

HMGCS2 in metabolic pathways was associated with overall survival in hepatocellular carcinoma: A LASSO-derived study

Science Progress

2021, Vol. 104(3) 1–21

© The Author(s) 2021

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/00368504211031749

journals.sagepub.com/home/sci

Rongrong Ding^{1*}, Tianyou Chen^{2*}, Yuan Zhang^{3*}, Xiaorong Chen³, Liping Zhuang⁴ and Zongguo Yang³ 

¹Department of Hepatobiliary Medicine, Shanghai Public Health Clinical Center, Fudan University, Shanghai, China

²Department of Interventional Medicine, Shanghai Public Health Clinical Center, Fudan University, Shanghai, China

³Department of Integrative Medicine, Shanghai Public Health Clinical Center, Fudan University, Shanghai, China

⁴Department of Integrative Oncology, Fudan University Shanghai Cancer Center, Shanghai, China

Abstract

This integrated bioinformatic study aimed to investigate potential prognostic candidates in hepatocellular carcinoma (HCC). In the GSE14520, GSE101685, and The Cancer Genome Atlas (TCGA) datasets, differentially expressed genes (DEGs) were identified and functional pathways of common DEGs were enriched. The least absolute shrinkage and selection operator (LASSO) model was used to screen the potential parameters associated with overall survival (OS) in HCC patients. Metabolic pathways were the most significantly enriched functional pathways of common DEGs in these three datasets. After LASSO model analysis, HMGCS2, UGP2, BCLC staging and TNM staging were screened as potential prognostic candidates for OS in HCC patients in

*These authors contributed equally to this work.

Corresponding authors:

Zongguo Yang, Department of Integrative Medicine, Shanghai Public Health Clinical Center, Fudan University, 2901 Caolang Road, Shanghai 201508, China.

Email: yangzongguo@shphc.org.cn

Liping Zhuang, Department of Integrative Oncology, Fudan University Shanghai Cancer Center, 270 Dong'an Road, Shanghai 200032, China.

Email: dr_zhuanglp@aliyun.com



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>)

which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

GSE14520. HMGCS2 in the metabolic pathway was significantly downregulated in tumor tissues and peripheral blood mononuclear cells in HCC patients (all $p < 0.05$). Cox regression model indicated that HMGCS2 might be associate with OS in HCC patients in GSE14520 and in the TCGA ($p = 0.029$ and $p = 0.05$, respectively). Kaplan–Meier analysis demonstrated that HMGCS2 downregulation in tumors contributed to an unfavorable OS in HCC patients, both in GSE14520 and in the TCGA ($p = 0.0001$ and $p = 0.0002$, respectively). Additionally, HMGCS2 was significantly downregulated in HCC patients with high alpha-fetoprotein (AFP), main tumor size >5 cm, multinodular, advanced tumor staging including BCLC, TNM and CLIP (all $p < 0.05$). HMGCS2 was involved in metabolic pathways, and downregulated HMGCS2 in tumors was associated with unfavorable OS in HCC patients.

Keywords

HMGCS2, hepatocellular carcinoma, overall survival, metabolic pathways, ketone body metabolism

Introduction

As one of the leading causes of cancer death, the incidence of hepatocellular carcinoma (HCC) has continued to increase globally in recent decades and will rise over the next 10–20 years.^{1–3} Mortality related to liver cancers including HCC has markedly increased by more than 2% annually since 2007,¹ and HCC is the seventh most frequently occurring cancer and the second most common cause of cancer death. Even though comprehensive treatment strategies have been developed, the prognosis of HCC is still not up to expectations.

Biomarkers in tumor tissues are a direct and cost-effective adjunct, especially in monitoring disease prognosis and treatment target selection in HCC patients.^{4,5} Identifying reliable prognostic biomarkers and therapeutic targets for HCC treatment is urgently required.^{6,7} Currently, only three biomarkers alpha-fetoprotein (AFP), core fucosylated AFP and des-gamma-carboxy prothrombin are widely used for HCC.^{6,8} However, high-throughput technologies and gene chips have become fast approaches for identifying differentially expressed genes (DEGs) and functional pathways and have led to a dramatic increase in the availability of molecular insights at multiple biological levels involved in the aggressiveness of HCC.^{9–11} Various bioinformatics databases in hepatology are currently available, and the accessibility of genomic sequencing data from liver tumors has provided us with invaluable resources.¹¹ Although many biomarkers identified through public databases have not been well validated and used widely in clinics,^{5,12,13} integrating data from different sources including the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) still helps us to facilitate the identification of promising biomarkers or therapeutic targets.^{11,14}

This study identified common DEGs in multiple datasets and investigated the most significant functional pathways. Clinical data were obtained from GSE14520 and TCGA liver hepatocellular carcinoma (LIHC). The least absolute shrinkage and selection operator (LASSO) model was used to identify potential candidates for predicting prognosis in HCC patients. We hope our findings will provide helpful insights into the pathogenesis and progression of HCC.

Table 1. Details of GEO and TCGA datasets included in this study.

| Datasets | Contributor(s) | Sample | HCC | Non-HCC | Platform |
|-----------|---------------------|--------|-----|---------|--|
| GSE14520 | Wang et al. (2009) | Tissue | 247 | 239 | Affymetrix Human Genome U133A 2.0 Array/Affymetrix HT Human Genome U133A Array |
| GSE33006 | Huang et al. (2011) | Tissue | 3 | 3 | Affymetrix Human Genome U133 Plus 2.0 Array |
| GSE45436 | Hsieh (2013) | Tissue | 97 | 37 | Affymetrix Human Genome U133 Plus 2.0 Array |
| GSE55092 | Melis et al. (2014) | Tissue | 49 | 91 | Affymetrix Human Genome U133 Plus 2.0 Array |
| GSE60502 | Kao (2014) | Tissue | 18 | 18 | Affymetrix Human Genome U133A Array |
| GSE84402 | Qin et al. (2016) | Tissue | 14 | 14 | Affymetrix Human Genome U133 Plus 2.0 Array |
| GSE101685 | Sen-Yung (2017) | Tissue | 24 | 8 | Affymetrix Human Genome U133 Plus 2.0 Array |
| GSE58208 | Hui (2014) | PBMCs | 10 | 17 | Affymetrix Human Genome U133 Plus 2.0 Array |
| GSE49515 | Hui (2013) | PBMCs | 10 | 10 | Affymetrix Human Genome U133 Plus 2.0 Array |
| TCGA-LIHC | TCGA | Tissue | 374 | 50 | RNAseq Htseq |

PBMCs: peripheral blood mononuclear cells.

Materials and methods

Ethics statement

This study is a secondary analysis based on public datasets. As declared in the original datasets, every subject provided verified informed consent. The informed consents and study protocol were reviewed and approved by the Institutional Review Board of the participating institutes.^{15,16}

Data source

Microarray series of GSE14520¹⁵ and GSE101685 were downloaded from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>, Table 1). In GSE14520, 220 hepatitis B virus (HBV) related HCC patients were included in this analysis after

excluding 22 cases without gene expression and 5 cases without clinical data. All liver tissue was obtained with informed consent from patients who underwent radical resection between 2002 and 2003 at the Liver Cancer Institute and Zhongshan Hospital, Fudan University. In GSE101685, 8 normal tissues and 24 HCC patients were selected for RNA extraction and showed the gene expression profile. Detailed information on these GEO datasets is summarized in Table 1. In addition, mRNA normalized counts data of LIHC in the TCGA database derived from the RNAseq Htseq platform were downloaded from the Genomic Data Commons Data Portal (<https://cancergenome.nih.gov/>). This dataset contains 424 samples with 374 liver cancer samples and 50 normal liver samples.

Clinical and gene expression data in tumor samples with Z scores normalized of TCGA-LIHC dataset were obtained from cBioPortal for cancer genomics (<http://www.cbioportal.org/>).^{17,18} After excluding non-HCC subjects and those without gene expression data, 361 cases were finally enrolled in the study. Clinicopathological features including age, gender, American Joint Committee on Cancer (AJCC) staging, body mass index, family history of cancer, risk factors, pathological grade, vascular invasion, child-pugh stage, AFP, hepatic inflammation, and Ishak scores were considered.

Identification of DEGs

Raw.CEL files of the microarray GSE14520 and GSE101685 datasets were normalized by the quantile method of robust multichip analysis (RMA) from the *R* affy package.¹⁹ Relative log expression (RLE) and normalized unscaled standard error (NUSE) tests were used for quality assessment of each sample in the GEO series. Mean gene expression data were calculated when multiple probes mapped onto the same gene name. Missing gene expression data were imputed with the *k*-nearest neighbor (KNN) method by the impute index in the *R* program.²⁰ DEGs between tumor and nontumor tissues were identified by the Limma package in the *R* program.²¹ In the TCGA dataset, the edgeR^{22,23} package in the *R* program was used to identify DEGs between tumor and nontumor samples, with the criterion of $|\log_2FC| > 1.0$ and adjusted p -value < 0.05 .

Common DEGs in GSE14520 and GSE101685 and the LIHC profile in the TCGA database were calculated in a Venn diagram from Bioinformatics & Evolutionary Genomics (<http://bioinformatics.psb.ugent.be/webtools/Venn/>). As shown in Table 1, GEO series with raw .CEL file format, including GSE33006, GSE45436, GSE55092, GSE60502, and GSE84402, were used to reevaluate the expression levels between tumor and nontumor tissues of screened candidates. GSE58208 and GSE49515 were used to calculate candidate gene expression in peripheral blood mononuclear cells between HCC and non-HCC patients (Table 1). Protein levels of screened candidates between HCC and normal samples detected by immunohistochemistry (IHC) in the Human Protein Atlas (<https://www.proteinatlas.org/>, HPA) dataset were also presented.

Enrichment analysis

As we described above, common DEGs in GSE14520 and GSE101685, and LIHC profile in the TCGA database were identified using a Venn diagram. Then, the R package clusterProfiler²⁴ was used to address the top 20 functional enrichment pathways of these common DEGs.²⁵ The pathways with the most significantly enriched genes were selected for LASSO model establishment.

The LASSO model establishment

A LASSO regression model was used to determine the most powerful prognostic markers for overall survival in HCC patients.²⁶ “glmnet” and “survival” packages were used for LASSO model establishment with family equals to “cox” and alpha equals to 1.²⁷ The model was validated with five-fold cross-validation. Both “lambda.1se” and “lambda.min” were used to assess the coefficient of parameters.²⁷

Nomogram model establishment

Parameters significantly associated with OS in HCC patients in the multivariate Cox model were included in the risk prediction model by nomogram with the “rms” package in the R program. Based on the Cox proportional hazards model, the “survival” package in the R program was used to calculate the cumulative risk of death. To calculate the concordance index and its 95% confidence intervals (CIs), the “survcomp” package was used. The bootstrap method was used for repeated sampling for internal verification of the model. A calibration plot was presented to evaluate the performance of the nomogram, which was also established in the “rms” package in the R program.

Statistical analysis

Descriptive statistics for the gene expression levels were presented as mean \pm standard deviation (SD). Differences in gene expression levels between the individual groups were analyzed using Student's *t* test for normally distributed continuous data and the Mann–Whitney test for non-normally distributed continuous data. When the number of groups is greater than or equals to three, one-way ANOVA was also considered for comparing differences. Parameters screened significance for OS in HCC patients by the LASSO model in GSE14520 were included in univariate and multivariate Cox regression. Associations between candidate genes and OS in HCC patients were also validated in TCGA-LIHC dataset by Cox model. The results were reported as hazard ratios (HR) with 95% CIs. The Kaplan–Meier method was used to compare OS between different groups, and the log-rank test was used to estimate the significances in survival between groups. Stata software version 16.0 (Stata Corp LLC, Texas, USA) was used for other statistics. $P < 0.05$ (two-tailed) was considered significant for all tests.

Results

Common genes identification and enrichment

DEGs of GSE14520 and GSE101685 and LIHC in TCGA profiles were identified, and common genes were calculated by Venn diagram. Two tumor samples (GSM2712021 and GSM2712027) in GSE101685 were deleted from the DEG analysis process based on the quality assessment of the RLE and NUSE methods (Supplemental Figure S1). Finally, 653 common DEGs were included (Figure 1(a)). Using the clusterProfiler package in the R program, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of these 653 common genes were investigated in GSE14520. The top 20 functional pathways of these common genes enriched were presented in Figure 1(b), and the significant genes enriched in the top 20 functional pathways were summarized in Table 2. As described in Figure 1, the most significant genes ($n = 65$) were enriched in metabolic pathways ($p = 0.001$ and adjusted $p = 0.048$, Figure 1(c)).

Identification of potential candidates for OS in HCC

In GSE14520, parameters including age, gender, alanine aminotransferase (ALT), cirrhosis, main tumor size, multinodular status, TNM staging, Barcelona Clinic Liver Cancer (BCLC) staging, Cancer of the Liver Italian Program (CLIP) staging, AFP and 65 significant genes enriched in metabolic pathways were included in the LASSO model (Figure 2(a)). To assess the coefficient of these parameters, BCLC staging and TNM staging were recruited when lambda.1se was used. After five-fold cross-validation, variables including HMGCS2 (3-hydroxy-3-methylglutaryl-coenzyme A synthase 2), UGP2 (UDP-glucose pyrophosphorylase), BCLC staging and TNM staging were screened to be linked to OS in HCC patients when λ took the minimum value (Figure 2(a)). The regression coefficient plot of factors by LASSO is shown in Figure 2(b).

As shown in Table 3, HMGCS2 (median cut-off), UGP2 (median cut-off), BCLC staging and TNM staging were all potential parameters associated with OS in HCC patients (all $p < 0.05$, Table 3). When these variables were included in the multivariate Cox model, high HMGCS2 levels in tumor tissues might contribute to favorable OS in HCC patients after adjusting for BCLC staging and TNM staging (HR = 0.57, 95% CI = 0.34–0.94, $p = 0.029$, Table 3). In TCGA-LIHC dataset, univariate Cox regression indicated that HMGCS2, UGP2, AJCC staging, and vascular invasion might be factors correlated with OS in HCC patients (all $p < 0.05$, Table 3). After adjusting for AJCC staging and vascular invasion conditions, HMGCS2 showed potential protective impact on OS in HCC patients in multivariate Cox model (HR = 0.63, 95% CI = 0.39–1.0, $p = 0.05$, Table 3). Unfortunately, UGP2 showed no significance for OS in HCC patients in the multivariate model, both in GSE14520 and TCGA-LIHC datasets ($p = 0.844$ and $p = 0.279$, respectively, Table 3).

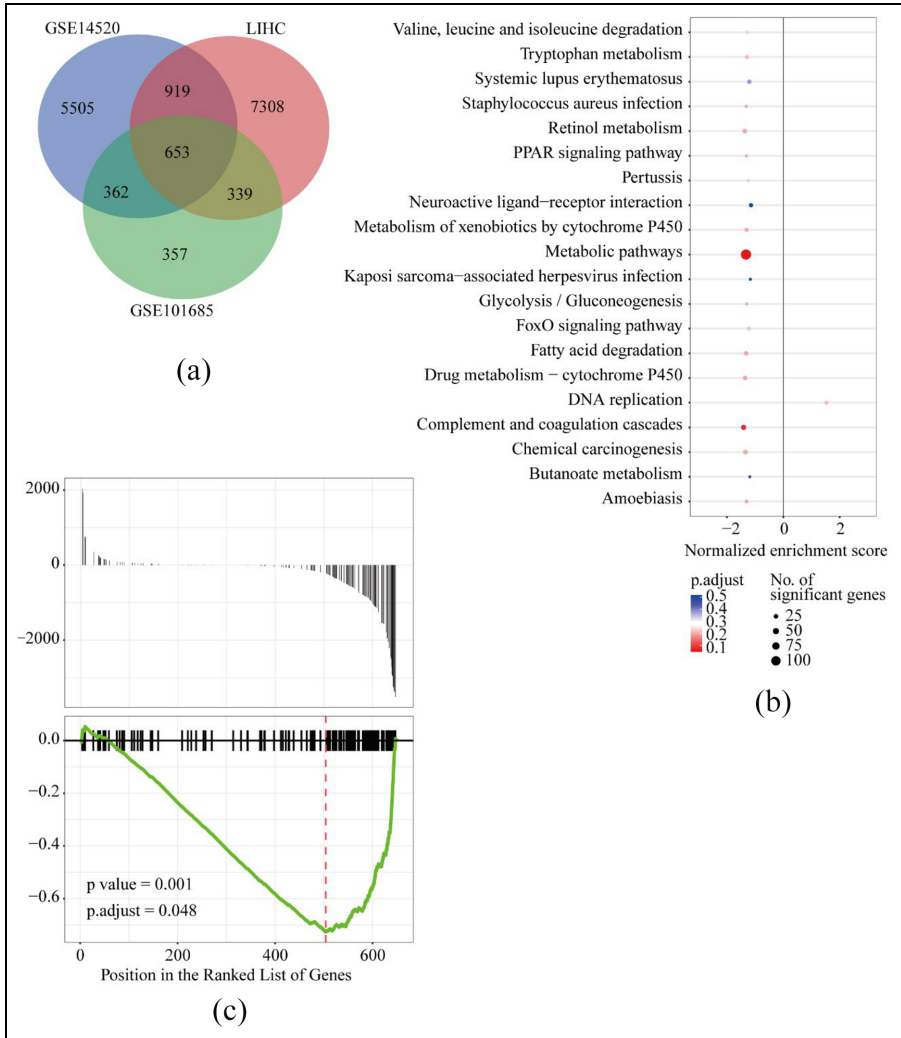


Figure 1. In GSE14520, LIHC and GSE101685, 653 differentially expressed genes (DEGs) were screened using Venn diagram (a), top 20 KEGG functional pathways were enriched by clusterProfiler package in R program (b), and the significances of metabolic pathways with most significant genes enriched (c).

Associations between *HMGCS2* and OS in HCC patients

HMGCS2 mRNA was significantly decreased in tumor tissues compared to nontumor tissues in multiple GEO series (all $p < 0.05$, Figure 3(a)). Moreover, *HMGCS2* mRNA was also downregulated in PBMCs in HCC patients compared to that in non-HCC patients (chronic hepatitis B patients and healthy individuals) in GSE58208 ($p < 0.0001$, Figure 3(a)), while no difference of *HMGCS2* from

Table 2. Significant genes enriched in the top 20 functional pathways.

| ID | Description | p value | Adj. p value | Enriched genes |
|----------|---|---------|--------------|---|
| hsa00280 | Valine, leucine and isoleucine degradation | 0.078 | 0.235 | ABAT; EHHADH; ACAT1; ACAA1; AOX1; ECHS1; ALDH6A1; HMGCS2; ALDH2 |
| hsa00380 | Tryptophan metabolism | 0.065 | 0.213 | ACAT1; CYP1A2; AOX1; ECHS1; CAT; ALDH8A1; TDO2; ALDH2 |
| hsa05322 | Systemic lupus erythematosus | 0.157 | 0.377 | C8A; C8B; C9; CIR; C1S; C6 |
| hsa05150 | Staphylococcus aureus infection | 0.045 | 0.196 | CFI; CIR; PLG; C1S |
| hsa00830 | Retinol metabolism | 0.019 | 0.196 | CYP4A11; RDH16; AOX1; CYP2A6; CYP3A4; ADH1C; ADH1A; CYP2C9; CYP2C8; ADH1B |
| hsa03320 | PPAR signaling pathway | 0.047 | 0.196 | ACAA1; CYP4A11; PKC2; ACSL1; HMGCS2; PKC1 |
| hsa05133 | Pertussis | 0.124 | 0.33 | CD14; SERPING1; CIR; C1S |
| hsa04080 | Neuroactive ligand-receptor interaction | 0.269 | 0.517 | AGTRI; VIPR1; GHR; PLG |
| hsa00980 | Metabolism of xenobiotics | 0.05 | 0.196 | CYP2A6; CYP3A4; ADH1C; ADH1A; CYP2C9; ADH1B; CYP2E1 |
| hsa01100 | Metabolism of xenobiotics by cytochrome P450 | 0.001 | 0.048 | AGL; CA2; KMO; CYP2C18; GSTZ1; CYP2B6; PEMT; GLUD1; GNE; GBA3; ACSM5; CYP2J2; CYP4F2; EPHX2; ACADSB; GLYAT; GCHI; GLS2; NAT2; BDH2; RGN; ABAT; GNMT; ADH6; GYS2; UGP2; FTCD; EHHADH; HAO1; ALAS1; HAO2; ACAT1; DPYS; CYP1A2; ACAA1; CYP4A11; PKC2; GRHPR; RDH16; AOX1; CYP2A6; ECHS1; CAT; ALDH6A1; ACSL1; FBP1; ALDH8A1; HMGCS2; MAT1A; AKR1D1; NNMT; BHMT; TAT; CYP3A4; TDO2; ALDH2; ADH1C; ADH1A; CYP2C9; ALDOB; ASS1; CYP2C8; PKC1; ADH1B; CYP2E1 |
| hsa05167 | Kaposi sarcoma-associated herpesvirus infection | 0.254 | 0.516 | GABARAPL1; RCANI; FOS; CXCL2 |
| hsa00010 | Glycolysis/gluconeogenesis | 0.053 | 0.196 | FBP1; ALDH2; ADH1C; ADH1A; ALDOB; PKC1; ADH1B |

(continued)

Table 2. (Continued)

| ID | Description | p value | Adj. p value | Enriched genes |
|----------|-------------------------------------|---------|--------------|---|
| hsa04068 | FoxO signaling pathway | 0.117 | 0.33 | IGF1; PCK2; CAT; PCK1 |
| hsa00071 | Fatty acid degradation | 0.047 | 0.196 | ADH6; EHHADH; ACAT1; ACAA1; CYP4A11; ECHS1; ACSL1; ALDH2; ADH1C; ADH1A; ADH1B |
| hsa00982 | Drug metabolism—cytochrome P450 | 0.025 | 0.196 | FMO3; CYP1A2; AOX1; CYP2A6; CYP3A4; ADH1C; ADH1A; CYP2C9; CYP2C8; ADH1B; CYP2E1 |
| hsa03030 | DNA replication | 0.067 | 0.213 | PCNA; MCM6; REC4; MCM3; MCM2; MCM7; RNASEH2A; MCM5; PRIM1; POLE2 |
| hsa04610 | Complement and coagulation cascades | 0.003 | 0.079 | C7; KLKB1; CFI; CPB2; C8A; C8B; C9; SERPING1; C1R; PLG; C1S; C6; FGA; F9 |
| hsa05204 | Chemical carcinogenesis | 0.031 | 0.196 | ADH6; CYP1A2; CYP2A6; CYP3A4; ADH1C; ADH1A; CYP2C9; CYP2C8; ADH1B; CYP2E1 |
| hsa00650 | Butanoate metabolism | 0.212 | 0.463 | BDH2; ABAT; EHHADH; ACAT1; ECHS1; HMGCS2 |
| hsa05146 | Amoebiasis | 0.051 | 0.196 | C8A; C8B; C9; CXCL2 |

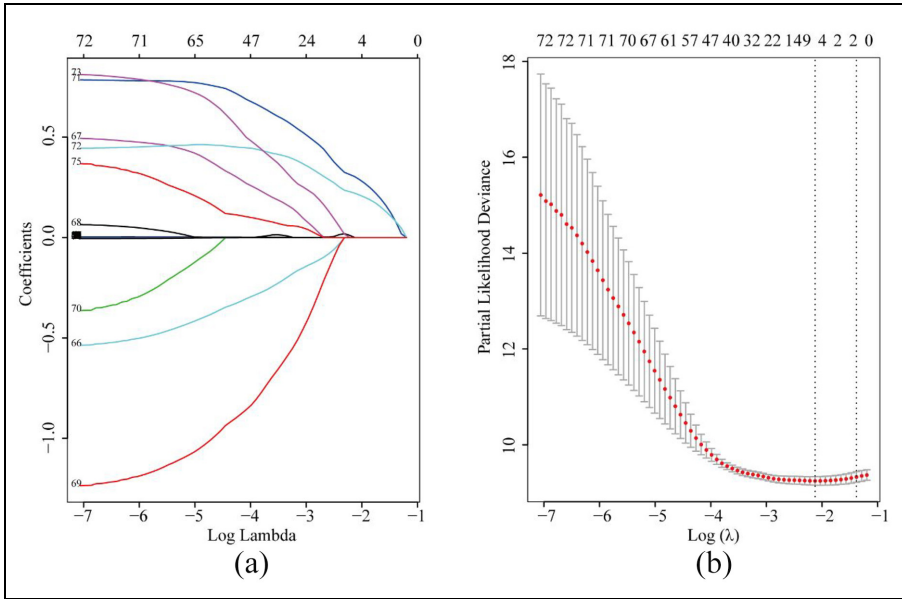


Figure 2. Parameters including significant genes in metabolic pathways and clinical data selection through LASSO regression (a) and elucidation of LASSO coefficient profiles for selected factors (b).

PBMCs between healthy individuals and HCC patients were observed in GSE49515 ($p > 0.05$, Figure 3(a)). In the HPA database, the antibody staining of HMGCS2 protein was high in 6/9 healthy individuals and 3/14 HCC patients (chi-square $p = 0.03$, Figure 3(b)).

In the GSE14520 dataset, HMGCS2 mRNA was significantly downregulated in tumor tissues compared to that in nontumor tissues ($p < 0.0001$, Figure 4(a)). Kaplan–Meier analysis revealed that HCC patients with low expression of HMGCS2 mRNA in tumors had unfavorable OS compared with those with high HMGCS2 ($p = 0.0001$, Figure 4(b)). In addition, we validate these findings in TCGA dataset. As shown in Figure 4(c), HMGCS2 mRNA expression was significantly lower expressed in HCC tumors than that in nontumors in TCGA dataset ($p < 0.0001$, Figure 4(c)). Survival analysis with the log-rank method also indicated that HCC patients with low HMGCS2 expression in tumors had significantly worse OS than those with high HMGCS2 expression ($p = 0.0002$, Figure 4(d)).

A nomogram model was established according to independent parameters including HMGCS2, BCLC staging and TNM staging (Figure 5(a)). According to the upper scale of each independent risk factor (-0.4 to 0.7), the corresponding score of this risk factor could be determined. The total score was obtained by adding the scores of each factor. Projecting downward from the total score, the corresponding mortality risk prediction probability value could be obtained. For example, in a patient with TNM stage III, BCLC stage 0–A and low HMGCS2,

Table 3. Cox regression models for identification of parameters associated with OS in HCC patients in GSE14520 and in the TCGA.

| Variables [#] | Univariate | | Multivariate | |
|-------------------------|-------------------|---------|------------------|---------|
| | HR (95% CI) | p value | HR (95% CI) | p value |
| GSE14520 dataset | | | | |
| HMGCS2, high versus low | 0.43 (0.27–0.667) | <0.001 | 0.57 (0.34–0.94) | 0.029 |
| UGP2, high versus low | 0.57 (0.37–0.89) | 0.013 | 0.95 (0.58–1.56) | 0.844 |
| BCLC staging | | | | |
| 0–A | Reference | 1.0 | Reference | 1.0 |
| B | 2.40 (1.28–4.53) | 0.007 | 1.31 (0.61–2.80) | 0.493 |
| C | 4.60 (2.72–7.76) | <0.001 | 2.51 (1.12–5.62) | 0.025 |
| TNM staging | | | | |
| I | Reference | 1.0 | Reference | 1.0 |
| II | 2.08 (1.21–3.59) | 0.008 | 1.87 (1.07–3.27) | 0.027 |
| III | 4.93 (2.82–8.60) | <0.001 | 2.37 (1.04–5.41) | 0.041 |
| TCGA-LIHC dataset* | | | | |
| HMGCS2, high versus low | 0.49 (0.35–0.71) | <0.001 | 0.63 (0.39–1.0) | 0.05 |
| UGP2, high versus low | 0.62 (0.44–0.88) | 0.007 | 0.77 (0.49–1.23) | 0.279 |
| AJCC staging | | | | |
| I | Reference | 1.0 | Reference | 1.0 |
| II | 1.49 (0.91–2.44) | 0.111 | 1.27 (0.68–2.36) | 0.454 |
| III | 2.82 (1.85–4.28) | <0.001 | 2.12 (1.25–3.59) | 0.005 |
| IV | 2.64 (1.38–5.04) | 0.003 | 2.65 (1.25–5.62) | 0.011 |
| Vascular invasion | | | | |
| None | Reference | 1.0 | Reference | 1.0 |
| Microvascular | 1.25 (0.80–1.97) | 0.327 | 0.96 (0.57–1.60) | 0.873 |
| Macrovascular | 2.24 (1.06–4.72) | 0.034 | 1.45 (0.66–3.19) | 0.361 |

[#] Only variables significantly associated with OS in HCC patients screened by univariate Cox regression were presented and included in the multivariate model.

*Variables including HMGCS2, UGP2, age, gender, body mass index, family history of cancer, risk factors, pathological grade, AJCC stage, vascular invasion, child-pugh stage, AFP, hepatic inflammation, and Ishak scores were included in the univariate Cox regression.

his/her total score was 0.685, and his/her mortality risk at 1-year, 3-year, and 5-year was 0.187, 0.492, and 0.634, respectively (Figure 5(a)). The concordance index of this model was 0.70 ± 0.03 (95% CI = 0.64–0.76). Calibration curves of 1-year, 3-year, and 5-year for internal verification of this nomogram with bootstrapping are described in Figure 5(b) to (d), respectively. The calibration curve displayed a good consistency between actual OS and predicted OS in 1-year, 3-year, and 5-year (Figure 5(b) to (d)).

Associations between HMGCS2 and clinicopathological features in HCC patients

In GSE14520, HMGCS2 mRNA expression was compared by clinicopathological features. As shown in Figure 5, HMGCS2 mRNA was significantly downregulated

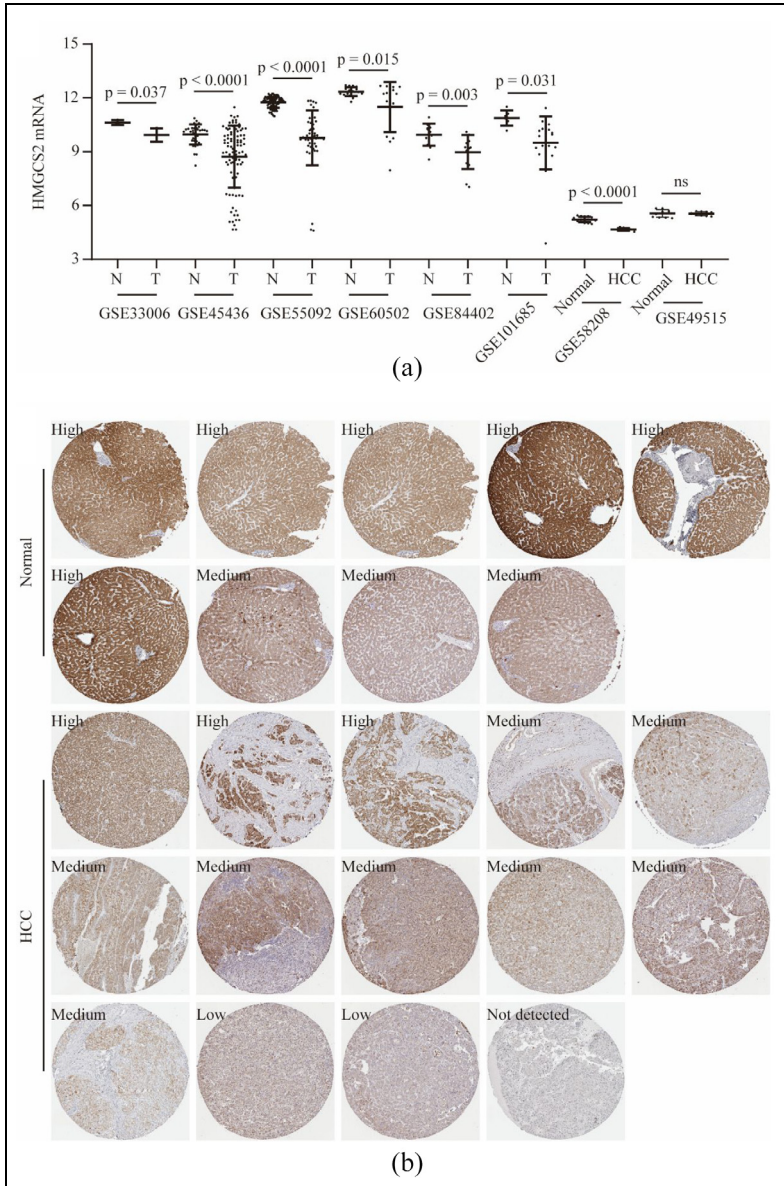


Figure 3. HMGCS2 mRNA was significantly lower in tumor tissues compared to that in nontumor tissues in HCC patients in GEO series, namely, GSE33006, GSE45436, GSE55092, GSE60502, GSE84402, and GSE101685 (all $p < 0.05$, a). In GSE58208, HMGCS2 mRNA was significantly downregulated in peripheral blood mononuclear cells (PBMCs) in HCC patients ($p < 0.0001$, a), while no significance was found in GSE49515 ($p > 0.05$, a). In the HPA database, the antibody staining of HMGCS2 protein was high in 6/9 normal livers and 3/14 HCC cases (chi-square $p = 0.03$, b).

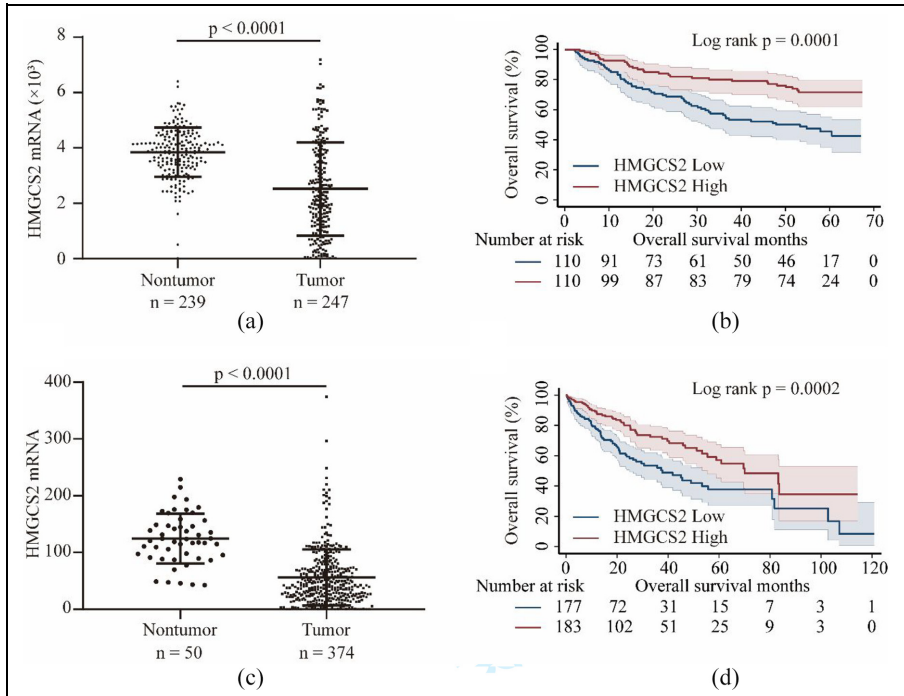


Figure 4. In GSE14520 dataset, HMGCS2 mRNA was significantly downregulated in tumor tissues than that in nontumor tissues (2521.42 ± 1683.83 vs 3846.15 ± 887.40 , $p < 0.0001$, a) and low HMGCS2 levels in tumors contributed to unfavorable overall survival (OS) in HCC patients (b). In TCGA-LICH profile, HMGCS2 mRNA was also significantly lower in HCC tumors than that in nontumor tissues (56.17 ± 49.39 vs 124.52 ± 44.06 , $p < 0.0001$, c) and was associated with worse OS (d).

in HCC patients with AFP >300 ng/ml, main tumor size >5 cm and/or multinodular ($p = 0.013$, $p = 0.01$, and $p = 0.002$, respectively, Figure 6(a) to (c)). Moreover, HMGCS2 mRNA was significantly lower in HCC patients with BCLC staging B and C than in those with BCLC staging 0–A ($p = 0.006$ and $p = 0.017$, respectively, Figure 6(d)). Similarly, HMGCS2 mRNA was significantly decreased in HCC patients with CLIP staging 1 and ≥ 2 compared to those with CLIP staging 0 ($p = 0.002$ and $p < 0.001$, respectively, Figure 6(e)). Additionally, HMGCS2 mRNA was significantly decreased according to the increase in TNM staging in HCC patients (all $p < 0.05$, Figure 6(f)).

Discussion

Metabolic reprogramming has been considered a hallmark in the cancer process.²⁸ Activation of oncogenes and loss of tumor suppressor genes enhanced metabolic

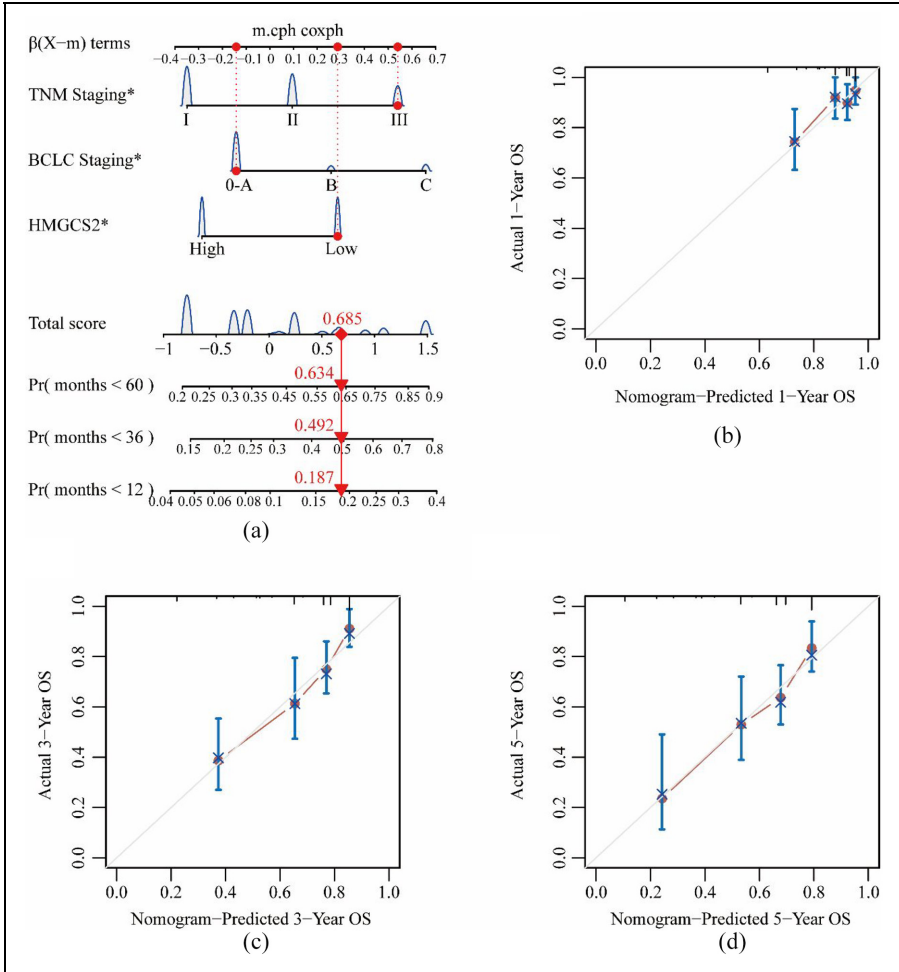


Figure 5. Screened by Cox regression model, TNM staging, BCLC staging, and HMGCS2 were included in competing risk model with nomogram for OS from HCC patients in GSE14520 (a), calibration curves for prediction 1-year (b), 3-year (c), and 5-year OS (d) in HCC patients in GSE14520.

reprogramming in human malignancies, leading to increased nutrient consumption, which could provide energy and promote biosynthetic pathways. Impaired metabolic pathways involved in cell proliferation and tumorigenesis progression,^{29–31} are consistently reprogrammed by mutations in MYC, TP53, Ras-related oncogenes, and the LKB1-AMP kinase (AMPK) and PI3 kinase (PI3K) signaling pathways.³² Our results based on three datasets also indicated that metabolic pathways were a vital mechanism in the development of HCC.

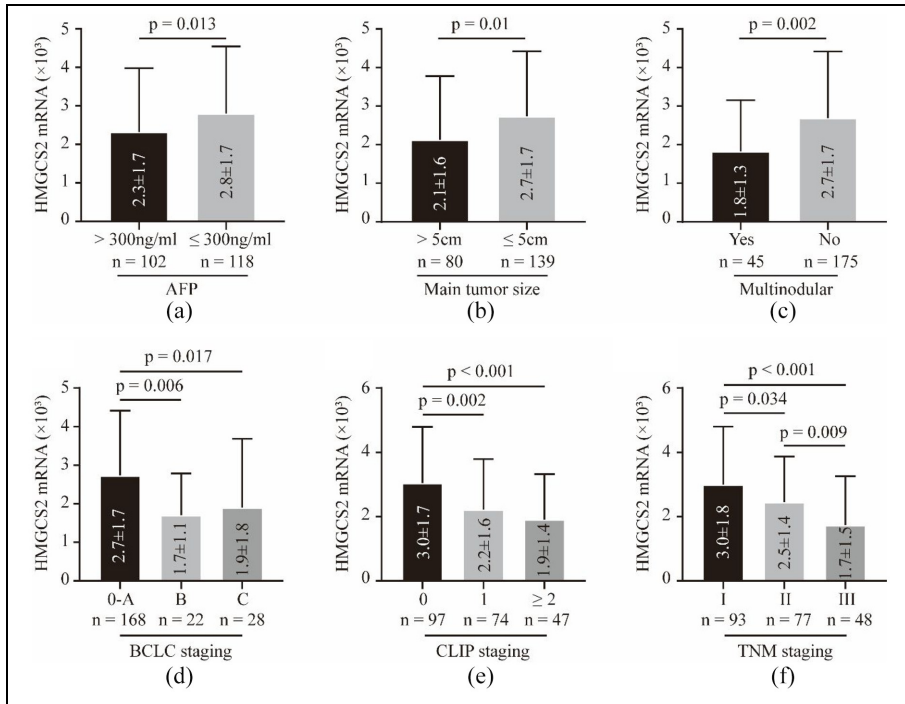


Figure 6. In GSE14520, HMGCS2 expression comparison by clinico-pathological features. HMGCS2 mRNA was significantly downregulated in HCC patients with AFP >300 ng/ml (a), main tumor size >5 cm (b), multinodular status (c), advanced BCLC staging (d), advanced TNM staging (e), and advanced CLIP staging (f).

HMGCS2 expression in human cancers is different. Upregulated HMGCS2 was found in estrogen receptor-negative breast cancer and apocrine carcinoma in the breast,^{33,34} whereas downregulated HMGCS2 was observed in poorly differentiated colon cancer.³⁵ Our results indicated that HMGCS2 was significantly downregulated in tumor tissues and expressed at lower levels in HCC patients with advanced pathological characteristics and clinical staging. While the expression of HMGCS2 mRNA in PBMCs from HCC patients was not verified. HMGCS2 downregulation in tumors contributed to unfavorable survival in HCC patients. These findings were in line with previous reports.^{36,37} In clinical practice, a low level of HMGCS2 in tumors was correlated with poor tumor differentiation, vascular invasion, higher pathological grades and clinical stages as well as worse survival in HCC patients.^{36,37} Wang et al indicated that ketogenic diet feeding could inhibit HCC tumor growth by increasing the HMGCS2 protein expression, while knockdown of HMGCS2 could enhance the cell growth ability in HCC cells. A reverse correlation was also observed between HMGCS2 expression and tumor size.³⁸

From an experimental perspective, HMGCS2 silencing could promote cell proliferation, migration, and xenograft tumorigenesis by enhancing the epithelial-mesenchymal transition process, activating the c-Myc/cyclinD1 and ERK/c-Jun pathways, and suppressing the caspase-dependent apoptosis pathway.^{36,37} HMGCS2 knockdown decreased ketone production, whereas HMGCS2 overexpression increased the level of intracellular ketones, resulting in the inhibition of cell growth and migration.³⁷ Moreover, a ketogenic diet strategy significantly inhibited liver cancer cell growth in mice.³⁷ In the hepatoma cell line HepG2, HMGCS2 could induce fatty acid β -oxidation (FAO) and ketogenesis. Additionally, RNAi-induced HMGCS2 downregulation abolished the PPAR α -mediated FAO process. Moreover, HMGCS2 activation specifically induced fibroblast growth factor 21 expression.³⁹ Tumors from HMGCS2 knockdown HCC cells showed increased lipid synthesis-related markers expression and increased lipids amount (fatty acids, triglycerides, and cholesterol), which were correlated with tumor weight.³⁸ Taken together, HMGCS2 might function as a tumor suppressor mainly through involvement in the regulation of ketone body signaling and interaction with the PPAR α pathway.³⁹

Additionally, HMGCS2 was significantly decreased in colorectal cancers and correlated with microvessel density, which indicated that HMGCS2 might negatively regulate tumor angiogenesis. In rectal cancer received neoadjuvant concurrent chemoradiotherapy, HMGCS2 overexpression was significantly associated with advanced tumor or nodal status and lower tumor regression grade and indicated poor prognosis for disease-free survival, local recurrence-free survival, and metastasis-free survival.⁴⁰ Knockdown of HMGCS2 contributed to endothelial cell tube formation in HT-29 cells.⁴¹ Conversely, a report by Chen et al indicated that higher HMGCS2 levels were significantly associated with advanced TNM stage, lymph node metastasis and tumor recurrence and poor survival in colorectal cancer patients and in oral squamous cell carcinoma patients.⁴² In prostate cancer, low HMGCS2 expression was significantly associated with advanced tumor pathological grade, distant metastasis and shorter disease-free survival and recurrence-free survival.⁴³ In esophageal squamous cell carcinoma (ESCC) patients, HMGCS2 downregulation occurred and was correlated with tumor cell differentiation, Pt status, and TNM stage. Downregulation of HMGCS2 contributed to a poorer OS and relapse-free survival in ESCC patients.⁴⁴

Even though controversial findings of HMGCS2 exist in colon cancer,^{35,41,42} a consensus has been reached that HMGCS2 is downregulated in HCC tumors and contributes to cancer aggressiveness and unfavorable outcomes in HCC patients. However, our research has some limitations. Firstly, no experiments were performed to address the effects of HMGCS2 on hepatoma cellular functions. Secondly, our follow-up data of HCC patients were not available, and our results from TCGA and GSE14520 datasets were not verified by prospective cohorts. Thirdly, all HCC patients in GSE14520 were HBV related, the extrapolation of this study needs further investigation. Hence, we suggest that well-designed studies

with large samples for investigating the predictive value of HMGCS2 are essentially needed.

In addition, our univariate Cox regression model revealed that a low level of UGP2 in tumor tissues might be a risk factor for OS in HCC patients. However, no significance was found in GSE14520 and the TCGA datasets when adjusting for cancer staging and HMGCS2 expression in the multivariate model. Recently, a single gene-centered study showed that UGP2 mRNA and protein were significantly downregulated in HCC tumors, and low UGP2 contributed to advanced tumor staging, aggressive cancer clinicopathology, and unfavorable survival in HCC patients.⁴⁵ Our findings were partly consistent with previous literature; however, the roles of UGP2 in HCC development still need to be elucidated in the future.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China (81803901 and 81903986). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Ethics statement

This study is a secondary analysis based on public datasets. As declared in the original datasets, every subject provided verified informed consent. The informed consents and study protocol were reviewed and approved by Institutional Review Board of the participating institutes.

ORCID iD

Zongguo Yang  <https://orcid.org/0000-0002-6623-4841>

Data availability

Datasets of the current study are available from The Cancer Genome Atlas (<https://portal.gdc.cancer.gov/>) and the NCBI Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>). The combined datasets were available from corresponding authors with reasonable request.

Supplemental material

Supplemental material for this article is available online.

References

1. Mejia JC and Pasko J. Primary liver cancers: intrahepatic cholangiocarcinoma and hepatocellular carcinoma. *Surg Clin North Am* 2020; 100(3): 535–549.
2. Omata M, Cheng AL, Kokudo N, et al. Asia-Pacific clinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update. *Hepatol Int* 2017; 11(4): 317–370.
3. Petrick JL, Kelly SP, Altekruse SF, et al. Future of hepatocellular carcinoma incidence in the United States forecast through 2030. *J Clin Oncol* 2016; 34(15): 1787–1794.
4. Xie S, Jiang X, Zhang J, et al. Identification of significant gene and pathways involved in HBV-related hepatocellular carcinoma by bioinformatics analysis. *PeerJ* 2019; 7: e7408.
5. Zhuang L, Yang Z and Meng Z. Upregulation of BUB1B, CCNB1, CDC7, CDC20, and MCM3 in tumor tissues predicted worse overall survival and disease-free survival in hepatocellular carcinoma patients. *Biomed Res Int* 2018; 2018: 7897346.
6. Guo W, Tan HY, Wang N, et al. Deciphering hepatocellular carcinoma through metabolomics: from biomarker discovery to therapy evaluation. *Cancer Manag Res* 2018; 10: 715–734.
7. Zhang B and Finn RS. Personalized clinical trials in hepatocellular carcinoma based on biomarker selection. *Liver Cancer* 2016; 5(3): 221–232.
8. West CA, Black AP and Mehta AS. Analysis of hepatocellular carcinoma tissue for biomarker discovery. In: Hoshida Y (ed.) *Hepatocellular carcinoma: translational precision medicine approaches*. Cham, CH: Humana Press, 2019, pp.93–107.
9. Roh SW, Abell GC, Kim KH, et al. Comparing microarrays and next-generation sequencing technologies for microbial ecology research. *Trends Biotechnol* 2010; 28(6): 291–299.
10. Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. *Science* 2013; 339(6127): 1546–1558.
11. Teufel A. Bioinformatics and database resources in hepatology. *J Hepatol* 2015; 62(3): 712–719.
12. Yang Z, Zhuang L, Szatmary P, et al. Upregulation of heat shock proteins (HSPA12A, HSP90B1, HSPA4, HSPA5 and HSPA6) in tumour tissues is associated with poor outcomes from HBV-related early-stage hepatocellular carcinoma. *Int J Med Sci* 2015; 12(3): 256–263.
13. Zhuang L, Zhang Y, Meng Z, et al. Oncogenic Roles of RAD51AP1 in tumor tissues related to overall survival and disease-free survival in hepatocellular carcinoma. *Cancer Control* 2020; 27(1): 1073274820977149.
14. Shen S, Lin Y, Yuan X, et al. Biomarker microRNAs for diagnosis, prognosis and treatment of hepatocellular carcinoma: a functional survey and comparison. *Sci Rep* 2016; 6: 38311.
15. Roessler S, Jia HL, Budhu A, et al. A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. *Cancer Res* 2010; 70(24): 10202–10212.
16. Roessler S, Long EL, Budhu A, et al. Integrative genomic identification of genes on 8p associated with hepatocellular carcinoma progression and patient survival. *Gastroenterology* 2012; 142(4): 957–966, e912.

17. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012; 2(5): 401–404.
18. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 2013; 6(269): 11.
19. Gautier L, Cope L, Bolstad BM, et al. Affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004; 20(3): 307–315.
20. Dong X, Lin L, Zhang R, et al. TOBMI: trans-omics block missing data imputation using a k-nearest neighbor weighted approach. *Bioinformatics* 2019; 35(8): 1278–1283.
21. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43(7): e47.
22. Nikolayeva O and Robinson MD. EdgeR for differential RNA-seq and ChIP-seq analysis: an application to stem cell biology. *Methods Mol Biol* 2014; 1150: 45–79.
23. Robinson MD, McCarthy DJ and Smyth GK. EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010; 26(1): 139–140.
24. Yu G, Wang LG, Han Y, et al. ClusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012; 16(5): 284–287.
25. Zhang C, Zheng Y, Li X, et al. Genome-wide mutation profiling and related risk signature for prognosis of papillary renal cell carcinoma. *Ann Transl Med* 2019; 7(18): 427.
26. Tibshirani R. The Lasso method for variable selection in the Cox model. *Stat Med* 1997; 16(4): 385–395.
27. Engebretsen S and Bohlin J. Statistical predictions with glmnet. *Clin Epigenetics* 2019; 11(1): 123.
28. Hanahan D and Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011; 144(5): 646–674.
29. Georgakopoulos-Soares I, Chartoumpakis DV, Kyriazopoulou V, et al. EMT factors and metabolic pathways in cancer. *Front Oncol* 2020; 10: 499.
30. Rathmell WK, Rathmell JC and Linehan WM. Metabolic pathways in kidney cancer: current therapies and future directions. *J Clin Oncol*. Epub ahead of print 29 October 2018. DOI: 10.1200/JCO.2018.79.2309.
31. La Vecchia S and Sebastian C. Metabolic pathways regulating colorectal cancer initiation and progression. *Semin Cell Dev Biol* 2020; 98: 63–70.
32. Boroughs LK and DeBerardinis RJ. Metabolic pathways promoting cancer cell survival and growth. *Nat Cell Biol* 2015; 17(4): 351–359.
33. Gromov P, Espinoza JA, Talman ML, et al. FABP7 and HMGCS2 are novel protein markers for apocrine differentiation categorizing apocrine carcinoma of the breast. *PLoS One* 2014; 9(11): e112024.
34. Wang J, Shidfar A, Ivancic D, et al. Overexpression of lipid metabolism genes and PBX1 in the contralateral breasts of women with estrogen receptor-negative breast cancer. *Int J Cancer* 2017; 140(11): 2484–2497.
35. Camarero N, Mascaró C, Mayordomo C, et al. Ketogenic HMGCS2 Is a c-Myc target gene expressed in differentiated cells of human colonic epithelium and down-regulated in colon cancer. *Mol Cancer Res* 2006; 4(9): 645–653.
36. Su SG, Yang M, Zhang MF, et al. MiR-107-mediated decrease of HMGCS2 indicates poor outcomes and promotes cell migration in hepatocellular carcinoma. *Int J Biochem Cell Biol* 2017; 91(Pt A): 53–59.

37. Wang YH, Liu CL, Chiu WC, et al. HMGCS2 mediates ketone production and regulates the proliferation and metastasis of hepatocellular carcinoma. *Cancers (Basel)* 2019; 11(12): 1876.
38. Wang YH, Suk FM and Liao YJ. Loss of HMGCS2 enhances lipogenesis and attenuates the protective effect of the ketogenic diet in liver cancer. *Cancers (Basel)* 2020; 12(7): 1797.
39. Vila-Brau A, De Sousa-Coelho AL, Mayordomo C, et al. Human HMGCS2 regulates mitochondrial fatty acid oxidation and FGF21 expression in HepG2 cell line. *J Biol Chem* 2011; 286(23): 20423–20430.
40. Lee YE, He HL, Shiue YL, et al. The prognostic impact of lipid biosynthesis-associated markers, HSD17B2 and HMGCS2, in rectal cancer treated with neoadjuvant concurrent chemoradiotherapy. *Tumour Biol* 2015; 36(10): 7675–7683.
41. Zou K, Hu Y, Li M, et al. Potential role of HMGCS2 in tumor angiogenesis in colorectal cancer and its potential use as a diagnostic marker. *Can J Gastroenterol Hepatol* 2019; 2019: 8348967.
42. Chen SW, Chou CT, Chang CC, et al. HMGCS2 enhances invasion and metastasis via direct interaction with PPARalpha to activate Src signaling in colorectal cancer and oral cancer. *Oncotarget* 2017; 8(14): 22460–22476.
43. Wan S, Xi M, Zhao HB, et al. HMGCS2 functions as a tumor suppressor and has a prognostic impact in prostate cancer. *Pathol Res Pract* 2019; 215(8): 152464.
44. Tang H, Wu Y, Qin Y, et al. Predictive significance of HMGCS2 for prognosis in resected Chinese esophageal squamous cell carcinoma patients. *Onco Targets Ther* 2017; 10: 2553–2560.
45. Hu Q, Shen S, Li J, et al. Low UGP2 expression is associated with tumour progression and predicts poor prognosis in hepatocellular carcinoma. *Dis Markers* 2020; 2020: 3231273.

Author biographies

Rongrong Ding is an Associate Chief Physician in Institute of Shanghai Public Health Clinical Center, Fudan University, and has long been engaged in the Department of Hepatobiliary Medicine. Her research focuses on diagnosis and treatment of liver diseases.

Tianyou Chen is an Attending Physician in Institute of Shanghai Public Health Clinical Center, Fudan University, and has long been engaged in the Department of Interventional Medicine. His research focuses on the interventional therapy of liver cancers.

Yuan Zhang is a Postgraduate in Institute of Shanghai Public Health Clinical Center, Fudan University. Her research focuses on the mechanisms of hepatocarcinogenesis.

Xiaorong Chen is a Chief Physician in Institute of Shanghai Public Health Clinical Center, Fudan University, and has long been engaged in the Department of Integrative Medicine. Her research focuses on the alternative therapies of the acute and chronic liver diseases.

Liping Zhuang is an Attending Physician in Institute of Fudan University Shanghai Cancer Center, and has been long engaged in the Department of Integrative Oncology. Her research focuses on the comprehensive treatment of liver cancers.

Zongguo Yang is an Attending Physician in Institute of Shanghai Public Health Clinical Center, Fudan University, and has long been engaged in the Department of Integrative Medicine. His research lies in the integrative treatment of liver diseases, and mechanisms of hepatocarcinogenesis.