Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Hub Gene Identification and Prognostic Model Establishment for Patients with HBV-related Hepatocellular Carcinoma

Lianmei Wang
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing
Jing Liu
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing
Zhong Xian
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing
Jingzhuo Tian
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing
Chunying Li
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing
Yushi Zhang
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing
Suyan Liu
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing
Jiayin Han
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing
Chen Pan
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing
Yan Yi
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing
Yong Zhao
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing
Hongbing Zhang
Institute of Basic Medical Sciences and School of Basic Medicine, Graduate School of Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing
Aihua Liang (✉ ahliang@icmm.ac.cn)
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing

---

Research Article

Keywords: HBV-related hepatocellular carcinoma (HBV-HCC), Prognosis, Protein–protein interaction (PPI), Hub genes, Machine learning

DOI: https://doi.org/10.21203/rs.3.rs-269454/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Hepatocellular carcinoma (HCC) is associated with poor 5-year survival. Chronic infection with hepatitis B virus (HBV) contributes to ~ 50% of HCC cases. Establishment of a prognostic model is pivotal for clinical therapy of HBV-related HCC (HBV–HCC). We downloaded gene-expression profiles from Gene expression omnibus (GEO) datasets with HBV-HCC patients and the corresponding controls. Integration of these differentially expressed genes (DEGs) was achieved with the Robustrankaggreg (RRA) method. DEGs functional analyses and pathway analyses was performed using the Gene ontology (GO) database, and the Kyoto encyclopedia of genes and genomes (KEGG) database respectively. DNA topoisomerase II alpha (TOP2A), Disks large-associated protein 5 (DLGAP5), RAD51 associated protein 1 (RAD51AP1), ZW10 interactor (ZWINT), BUB1 mitotic checkpoint serine/threonine kinase B (BUB1B), Cyclin B1 (CCNB1), Forkhead box M1 (FOXM1), Cyclin B2 (CCNB2), Aurora kinase A (AURKA), and Cyclin-dependent kinase 1 (CDK1) were identified as the top-ten hub genes. These hub-genes were verified by the Liver cancer-riken, JP project from international cancer genome consortium (ICGC-LIRI-JP) project, The Cancer genome atlas (TCGA) HCC cohort, and Human protein profiles dataset. FOXM1 and CDK1 were found to be prognostic-related molecules for HBV-HCC patients. The expression patterns of FOXM1 and CDK1were consistently in human and mouse. Furthermore, a nomogram model based on histology grade, pathology stage, sex and, expression of FOXM1 and CDK1 was built to predict the prognosis for HBV–HCC patients. The nomogram model could be used to predict the prognosis of HBV-HCC cases.

Introduction

HCC is the second most lethal cancer worldwide, and carries 5-year survival of 18% [1]. About 3.5% of the world population suffers chronic infection with HBV [2], and ~ 50% of such cases lead to HCC [3]. In addition, male sex, cirrhosis, diabetes mellitus, aflatoxin exposure, alcohol intake, and tobacco consumption are risk factors for HCC [3–5].

Sodium taurocholate cotransporting polypeptide is the receptor for HBVs entering hepatocytes [6]. Long-term HBV infection, augmentation of HBV replication, genome integration of HBV, HBV mutants, and HBV-encoded oncoproteins work in HBV-HCC [3]. Universal vaccination of newborns against the HBV has reduced the prevalence of HCC dramatically [7]. Antiviral therapy with nucleoside analogs or nucleotide analogs suppresses HBV replication significantly in patients with chronic HBV infection [2]. Sorafenib, lenvatinib, regorafenib, and HBV-related T-cell immunotherapies are potential therapeutic methods for HBV-HCC [2,8–10]. Early detection is important for cancer treatment, so biomarker detection is pivotal for HBV-HCC prevention.

Bioinformatics analysis integrated with data on RNA expression and clinical characteristics are tools to predict the prognosis of diseases. Data on RNA expression can be obtained with next-generation sequencing and microarrays [11]. GEO is a repository of data on high-throughput gene expression and

hybridization arrays and chips. To avoid inconsistent results due to different experiments, GEO datasets can be incorporated with the RRA method for comparison of sequenced genes [12].

In the present study, we downloaded three datasets from GEO datasets. We chose 253 HBV-related samples from these datasets (128 control samples and 125 tumor samples). Microarray data were analyzed with R (https://www.r-project.org/). DEGs from these datasets were integrated using the RRA method (https://cran.rstudio.com/bin/windows/contrib/3.5/RobustRankAggreg_1.1.zip). DEGs were evaluated by the database for annotation, visualization and integrated discovery (DAVID; https://david.ncifcrf.gov/). The GO database (www.geneontology.org/) focuses on the function of the genes and gene products, and the KEGG database (www.genome.jp/kegg/pathway.html) focuses on enrichment of pathways: both were employed for DEGs. We pinpointed hub genes with Cytoscape (https://cytoscape.org/) from DEGs analyzed by the search tool for the retrieval of interacting genes/proteins (STRING) database (https://string-db.org/). Then hub genes were verified with ICGC (https://dcc.icgc.org/)-LIRI-JP, TCGA (https://portal.gdc.cancer.gov/) HCC project, and Human protein profiles dataset (www.proteinatlas.org). Furthermore, a prognosis model was constructed based on biomarkers and clinical factors. The nomogram model we present might be applied as a predictive tool for the prognosis of HBV–HCC patients clinically.

Results

Screening of DEGs in HBV–HCC patients.

The HBV–HCC datasets GSE55092, GSE62232, and GSE121248 were normalized, and the DEGs were selected by corrected $P < 0.05$ and $|\log FC| > 2$ with 'limma'. The GSE55092 dataset involved 1289 DEGs, 496 of which showed upregulated expression and 793 exhibited downregulated expression (Fig. 1a). The GSE62232 dataset involved 1583 DEGs, of which 768 showed upregulated expression and 815 exhibited downregulated expression (Fig. 1b). The GSE121248 dataset involved 552 DEGs, of which 165 showed upregulated expression and 387 exhibited downregulated expression (Fig. 1c). The cluster heatmaps of the top-100 genes were shown in Fig. 1.

The expression matrix of these three datasets was obtained using 'limma', and sorted by log FC. The integrated DEGs were screened through 'RRA' according to corrected $P < 0.05$ and $|\log FC| > 1$. A total of 135 DEGs were identified, 57 of which showed upregulated expression, and 78 of which exhibited downregulated expression (Supplementary Table S1 online). We summarized the top-20 genes with upregulated and downregulated expression in the heatmap shown as Supplementary Fig. S1

Enrichment analyses using GO and KEGG databases. Functional annotation of integrated DEGs was analyzed using DAVID. Functional annotation was based on three categories: Biological process (BP), Molecular function (MF), and Cell component (CC). Annotation of genes with upregulated and downregulated expression was listed in Supplementary Table S2 and Table S3, respectively.

The top-five MF of genes with upregulated expression were “ATP binding”, “Protein kinase activity”, “Protein kinase binding”, “ATP-dependent microtubule motor activity”, and “Peptidase inhibitor activity” (Fig. 2A). The top-five CC of genes with upregulated expression were “Cytoplasm”, “Nucleus”, “Membrane”, “Midbody”, and “Centrosome” (Fig. 2A). The top-five BP of genes with upregulated expression were “Cell division”, “Mitotic nuclear division”, “Positive regulation of apoptotic process”, “Mitotic spindle organization”, and “Cytokinesis” (Fig. 2A). The genes with upregulated expression were shown in Fig. 2B.

The top-five MF of genes with downregulated expression were “Heme binding”, “Iron ion binding”, “Zinc ion binding”, “Oxidoreductase activity”, and “Monooxygenase activity” (Fig. 2C). The top-five CC of genes with downregulated expression were “Extracellular exosome”, “Extracellular region”, “Extracellular space”, “Endoplasmic reticulum membrane”, and “Organelle membrane” (Fig. 2C). The top-five BP of genes with downregulated expression were “Oxidation–reduction process”, “Cellular response to cadmium ion”, “Negative regulation of growth”, “Cellular response to zinc ion”, and “Xenobiotic metabolic process” (Fig. 2C). The genes with downregulated expression were shown in Fig. 2D.

Pathway analyses using the KEGG database were undertaken using DAVID and visualized with ‘ggplot2’ and Cytoscape 3.7.1. In these DEGs, the pathways that were augmented were “Caffeine metabolism”, “Cell cycle”, “p53 signaling pathway”, “Oocyte meiosis”, “Mineral absorption”, “Linoleic acid metabolism”, “Progesterone-mediated oocyte maturation”, “Retinol metabolism”, “Prion diseases”, “Chemical carcinogenesis”, “Tryptophan metabolism”, “Drug metabolism-cytochrome P450”, “Bile secretion”, “Metabolism of xenobiotics by cytochrome P450”, “Steroid hormone biosynthesis”, ‘Arachidonic acid metabolism”, “Glycolysis/gluconeogenesis”, “Complement and coagulation cascades”, “HTLV-I infection”, “The 5' adenosine monophosphate-activated protein kinase (AMPK) signaling pathway”, “Insulin signaling pathway”, “Proteoglycans in cancer”, and “Metabolic pathways” (Fig. 2E). The genes involved in these pathways were visualized in Fig. 2F.

Investigation of DEGs using PPI networks.

We identified 135 DEGs in the GSE55092, GSE62232, and GSE121248 datasets, and constructed PPI networks through the STRING database. We screened seven functional modules from the PPI network through MCODE within Cytoscape. The main function modules are shown in Supplementary Fig. S2 A. The genes in this module that were enriched according to the KEGG database were “Cell cycle”, “Oocyte meiosis”, “p53 signaling pathway”, “Progesterone-mediated oocyte maturation”, “HTLV-I infection”, and “Viral carcinogenesis” (see Supplementary Fig. S2 B online). TOP2A, DLGAP5, RAD51AP1, ZWINT, BUB1B, CCNB1, FOXM1, CCNB2, AURKA and CDK1 were selected as top-ten hub genes through Cytoscape according to the value of degree. The expression heatmap of hub genes in these datasets was shown as Fig. 3.

Verification of hub genes.

We verified these ten hub genes in TCGA HCC cohort, ICGC-LIRI-JP dataset, and Human Protein Profiles. We identified DEGs through 'edgeR' in the ICGC-LIRI-JP and TCGA cohort. Expression of these ten hub genes was increased dramatically in ICGC-LIRI-JP (Fig. 4A) and TCGA (Fig. 4B) project. We also confirmed these ten hub genes protein expression patterns in the Human Protein Profiles, FOXM1 and AURKA were robustly positive in HCC samples compared with normal tissues, TOP2A, DLGA5, CCNB1, CCNB2, and CDK1 were moderately increased in HCC cohort in contrast to normal patients (Fig. 4C). However, RAD51AP1, ZWINT and BUB1B were not found in the website.

Construction and verification of a two-gene prognostic model.

Survival of hub genes was assessed using the ICGC-LIRI-JP cohort. All of these ten hub genes were correlated significantly with patient survival (see Supplementary Fig. S3 online). Expression of these ten hub genes in these 232 HCC patients was documented for multivariate cox regression analyses.

These patients were separated into low or high-risk groups according to the median risk score as the cut-off value.
FOXM1 and CDK1 were identified as independent prognostic factors related to patient survival with $P < 0.05$; their coefficients were 0.108 and 0.129, and hazard ratios were 1.114 and 1.138 respectively. The Kaplan–Meier curve for Overall survival (OS) in the high- and low-risk groups based on the genes of FOXM1 and CDK1 was dramatically different ($P = 6.878 \times 10^{-7}$) (Fig. 6A). The prognostic capacity based on the genes of FOXM1 and CDK1 was analyzed by the area Under the curve (AUC) of ROC (Fig. 5A). The AUC for survival at 1, 2 and 3 years was 0.769, 0.786 and 0.815, respectively. A higher AUC indicates a better forecasting model than a lower AUC, so our two-gene model showed high sensitivity and specificity to assess the prognosis.

Furthermore, we verified the two-gene prognostic model by TCGA HCC cohort, The Kaplan–Meier curve for OS was significant different ($P = 2.718 \times 10^{-3}$) between the high- and low-risk groups (Fig. 5B). The AUC for survival at 1, 2 and 3 years was 0.639, 0.636 and 0.644, respectively (Fig. 5B). Therefore, we can conclude that FOXM1 and CDK1 are prognosis molecules for ICGC-LIRI-JP and TCGA HCC project.

We established a liver PTEN knockout (KO) mouse, extrons 4 and 5 of PTEN were deleted (Fig. 6A), and these mice developed liver tumors at 12 months (Fig. 6B). mRNA (Foxm1 and Cdk1) expression increased significantly in the liver tumor tissue of KO mice in contrast to the normal liver tissue of CTL mice (Fig. 6C). These results indicated that the liver tumor expression patterns of FOXM1 and CDK1 were consistently in human and mouse.

Establishment of a prognostic model using a nomogram.

We collected the complete clinical data of 213 patients in the ICGC-LIRI-JP dataset. Multivariate cox regression was undertaken to evaluate the two-gene prognostic model and clinical factors (histology grade, pathology stage, age, cancer history, malignancy and sex) for the prognosis. The histology grade, pathology stage, sex and two-gene prognostic model were independent prognostic factors with $P < 0.05$ (Fig. 7).

We constructed a nomogram with the four independent prognostic factors selected from multivariate cox regression analyses to predict OS at 1, 2 and 3 years in the ICGC-LIRI-JP cohort (Fig. 8A). The gray line in the calculation curve indicated the best prediction, and the nomogram model (described with a red line) matched well with the gray line (Fig. 8B). We also analyzed the AUC of the nomogram model, and the AUC for OS at 1, 2 and 3 years was 0.837, 0.804 and 0.829, respectively (Fig. 8C). Therefore, we concluded that the nomogram model performed well.

We evaluated the nomogram model through DCA. The nomogram model based on a combination of histology grade, pathology stage, sex and two-gene prognostic model performed better to predict OS than that using a single factor (Fig. 8D). Hence, the nomogram model could facilitate preoperative consultation of the patient, decision-making by the clinician and the follow-up schedule.

Discussion

We integrated DEGs obtained from HBV-HCC-related GEO datasets through the RRA method, conducted functional analyses using the GO database, pathway analyses using the KEGG database, and selected hub genes with these DEGs. CDK1, FOXM1, TOP2A, DLGAP5, RAD51AP1, ZWINT, BUB1B, CCNB1, CCNB2 and AURKA were identified as the top ten hub genes. FOXM1 and CDK1 were selected as prognosis-related molecules for HBV–HCC, and were verified using the ICGC-LIRI-JP, TCGA HCC cohort, Human protein profiles dataset, and PTEN knockout mouse liver tissues. In addition, a nomogram model was built to predict the prognosis for HBV–HCC patients.

Integrated DEGs were clustered according to pathway analyses, and alterations were detected in “glycolysis”, “AMPK signaling pathway”, “p53 signaling pathway”, “cell cycle”, and “bile secretion”. Greater consumption of glucose in tumors than that in normal tissues was discovered first by the German physiologist Otto Warburg [13]. The universal metabolic feature of cancer cells is preferential conversion of glucose to lactate instead of entry into the tricarboxylic-acid cycle [14]. This action provides carbon sources for rapid growth of tumor cells and an acidic micro-environment that aids the invasion and metastasis of cancer cells [15]. The p53, AMPK and mechanistic target of rapamycin (mTOR) signaling pathways work collectively to modulate aerobic glycolysis in tumors. AMPK is a Ser/Thr kinase and a heterotrimer complex containing catalytic α as well as regulatory β and γ subunits [16]. AMPK is activated by liver kinase B1 (LKB1) through phosphorylation of Thr172, located in the catalytic α subunit [16]. AMPK is switched on if there is a deficiency of energy and nutrients, attenuates aerobic glycolysis, and inhibits HCC [17,18]. AMPK and mTOR signaling pathways antagonistically sense cellular energy and nutrients and regulate cell growth. mTOR (another Ser/Thr kinase) is switched on if there is abundant energy and

nutrients, and promotes tumor growth [19]. P53 is one of the most frequent mutation sites in HCC [1]. P53, as a tumor suppressor, regulates cell proliferation or activates senescence and apoptosis [20]. P53 null mice develop leukemias and sarcomas [21]. P53 contributes to the cancer hallmark of evading growth suppressors [20]. P53 deregulates aerobic glycolysis by targeting the key enzymes of glucose metabolism, including glucose transporter 3, hexokinase 2, phosphoglycerate mutase, and pyruvate kinase isoform M2 [22]. LKB1/AMPK increases P53 expression and blocks the mTOR signaling pathway by stimulating cell-cycle arrest and apoptosis [22]. Alterations of the synthesis, transport, and metabolism of bile acids are universal in liver inflammation and HCC [23]. Secretion of bile acids from the liver into the bile duct is mediated by bile salt export pump (BSEP), expression of which is regulated by the farnesoid X receptor (FXR) [24]. Deficiency of BSEP or the FXR induces tumor genesis in humans and mice [25,26]. We isolated hub genes through analyses of PPI networks, and CDK1, FOXM1, TOP2A, DLGAP5, RAD51AP1, ZWINT, BUB1B, CCNB1, CCNB2 and AURKA were pinpointed as the top ten hub genes. CDK1, TOP2A, BUB1B and CCNB1 are involved in the cell cycle, which is usually aberrant in tumor cells. Some of the hub genes involved in HCC identified by other scholars are identical to those identified in the present study [27,28], but a nomogram model was not constructed in those studies.

We identified CDK1 and FOXM1 as pivotal for the prognosis of HBV–HCC patients. CDK1 belongs to Ser/Thr kinase family, and is indispensable for the centrosome cycle and onset of mitosis [29]]. Deletion of CDK1 causes lethality in mouse embryos in morula and blastocyst stages [30]. CDK1 activated by cyclin A and cyclin B drives the phosphorylation of thousands of proteins to mitosis [31]. Aberrant regulation of CDK1 expression triggers genomic instability and chromosomal instability, which are signature features of chromosomally unstable tumors [32]. CIP/KIP and INK4 family proteins are inhibitors of CDK1. The CIP/KIP protein family is composed of P21, P27 and P57, whereas the INK4 protein family is composed of INK4A, INK4B, INK4C and INK4D [32]. Mice with ablation of P21 or P27 develop tumors spontaneously [33,34], which can be attributed (at least in part) to CDK1 activation.

FOXM1 has multiple roles in the growth, angiogenesis, metabolism, migration, DNA-damage response, development and progression of tumor cells. FOXM1 expression is augmented in cancer of the liver, lung, breast, stomach, brain, colon, pancreas, prostate gland, and blood [35–37]. FOXM1 depletion in mice can lead to embryonic lethality, and FOXM1 is necessary for mitosis of hepatoblast-like precursor cells and liver regeneration [38,39]. In addition, FOXM1 provokes the activity of CDK1 and PLK1 to promote cells to enter mitosis [40,41].

Conclusion

In the present study, we identified FOXM1 and CDK1 as molecular biomarkers of HCC, and created a nomogram model as a predictive tool for the prognosis of HBV–HCC patients.

Materials And Methods

Data source.

GSE55092, GSE62232, and GSE121248 gene-expression profiles were downloaded in GEO datasets. We chose 120 samples (81 samples were non-tumor control samples, and 39 were HBV–HCC samples) from GSE55092, 26 samples (10 samples were non-tumor control samples, and 16 were HBV-HCC samples) from GSE62232, and 107 samples (37 samples were non-tumor control samples, and 70 were HBV-HCC samples) from GSE121248 in our work. The platform of GSE55092, GSE62232, and GSE121248 was the GPL570 [HG-U133_plus_2] affymetrix human genome U133 plus 2.0 array. ICGC-LIRI-JP contains 202 normal and 243 cancer patient samples, and all the patients were HBV- and HCV-infected. TCGA HCC project includes 371 HCC patient tissues and 50 adjacent non-tumorous liver tissues.

Screening of DEGs.

We analyzed the data using R 4.0.2. Normalization and DEGs analyses were done with 'limma R' or 'edgeR' within Bioconductor (www.bioconductor.org/). All gene-expression data were subjected to $\log_2$ transformation. DEGs were screened with corrected $P < 0.05$ and $|\log \text{fold change (FC)}| > 2$. RRA within R was used to integrate the DEGs in the microarray data.

Pathway analyses of DEGs using GO and KEGG databases.

Gene annotation (using the GO database) and pathway enrichment (using the KEGG database) of DEGs were analyzed with DAVID 6.8 and $P < 0.05$ was considered significant. Visualization was undertaken with 'ggplot2' within R and Cytoscape 3.7.1.

Construction of protein-protein interaction (PPI) networks and module analyses.

PPI networks were created using the STRING database. Cytoscape 3.7.1 was applied to screen for hub genes according to degrees. We analyzed the modules in PPI networks. The MCODE default parameters were 'Degree cutoff = 2', 'Node score cutoff = 0.2', 'K-core = 2', and 'Max. depth = 100' [42].

Prognostic analyses of hub genes.

Survival analyses of hub genes was done through R based on the ICGC-LIRI-JP dataset. Cox regression analyses were undertaken through genes selected from survival data with $P < 0.05$. Correlation of the prognosis with candidate genes was analyzed and visualized by 'survival', 'survminer' and 'ggplot2' packages within R. Time-dependent analyses of Receiver operating characteristic (ROC) curves by

survival were done using the ‘ROC’ package within R. The prognostic model was established based on the ICGC-LIRI-JP dataset.

Nomogram.

Multivariate Cox regression analyses was undertaken to select significant factors using a two-gene prognostic model and clinical characteristics for the prognosis. The result of multivariate cox regression was visualized with forest plots. The nomogram, calculation plot, ROC plot and Decision curve analysis (DCA) were calculated and visualized with ‘rms’, ‘foreign’, ‘survival’, ‘regplot’, ‘mstate’, ‘survivalROC’, ‘survcomp’, ‘Hmisc’, ‘grid’, ‘lattice’, ‘Formula’, ‘ggplot2’ and ‘rmda’ packages within R.

Animals.

Mice were maintained in cages in a room equipped with an air-filtering system, and they were kept on a 12-h light/dark cycle. The mice were fed with standard food and given with sterilized water. Alb-Cre mice [43] (Stock NO.003574; strain background C57BL/6J) were obtained from Jackson laboratory. $PTEN^{floxp/floxp}$ mice (strain background C57BL/6J) [44] were kindly provided by Dr Hongbing Zhang (Institute of basic medical sciences and school of basic medicine, Graduate school of peking union medical college, Chinese academy of medical sciences, China).

Quantitative real-time PCR.

RNA was extracted from liver tissues by a Total RNA kit (OMEGA, Norcross, GA, USA) and reverse transcripted with an oligo-dT primer (Toyobo, Osaka, Japan). qPCR was performed as previously described [19] using a Roche 480 instrument and SYBR green PCR master mix (Roche, Mannheim, Germany), Quantification was performed with the $^{\Delta\Delta}$CT method. The primers (Sangon Biotech, Beijing, China) were listed in Supplementary Table S1.

Statistical analysis.

Data are shown as the mean ± SEM. Statistical analyses were performed using Prism 8.0 software (Graph-pad software inc.) or R, and P values less than 0.05 were considered significant.

Consent for publication.

All the authors have consented for the publication.

Declarations

Consent for publication.

All the authors have consented for the publication.

Acknowledgements

Not applicable

Funding

Funding information is not applicable

Data availability

Publicly datasets were analyzed in the present study. The data can be found at https://dcc.icgc.org/; https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga and www.ncbi.nlm.nih.gov/geo/.

Authors' contributions

LW and AL designed the current study. LW, JL, ZX, JT, CL, YZ, SL, JH, CP, YY, and YZ acquired and analyzed the data. HB provided the material. LW wrote and revised the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

This study was carried out in strict accordance with the recommendations of ethical guidelines and regulations for the use of laboratory animals and cells issued by the Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China. All animal-related protocol was approved by the Committee on the Ethics of Animal Experiments of the Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences. The authors complied with the ARRIVE guidelines.

Patient consent for publication

Not applicable

Competing interests

The authors declare no competing interests.

References

1. Craig, A. J., von Felden, J., Garcia-Lezana, T., Sarcognato, S. & Villanueva, A. Tumour evolution in hepatocellular carcinoma. *Nat Rev Gastroenterol Hepatol.***17**, 139–152 https://doi.org/10.1038/s41575-019-0229-4 (2020).
2. Tan, A. T. & Schreiber, S. Adoptive T-cell therapy for HBV-associated HCC and HBV infection. *Antiviral Res.***176**, 104748 https://doi.org/10.1016/j.antiviral.2020.104748 (2020).
3. Xie, Y., Hepatitis, B. & Virus-Associated Hepatocellular Carcinoma. *Adv Exp Med Biol.***1018**, 11–21 https://doi.org/10.1007/978-981-10-5765-6_2 (2017).
4. Global Burden of Disease *et al.* Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol.***5**, 1749–1768 https://doi.org/10.1001/jamaoncol.2019.2996 (2019).
5. Marrero, J. A. *et al.* Alcohol, tobacco and obesity are synergistic risk factors for hepatocellular carcinoma. *J Hepatol.***42**, 218–224 https://doi.org/10.1016/j.jhep.2004.10.005 (2005).
6. Yan, H. *et al.* Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. *Elife.***3**, https://doi.org/10.7554/eLife.00049 (2012).
7. Chang, M. H. *et al.* Universal hepatitis B vaccination in Taiwan and the incidence of hepatocellular carcinoma in children. Taiwan Childhood Hepatoma Study Group. *N Engl J Med.***336**, 1855–1859 https://doi.org/10.1056/NEJM199706263362602 (1997).
8. Hsu, C. H., Shen, Y. C., Shao, Y. Y., Hsu, C. & Cheng, A. L. Sorafenib in advanced hepatocellular carcinoma: current status and future perspectives. *J Hepatocell Carcinoma.***1**, 85–99 https://doi.org/10.2147/JHC.S45040 (2014).
9. Kudo, M. *et al.* Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet.***391**, 1163–1173 https://doi.org/10.1016/S0140-6736(18)30207-1 (2018).
10. Bruix, J. *et al.* Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet.***389**, 56–66 https://doi.org/10.1016/S0140-6736(16)32453-9 (2017).
11. Gao, X. *et al.* Identification of key candidate genes and biological pathways in bladder cancer. *PeerJ.***6**, e6036 https://doi.org/10.7717/peerj.6036 (2018).
12. Kolde, R., Laur, S., Adler, P. & Vilo, J. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics.***28**, 573–580 https://doi.org/10.1093/bioinformatics/btr709 (2012).
13. Warburg, O., Wind, F. & Negelein, E. The Metabolism of Tumors in the Body. *J Gen Physiol.***8**, 519–530 https://doi.org/10.1085/jgp.8.6.519 (1927).

14. Pavlova, N. N. & Thompson, C. B. The Emerging Hallmarks of Cancer Metabolism. *Cell Metab.***23**, 27–47 https://doi.org/10.1016/j.cmet.2015.12.006 (2016).
15. Gatenby, R. A. & Gawlinski, E. T. The glycolytic phenotype in carcinogenesis and tumor invasion: insights through mathematical models. *Cancer Res.***63**, 3847–3854 (2003).
16. Gonzalez, A., Hall, M. N., Lin, S. C. & Hardie, D. G. AMPK and TOR: The Yin and Yang of Cellular Nutrient Sensing and Growth Control. *Cell Metab.***31**, 472–492 https://doi.org/10.1016/j.cmet.2020.01.015 (2020).
17. Feng, J. *et al.* Emerging roles and the regulation of aerobic glycolysis in hepatocellular carcinoma. *J Exp Clin Cancer Res.***39**, 126 https://doi.org/10.1186/s13046-020-01629-4 (2020).
18. Faubert, B. *et al.* AMPK is a negative regulator of the Warburg effect and suppresses tumor growth in vivo. *Cell Metab.***17**, 113–124 https://doi.org/10.1016/j.cmet.2012.12.001 (2013).
19. Ma, A. *et al.* Tsc1 deficiency-mediated mTOR hyperactivation in vascular endothelial cells causes angiogenesis defects and embryonic lethality. *Hum Mol Genet.***23**, 693–705 https://doi.org/10.1093/hmg/ddt456 (2014).
20. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell.***144**, 646–674 https://doi.org/10.1016/j.cell.2011.02.013 (2011).
21. Ghebranious, N. & Donehower, L. A. Mouse models in tumor suppression. *Oncogene.***17**, 3385–3400 https://doi.org/10.1038/sj.onc.1202573 (1998).
22. Gomes, A. S., Ramos, H., Soares, J. & Saraiva, L. p53 and glucose metabolism: an orchestra to be directed in cancer therapy. *Pharmacol Res.***131**, 75–86 https://doi.org/10.1016/j.phrs.2018.03.015 (2018).
23. Jia, W., Xie, G. & Jia, W. Bile acid-microbiota crosstalk in gastrointestinal inflammation and carcinogenesis. *Nat Rev Gastroenterol Hepatol.***15**, 111–128 https://doi.org/10.1038/nrgastro.2017.119 (2018).
24. Dawson, P. A., Lan, T. & Rao, A. Bile acid transporters. *J Lipid Res.***50**, 2340–2357 https://doi.org/10.1194/jlr.R900012-JLR200 (2009).
25. Iannelli, F. *et al.* Massive gene amplification drives paediatric hepatocellular carcinoma caused by bile salt export pump deficiency. *Nat Commun.***5**, 3850 https://doi.org/10.1038/ncomms4850 (2014).
26. Yang, F. *et al.* Spontaneous development of liver tumors in the absence of the bile acid receptor farnesoid X receptor. *Cancer Res.***67**, 863–867 https://doi.org/10.1158/0008-5472.CAN-06-1078 (2007).
27. Xie, W. *et al.* Nine hub genes related to the prognosis of HBV-positive hepatocellular carcinoma identified by protein interaction analysis. *Ann Transl Med.***8**, 478 https://doi.org/10.21037/atm.2020.03.94 (2020).
28. Chen, Z. *et al.* Identification of Potential Key Genes for Hepatitis B Virus-Associated Hepatocellular Carcinoma by Bioinformatics Analysis. *J Comput Biol.***26**, 485–494 https://doi.org/10.1089/cmb.2018.0244 (2019).

29. Xie, B., Wang, S., Jiang, N. & Li, J. J. Cyclin B1/CDK1-regulated mitochondrial bioenergetics in cell cycle progression and tumor resistance. *Cancer Lett.***443**, 56–66 https://doi.org/10.1016/j.canlet.2018.11.019 (2019).
30. Santamaria, D. *et al.* Cdk1 is sufficient to drive the mammalian cell cycle. *Nature.***448**, 811–815 https://doi.org/10.1038/nature06046 (2007).
31. Crncec, A. & Hochegger, H. Triggering mitosis. *FEBS Lett.***593**, 2868–2888 https://doi.org/10.1002/1873-3468.13635 (2019).
32. Malumbres, M. & Barbacid, M. Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer.***9**, 153–166 https://doi.org/10.1038/nrc2602 (2009).
33. Martin-Caballero, J., Flores, J. M., Garcia-Palencia, P. & Serrano, M. Tumor susceptibility of p21(Waf1/Cip1)-deficient mice. *Cancer Res.***61**, 6234–6238 (2001).
34. Fero, M. L. *et al.* A syndrome of multiorgan hyperplasia with features of gigantism, tumorigenesis, and female sterility in p27(Kip1)-deficient mice. *Cell.***85**, 733–744 https://doi.org/10.1016/s0092-8674(00)81239-8 (1996).
35. Kim, I. M. *et al.* The Forkhead Box m1 transcription factor stimulates the proliferation of tumor cells during development of lung cancer. *Cancer Res.***66**, 2153–2161 https://doi.org/10.1158/0008-5472.CAN-05-3003 (2006).
36. Kalinina, O. A. *et al.* Sustained hepatic expression of FoxM1B in transgenic mice has minimal effects on hepatocellular carcinoma development but increases cell proliferation rates in preneoplastic and early neoplastic lesions. *Oncogene.***22**, 6266–6276 https://doi.org/10.1038/sj.onc.1206640 (2003).
37. Song, X., Kenston, F., Zhao, S. S., Yang, J., Gu, Y. & D. & Roles of FoxM1 in cell regulation and breast cancer targeting therapy. *Med Oncol.***34**, 41 https://doi.org/10.1007/s12032-017-0888-3 (2017).
38. Krupczak-Hollis, K. *et al.* The mouse Forkhead Box m1 transcription factor is essential for hepatoblast mitosis and development of intrahepatic bile ducts and vessels during liver morphogenesis. *Dev Biol.***276**, 74–88 https://doi.org/10.1016/j.ydbio.2004.08.022 (2004).
39. Tang, S. Y., Jiao, Y. & Li, L. Q. [Significance of Forkhead Box m1b (Foxm1b) gene in cell proliferation and carcinogenesis]. *Ai Zheng.***27**, 894–896 (2008).
40. Fu, Z. *et al.* Plk1-dependent phosphorylation of FoxM1 regulates a transcriptional programme required for mitotic progression. *Nat Cell Biol.***10**, 1076–1082 https://doi.org/10.1038/ncb1767 (2008).
41. Zhang, N. & Pati, D. Separase Inhibitor Sepin-1 Inhibits Foxm1 Expression and Breast Cancer Cell Growth. *J Cancer Sci Ther.***10**, https://doi.org/10.4172/1948-5956.1000517 (2018).
42. Bader, G. D. & Hogue, C. W. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics.***4**, 2 https://doi.org/10.1186/1471-2105-4-2 (2003).
43. Postic, C. *et al.* Dual roles for glucokinase in glucose homeostasis as determined by liver and pancreatic beta cell-specific gene knock-outs using Cre recombinase. *J Biol Chem.***274**, 305–315 https://doi.org/10.1074/jbc.274.1.305 (1999).

44. Tong, Z. *et al.* Pancreas-specific Pten deficiency causes partial resistance to diabetes and elevated hepatic AKT signaling. *Cell Res.***19**, 710–719 https://doi.org/10.1038/cr.2009.42 (2009).

Figures

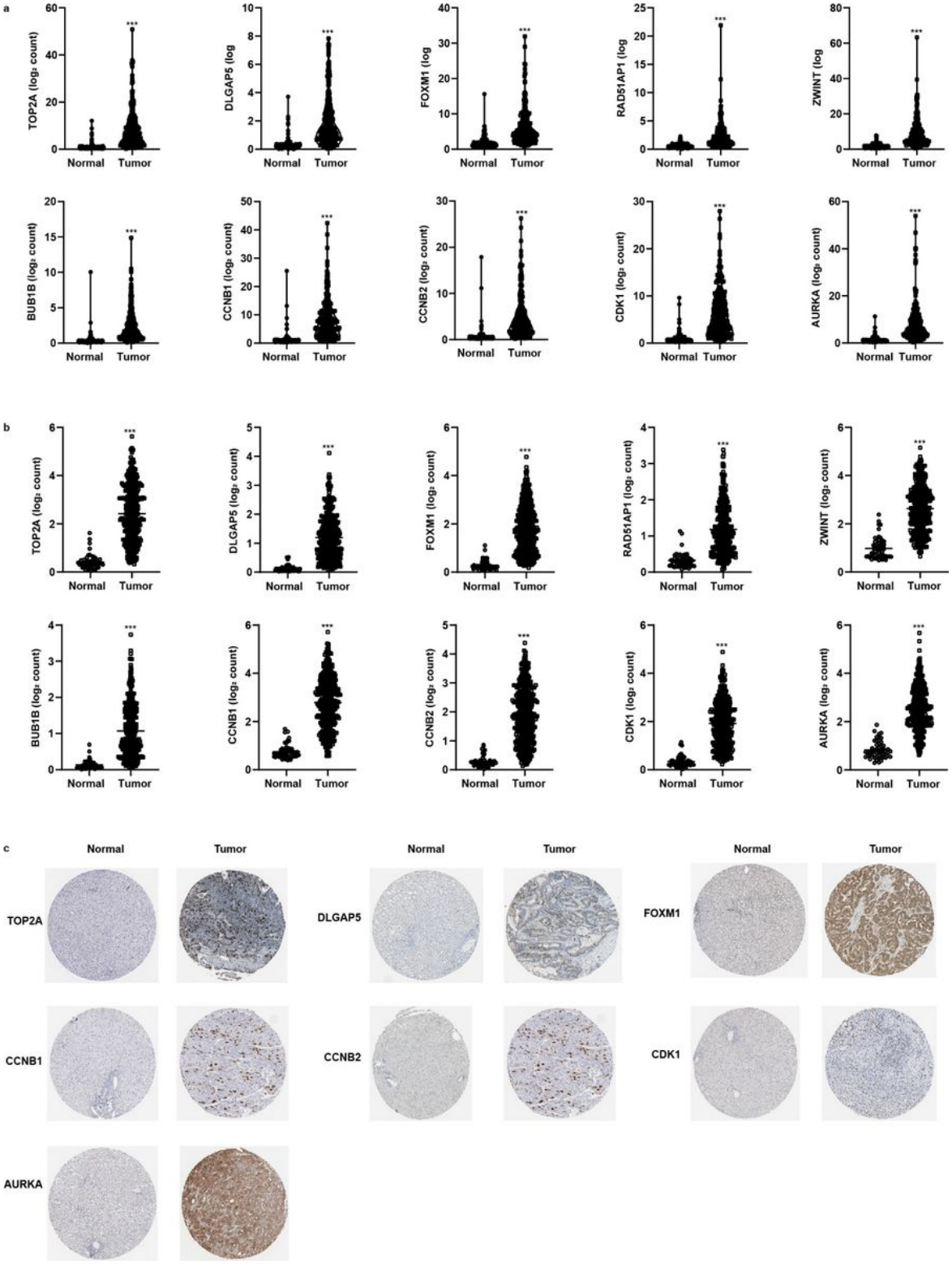

Figure 4

mRNA and protein expression patterns of the indicated genes. The mRNA levels in ICGC-LIRI-JP cohort (A), and TCGA HCC dataset (B), and protein levels in Human Protein Profiles dataset (C). n=202 in normal group, n=243 in tumor group in (A); n=50 in normal group, n=371 in tumor group in (B). ***P ≤ 0.001 vs. normal patients.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfigures.docx