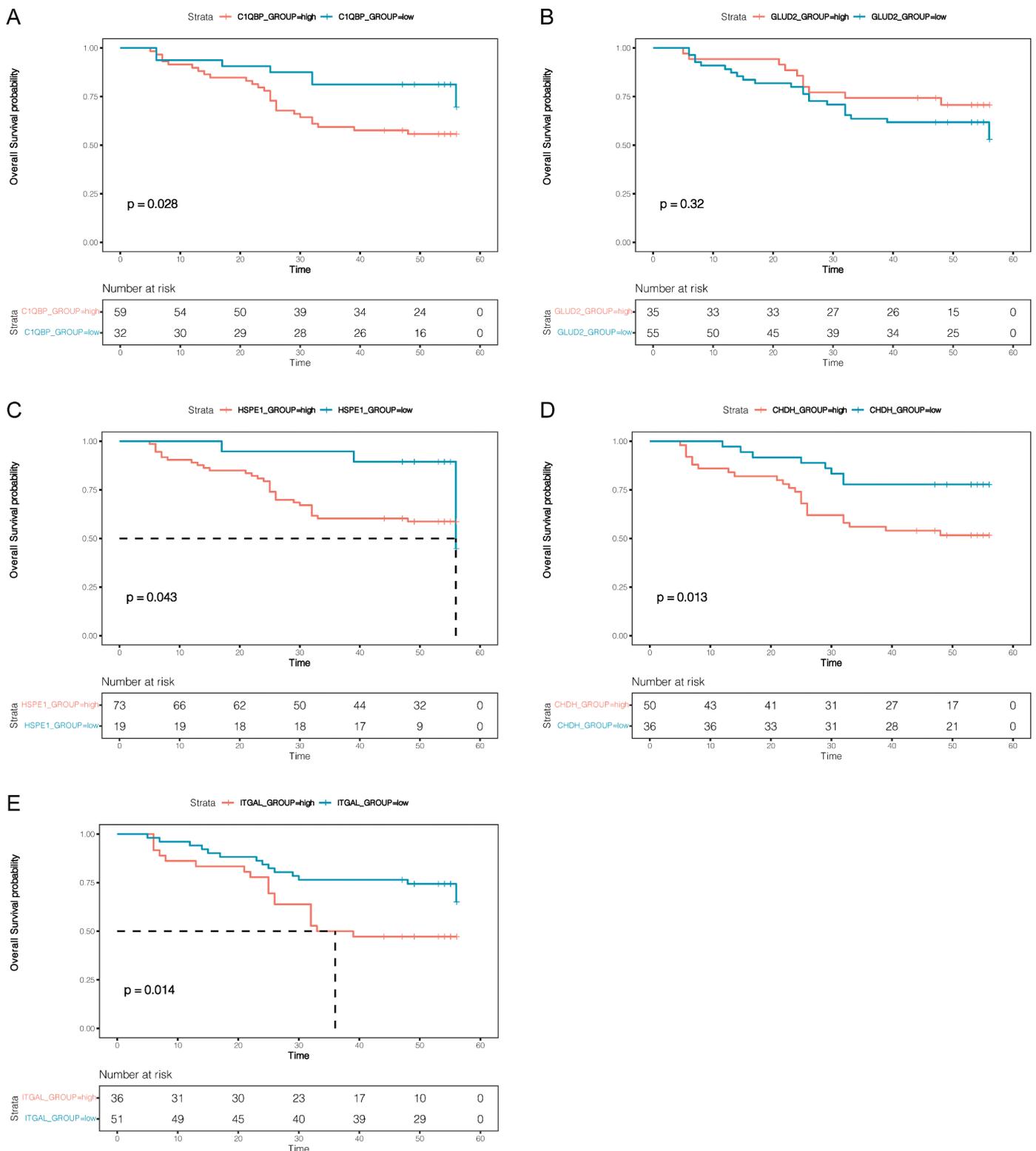


Fig. 6. Kaplan–Meier analysis of OS in patients with low and high protein level in HCC tissue microarray according to IHC scores. Log-rank $P < 0.05$ was considered statistical significant. HCC, Hepatocellular carcinoma; OS, overall survival; Immunohistochemistry, IHC.

additional multi-center, prospective, large sample cohorts are warrant to verify the prognostic efficacy. Taken together, these results indicated the great therapeutic potential of the four candidates in HCC.

The present study has some limitations. First, this study was conducted at a single center with a relatively small sample size (twelve paired tumor and adjacent noncancerous tissues) due to difficulties in sample collection that may lead to possible sample bias and diminish the

reliability of the conclusions. Second, the study design was mainly based on a proteomics technique with limited consideration of molecular alternations at other omics levels (e.g., single-cell sequencing technology that possesses single-cell resolution to characterize the functional state over traditional bulk sequencing analysis technologies) [48,49]. Further multicenter studies using multiomics techniques to investigate larger samples are required to validate our findings. Despite the

forementioned limitations, our results challenged the biological plausibility of describing adjacent noncancerous tissues as biologically normal components at the protein level in patients with HBV-related HCC.

In summary, our current study provides a new understanding of the aberrant expression pattern in adjacent noncancerous tissues from patients with HCC based on AFP subgroups. We revealed that a serum AFP level equal to 400 ng/mL was a pivotal turning point in biological functions both in tumor and adjacent noncancerous tissues. Our study identified four druggable targets utilized as prognostic markers and potential therapeutic targets that were mainly involved in metabolism and immunity and may provide opportunities for exploring innovative treatment targets on the background of precision and translational medicine.

Statement

The study was approved by the Research Ethics Committee of The First Affiliated Hospital, Zhejiang University School of Medicine, and written informed consent was obtained from each patient. The study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki (6th revision, 2008) as reflected in a priori approval by the institution's human research committee.

CRedit authorship contribution statement

Xuyong Wei: Conceptualization, Data curation, Writing – original draft, Writing – review & editing, Funding acquisition. **Renyi Su:** Conceptualization, Data curation, Formal analysis, Data curation, Writing – original draft. **Mengfan Yang:** Writing – review & editing. **Binhua Pan:** Data curation, Formal analysis, Visualization. **Jun Lu:** Formal analysis, Visualization. **Hanchao Lin:** Writing – review & editing. **Wenzhi Shu:** Writing – review & editing. **Rui Wang:** Formal analysis, Visualization. **Xiao Xu:** Conceptualization, Writing – review & editing, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was supported by the National Key Research and Development Program of China (No. 2021YFA1100500), the Major Research Plan of the National Natural Science Foundation of China (No.92159202), National High-tech R&D Program of China (863 Program) (No.2012AA020204), Youth Program of National Natural Science Foundation of China (No.81702858); Youth Program of National Natural Science Foundation of Zhejiang Province (No.LQ17H160006).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tranon.2022.101422.

References

- [1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2020, *CA Cancer J. Clin.* 70 (1) (2020) 7–30, <https://doi.org/10.3322/caac.21590>. Jan.
- [2] J.C. Nault, A. Villanueva, Biomarkers for hepatobiliary cancers, *Hepatology* 73 (Suppl 1) (2021) 115–127, <https://doi.org/10.1002/hep.31175>. Jan.
- [3] A. Villanueva, Hepatocellular carcinoma, *N. Engl. J. Med.* 380 (15) (2019) 1450–1462, <https://doi.org/10.1056/NEJMr1713263>. Apr 11.
- [4] G.J. Mizejewski, Biological role of alpha-fetoprotein in cancer: prospects for anticancer therapy, *Expert Rev. Anticancer Ther.* 2 (6) (2002) 709–735, <https://doi.org/10.1586/14737140.2.6.709>. Dec.
- [5] P.R. Galle, F. Foerster, M. Kudo, et al., Biology and significance of alpha-fetoprotein in hepatocellular carcinoma, *Liver Int.* 39 (12) (2019) 2214–2229, <https://doi.org/10.1111/liv.14223>. Dec.
- [6] S.L. Chan, F.K. Mo, P.J. Johnson, et al., New utility of an old marker: serial alpha-fetoprotein measurement in predicting radiologic response and survival of patients with hepatocellular carcinoma undergoing systemic chemotherapy, *J. Clin. Oncol.* 27 (3) (2009) 446–452, <https://doi.org/10.1200/JCO.2008.18.8151>. Jan 20.
- [7] M. Kudo, R.S. Finn, S. Qin, et al., Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial, *Lancet* 391 (10126) (2018) 1163–1173, [https://doi.org/10.1016/S0140-6736\(18\)30207-1](https://doi.org/10.1016/S0140-6736(18)30207-1). Mar 24.
- [8] J.M. Llovet, S. Ricci, V. Mazzaferro, et al., Sorafenib in advanced hepatocellular carcinoma, *N. Engl. J. Med.* 359 (4) (2008) 378–390, <https://doi.org/10.1056/NEJMoa0708857>. Jul 24.
- [9] T. Chen, X. Dai, J. Dai, et al., AFP promotes HCC progression by suppressing the HuR-mediated Fas/FADD apoptotic pathway, *Cell Death Dis.* 11 (10) (2020) 822, <https://doi.org/10.1038/s41419-020-03030-7>. Oct 2.
- [10] R. Gao, C. Cai, J. Gan, et al., miR-1236 down-regulates alpha-fetoprotein, thus causing PTEN accumulation, which inhibits the PI3K/Akt pathway and malignant phenotype in hepatoma cells, *Oncotarget* 6 (8) (2015) 6014–6028, <https://doi.org/10.18632/oncotarget.3338>. Mar 20.
- [11] I. Gotsman, D. Israeli, R. Alper, E. Rabbani, D. Engelhardt, Y. Ilan, Induction of immune tolerance toward tumor-associated-antigens enables growth of human hepatoma in mice, *Int. J. Cancer* 97 (1) (2002) 52–57, <https://doi.org/10.1002/ijc.1576>. Jan 1.
- [12] Y. Lu, M. Zhu, W. Li, et al., Alpha fetoprotein plays a critical role in promoting metastasis of hepatocellular carcinoma cells, *J. Cell. Mol. Med.* 20 (3) (2016) 549–558, <https://doi.org/10.1111/jcmm.12745>. Mar.
- [13] F. Chen, D.S. Chandrashekar, S. Varambally, C.J. Creighton, Pan-cancer molecular subtypes revealed by mass-spectrometry-based proteomic characterization of more than 500 human cancers, *Nat. Commun.* 10 (1) (2019) 5679, <https://doi.org/10.1038/s41467-019-13528-0>. Dec 12.
- [14] J. Shen, F. Jiang, Y. Yang, et al., 14-3-3eta is a novel growth-promoting and angiogenic factor in hepatocellular carcinoma, *J. Hepatol.* 65 (5) (2016) 953–962, <https://doi.org/10.1016/j.jhep.2016.05.017>. Nov.
- [15] S.Y. Wen, L.N. Zhang, X.M. Yang, et al., LRG1 is an independent prognostic factor for endometrial carcinoma, *Tumour Biol.* 35 (7) (2014) 7125–7133, <https://doi.org/10.1007/s13277-014-1953-6>. Jul.
- [16] X.Y. Huang, A.W. Ke, G.M. Shi, et al., alphaB-crystallin complexes with 14-3-3zeta to induce epithelial-mesenchymal transition and resistance to sorafenib in hepatocellular carcinoma, *Hepatology* 57 (6) (2013) 2235–2247, <https://doi.org/10.1002/hep.26255>. Jun.
- [17] Q. Gao, H. Zhu, L. Dong, et al., Integrated proteogenomic characterization of HBV-related hepatocellular carcinoma, *Cell* 179 (2) (2019) 561–577, <https://doi.org/10.1016/j.cell.2019.08.052>. Oct 3e22.
- [18] J.R. Wisniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis, *Nat. Methods* 6 (5) (2009) 359–362, <https://doi.org/10.1038/nmeth.1322>. May.
- [19] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, *OMICS* 16 (5) (2012) 284–287, <https://doi.org/10.1089/omi.2011.0118>. May.
- [20] A. Subramanian, P. Tamayo, V.K. Mootha, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, *Proc. Natl. Acad. Sci. U. S. A.* 102 (43) (2005) 15545–15550, <https://doi.org/10.1073/pnas.0506580102>. Oct 25.
- [21] Y. Tao, C. Tian, N. Verma, et al., Recovering intrinsic fragmental vibrations using the generalized subsystem vibrational analysis, *J. Chem. Theory Comput.* 14 (5) (2018) 2558–2569, <https://doi.org/10.1021/acs.jctc.7b01171>. May 8.
- [22] K. Tunyasuvunakool, J. Adler, Z. Wu, et al., Highly accurate protein structure prediction for the human proteome, *Nature* 596 (7873) (2021) 590–596, <https://doi.org/10.1038/s41586-021-03828-1>. Aug.
- [23] L. Kadalayil, R. Benini, L. Pallan, et al., A simple prognostic scoring system for patients receiving transarterial embolisation for hepatocellular cancer, *Ann. Oncol.* 24 (10) (2013) 2565–2570, <https://doi.org/10.1093/annonc/mdt247>. Oct.
- [24] S. Merani, P. Majno, N.M. Kneteman, et al., The impact of waiting list alpha-fetoprotein changes on the outcome of liver transplant for hepatocellular carcinoma, *J. Hepatol.* 55 (4) (2011) 814–819, <https://doi.org/10.1016/j.jhep.2010.12.040>. Oct.
- [25] X. Xu, D. Lu, Q. Ling, et al., Liver transplantation for hepatocellular carcinoma beyond the Milan criteria, *Gut* 65 (6) (2016) 1035–1041, <https://doi.org/10.1136/gutjnl-2014-308513>. Jun.
- [26] N. Brau, R.K. Fox, P. Xiao, et al., Presentation and outcome of hepatocellular carcinoma in HIV-infected patients: a U.S.-Canadian multicenter study, *J. Hepatol.* 47 (4) (2007) 527–537, <https://doi.org/10.1016/j.jhep.2007.06.010>. Oct.
- [27] D. Aran, R. Camarda, J. Odegaard, et al., Comprehensive analysis of normal adjacent to tumor transcriptomes, *Nat. Commun.* 8 (1) (2017) 1077, <https://doi.org/10.1038/s41467-017-01027-z>. Oct 20.
- [28] S. Russi, G. Calice, V. Ruggieri, et al., Gastric normal adjacent mucosa versus healthy and cancer tissues: distinctive transcriptomic profiles and biological features, *Cancers* 11 (9) (2019), <https://doi.org/10.3390/cancers11091248> (Basel) Aug 26.
- [29] R. Sanz-Pamplona, A. Berenguer, D. Cordero, et al., Aberrant gene expression in mucosa adjacent to tumor reveals a molecular crosstalk in colon cancer, *Mol. Cancer* 13 (2014) 46, <https://doi.org/10.1186/1476-4598-13-46>. Mar 5.

- [30] S.B. Storrs, W.P. Kolb, R.N. Pinckard, M.S. Olson, Characterization of the binding of purified human Clq to heart mitochondrial membranes, *J. Biol. Chem.* 256 (21) (1981) 10924–10929. Nov 10.
- [31] Y. Bai, W. Wang, S. Li, et al., ClQBP promotes homologous recombination by stabilizing MRE11 and controlling the assembly and activation of MRE11/RAD50/NBS1 complex, *Mol. Cell* 75 (6) (2019) 1299–1314, <https://doi.org/10.1016/j.molcel.2019.06.023>. Sep 19e6.
- [32] K. Itahana, Y. Zhang, Mitochondrial p32 is a critical mediator of ARF-induced apoptosis, *Cancer Cell* 13 (6) (2008) 542–553, <https://doi.org/10.1016/j.ccr.2008.04.002>. Jun.
- [33] J. Zuniga, I. Buendia-Roldan, Y. Zhao, et al., Genetic variants associated with severe pneumonia in A/H1N1 influenza infection, *Eur. Respir. J.* 39 (3) (2012) 604–610, <https://doi.org/10.1183/09031936.00020611>. Mar.
- [34] G. Legname, G. Fossati, G. Gromo, N. Monzini, F. Marucci, D. Modena, Expression in *Escherichia coli*, purification and functional activity of recombinant human chaperonin 10, *FEBS Lett.* 361 (2-3) (1995) 211–214, [https://doi.org/10.1016/0014-5793\(95\)00184-b](https://doi.org/10.1016/0014-5793(95)00184-b). Mar 20.
- [35] A.M. Czarnecka, C. Campanella, G. Zummo, F. Cappello, Heat shock protein 10 and signal transduction: a "capsula eburnea" of carcinogenesis? *Cell Stress Chaperones* 11 (4) (2006) 287–294, <https://doi.org/10.1379/csc-200.1>. Winter.
- [36] A.F. Kovacs, N. Fekete, L. Turiak, et al., Unravelling the role of trophoblastic-derived extracellular vesicles in regulatory T cell differentiation, *Int. J. Mol. Sci.* 20 (14) (2019), <https://doi.org/10.3390/ijms20143457>. Jul 14.
- [37] K.M. Lin, J.M. Hollander, V.Y. Kao, B. Lin, L. Macpherson, W.H. Dillmann, Myocyte protection by 10 kD heat shock protein (Hsp10) involves the mobile loop and attenuation of the Ras GTP-ase pathway, *FASEB J.* 18 (9) (2004) 1004–1006, <https://doi.org/10.1096/fj.03-0348fje>. Jun.
- [38] F. Burki, H. Kaessmann, Birth and adaptive evolution of a hominoid gene that supports high neurotransmitter flux, *Nat. Genet.* 36 (10) (2004) 1061–1063, <https://doi.org/10.1038/ng1431>. Oct.
- [39] Q. Li, S. Guo, X. Jiang, et al., Mice carrying a human *GLUD2* gene recapitulate aspects of human transcriptome and metabolome development, *Proc. Natl. Acad. Sci. U. S. A.* 113 (19) (2016) 5358–5363, <https://doi.org/10.1073/pnas.1519261113>. May 10.
- [40] Z. Petraki, S. Droubogiannis, K. Mylonaki, G. Chlouverakis, A. Plaitakis, C. Spanaki, Transgenic expression of the positive selected human *GLUD2* gene improves in vivo glucose homeostasis by regulating basic insulin secretion, *Metabolism* 100 (2019), 153958, <https://doi.org/10.1016/j.metabol.2019.153958>. Nov.
- [41] J.B. Spinelli, H. Yoon, A.E. Ringel, S. Jeanfavre, C.B. Clish, M.C. Haigis, Metabolic recycling of ammonia via glutamate dehydrogenase supports breast cancer biomass, *Science* 358 (6365) (2017) 941–946, <https://doi.org/10.1126/science.aam9305>. Nov 17.
- [42] X.J. Ma, S.G. Hilsenbeck, W. Wang, et al., The *HOXB13:LL17BR* expression index is a prognostic factor in early-stage breast cancer, *J. Clin. Oncol.* 24 (28) (2006) 4611–4619, <https://doi.org/10.1200/JCO.2006.06.6944>. Oct 1.
- [43] X.J. Ma, Z. Wang, P.D. Ryan, et al., A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen, *Cancer Cell* 5 (6) (2004) 607–616, <https://doi.org/10.1016/j.ccr.2004.05.015>. Jun.
- [44] S. Park, S.G. Choi, S.M. Yoo, J.H. Son, Y.K. Jung, Choline dehydrogenase interacts with SQSTM1/p62 to recruit LC3 and stimulate mitophagy, *Autophagy* 10 (11) (2014) 1906–1920, <https://doi.org/10.4161/auto.32177>.
- [45] A.L. Corbi, R.S. Larson, T.K. Kishimoto, T.A. Springer, C.C. Morton, Chromosomal location of the genes encoding the leukocyte adhesion receptors LFA-1, Mac-1 and p150,95. Identification of a gene cluster involved in cell adhesion, *J. Exp. Med.* 167 (5) (1988) 1597–1607, <https://doi.org/10.1084/jem.167.5.1597>. May 1.
- [46] Q. Lu, D. Ray, D. Gutsch, B. Richardson, Effect of DNA methylation and chromatin structure on *ITGAL* expression, *Blood* 99 (12) (2002) 4503–4508, <https://doi.org/10.1182/blood.v99.12.4503>. Jun 15.
- [47] R.W. Ross, M.D. Galsky, H.I. Scher, et al., A whole-blood RNA transcript-based prognostic model in men with castration-resistant prostate cancer: a prospective study, *Lancet Oncol.* 13 (11) (2012) 1105–1113, [https://doi.org/10.1016/S1470-2045\(12\)70263-2](https://doi.org/10.1016/S1470-2045(12)70263-2). Nov.
- [48] Q. Song, J. Su, L.D. Miller, W. Zhang, scLM: automatic detection of consensus gene clusters across multiple single-cell datasets, *Genom. Proteom. Bioinform.* 19 (2) (2021) 330–341, <https://doi.org/10.1016/j.gpb.2020.09.002>. Apr.
- [49] Q. Song, J. Su, W. Zhang, scGCN is a graph convolutional networks algorithm for knowledge transfer in single cell omics, *Nat. Commun.* 12 (1) (2021) 3826, <https://doi.org/10.1038/s41467-021-24172-y>. Jun 22.