Identification of Flap endonuclease 1 as a potential core gene in hepatocellular carcinoma by integrated bioinformatics analysis

Chuanfei Li [1], **Feng Qin** [2], **Hao Hong** [3], **Hui Tang** [4], **Xiaoling Jiang** [5], **Shuangyan Yang** [4], **Zhechuan Mei** [1], **Di Zhou** [6]
Corresp. 6

[1] Department of Gastroenterology, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China
[2] Department of Infectious Diseases, The People's Hospital of Shi Zhu, Chongqing, China
[3] Department of Orthopaedics, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China
[4] Department of Infectious Diseases, Institute for Viral Hepatitis, The Key Laboratory of Molecular Biology for Infectious Diseases, Chinese Ministry of Education, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China
[5] Tongnan District People' Hospital, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China
[6] Department of Radiology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China

Corresponding Author: Di Zhou
Email address: zhoudi@cqmu.edu.cn

Hepatocellular carcinoma (HCC) is a common yet deadly form of malignant cancer. However, the specific mechanisms involved in HCC diagnosis have not yet fully elucidated. Herein, we screened four publically available Gene Expression Omnibus (GEO) expression profiles (GSE14520, GSE29721, GSE45267 and GSE60502), and used them to identify 409 differentially expressed genes (DEGs), including 142 and 267 up- and down-regulated genes, respectively. The DAVID database was used to look for functionally enriched pathways among DEGs, and the STRING database and Cytoscape platform were used to generate a protein-protein interaction (PPI) network for these DEGs. The cytoHubba plug-in was utilized to detect185 hub genes, and three key clustering modules were constructed with the MCODE plug-in. Gene functional enrichment analyses of these three key clustering modules were further performed, and nine core genes including BIRC5, DLGAP5, DTL, FEN1, KIAA0101, KIF4A, MCM2, MKI67, and RFC4, were identified in the most critical cluster. Subsequently, the hierarchical clustering and expression of core genes in TCGA liver cancer tissues were analyzed using the UCSC Cancer Genomics Browser, and whether elevated core gene expression was linked to a poor prognosis in HCC patients was assessed using the GEPIA database. The PPI of the nine core genes revealed an interaction between FEN1, MCM2, RFC4, and BIRC5. Furthermore, FEN1 expression was positively correlated with that of three other three core genes in TCGA liver cancer tissues. FEN1 expression in HCC and other tumor types was assessed with the FIREBROWSE and ONCOMINE databases, and results were verified in HCC samples and hepatoma cells. FEN1

levels were also positively correlated with tumor size and incidence of distant metastasis. In conclusion, we identified nine core genes associated with HCC development, offering novel insight into HCC progression. In particular, the aberrantly elevated FEN1 may represent a potential biomarker for HCC diagnosis and treatment.

Identification of Flap endonuclease 1 as a potential core gene in hepatocellular carcinoma by integrated bioinformatics analysis

Chuanfei Li[1], Feng Qin[2], Hao Hong[3], Hui Tang[4], Xiaoling Jiang[5], Shuangyan Yang[4], Zhechuan Mei[1], Di Zhou[6]

[1] Department of Gastroenterology, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China
[2] Department of Infectious Diseases, The People's Hospital of Shi Zhu, Chongqing, China
[3] Department of Orthopaedics, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China
[4] Department of Infectious Diseases, Institute for Viral Hepatitis, The Key Laboratory of Molecular Biology for Infectious Diseases, Chinese Ministry of Education, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China
[5] Tongnan District People' Hospital, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China
[6] Department of Radiology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China

Corresponding Author:
Di Zhou
No.1 Youyi Road, Yuzhong District, Chongqing, 400016, China
Email address: zhoudi@cqmu.edu.cn

ABSTRACT

Hepatocellular carcinoma (HCC) is a common yet deadly form of malignant cancer. However, the specific mechanisms involved in HCC diagnosis have not yet fully elucidated. Herein, we screened four publically available Gene Expression Omnibus (GEO) expression profiles (GSE14520, GSE29721, GSE45267 and GSE60502), and used them to identify 409 differentially expressed genes (DEGs), including 142 and 267 up- and down-regulated genes, respectively. The DAVID database was used to look for functionally enriched pathways among DEGs, and the STRING database and Cytoscape platform were used to generate a protein-protein interaction (PPI) network for these DEGs. The cytoHubba plug-in was utilized to detect185 hub genes, and three key clustering modules were constructed with the MCODE plug-in. Gene functional enrichment analyses of these three key clustering modules were further performed, and nine core genes including BIRC5, DLGAP5, DTL, FEN1, KIAA0101, KIF4A, MCM2, MKI67, and RFC4, were identified in the most critical cluster. Subsequently, the

hierarchical clustering and expression of core genes in TCGA liver cancer tissues were analyzed using the UCSC Cancer Genomics Browser, and whether elevated core gene expression was linked to a poor prognosis in HCC patients was assessed using the GEPIA database. The PPI of the nine core genes revealed an interaction between FEN1, MCM2, RFC4, and BIRC5. Furthermore, FEN1 expression was positively correlated with that of three other three core genes in TCGA liver cancer tissues. FEN1 expression in HCC and other tumor types was assessed with the FIREBROWSE and ONCOMINE databases, and results were verified in HCC samples and hepatoma cells. FEN1 levels were also positively correlated with tumor size and incidence of distant metastasis. In conclusion, we identified nine core genes associated with HCC development, offering novel insight into HCC progression. In particular, the aberrantly elevated FEN1 may represent a potential biomarker for HCC diagnosis and treatment.

Subjects Bioinformatics, Gastroenterology and Hepatology
Keywords Hepatocellular carcinoma, Core genes, Bioinformatics analysis, Flap endonuclease 1

INTRODUCTION

Hepatocellular carcinoma remains among the most common and deadly forms of cancer globally, posing a significant threat to human life (Forner et al. 2018). HCC and other tumors develop as the result of the long-term accumulation of genetic mutations. Although a large number of biomarkers for the diagnosis of HCC have been identified (C et al. 2018), the specific molecular mechanisms governing HCC development, recurrence, and treatment remain obscure. Therefore, it is essential to identify and exploit novel biomarkers involved in HCC onset and progression to better understand the pathogenesis of HCC.

Between human genome sequencing efforts and the rapid development of gene sequencing technologies, precision medicine has risen to prominence and been widely employed in the field of oncology (IR et al. 2017; R et al. 2019). Precision medicine relies upon initially exploring potential therapeutic targets via high-throughput sequencing technologies (PN, 2012), as these technologies allow for the large-scale investigation of altered gene expression in the context of disease (Chen et al. 2010). However, sequencing results are often limited and inconsistent owing to the heterogeneity of samples in independent studies, and due to the fact that most studies focus on one cohort. As such, this study sought to analyze genes involved in liver cancer development using a range of available liver cancer-related gene chip datasets, with the goal of identifying potential novel molecular targets for liver cancer treatment and diagnosis.

For the purposes of this study, four HCC related Gene Expression Omnibus (GEO) database datasets were downloaded: GSE14520 (S et al. 2012; S et al. 2010), GSE29721 (B et al. 2011), GSE45267 (CL et al. 2018) and GSE60502 (YH et al. 2014). By analyzing these four datasets, we were able to identify 409 differentially expressed genes (DEGs) linked with HCC, of which 142 and 267 were up- and down-regulated, respectively. The DAVID database was then used for a functional enrichment analysis of these DEGs, while the STRING database and Cytoscape were utilized to generate a protein-protein interaction (PPI) network, and three clustering

modules were filtered out with the MCODE plug-in, among which clustering module 1 was most associated with HCC. In addition, nine core genes including BIRC5, DLGAP5, DTL, FEN1, KIAA0101, KIF4A, MCM2, MKI67, and RFC4, were identified within clustering module 1, and these corresponded to the hub genes in our PPI network. Using the GEPIA database, we performed survival analyses of patients based on expression of these nine core genes, revealing that their overexpression was linked to a poorer prognosis in HCC patients. When we surveyed the literature surrounding these genes, we found that all except for FEN1 had previously been confirmed to play a vital role in HCC. Existing studies have shown that FEN1 is highly expressed in various cancers, such as brain (Nikolova et al. 2009), lung (He et al. 2017; Zhang et al. 2018), breast (Abdel-Fatah et al. 2014), gastric (Wang et al. 2014), prostate (Lam et al. 2006) and pancreatic cancer (Isohookana et al. 2018), but its expression and role in HCC remains unclear. Our previous data indicated that FEN1 expression was elevated in HCC tumors, and this was confirmed upon comparing HCC samples and hepatoma cells to appropriate controls. Together, the results of these analyses suggest that FEN1 may be a core gene orchestrating the progression of HCC.

MATERIALS & METHODS

Data collection and DEG validation

Four liver cancer-related datasets (GSE14520, GSE29721, GSE45267 and GSE60502), were downloaded from GEO (http://www.ncbi.nlm.nih.gov/geo/). These datasets contained a total of 299 tumor samples and 289 non-tumor samples. GSE14520 contained 225 liver cancer samples and 220 adjacent controls; GES29721 contained 10 pairs of liver cancer samples and adjacent control tissue; GSE45267 contained 46 liver cancer samples and 41 adjacent controls; and GSE60502 contained 18 pairs of liver cancer tissues and adjacent controls. The specific platform information for these four datasets is compiled in Table 1. GEO2R was utilized to identify DEGs in these studies, using the screening criteria: $|\log FC \text{ (fold change)}| \geq 1$, $P<0.05$, and adjusted $P<0.05$. In addition, genes with multiple probe set or probe sets lacking matched gene symbols were removed or averaged, respectively. Then, the overall DEGs, as well as those that were up- and down-regulated in the four datasets were intersected and visualized using Funrich (v 3.0, http://funrich.org/index.html).

Functional enrichment analyses

Gene ontology analyses focuses on three domains: biological processes (BP), cellular components (CC), and molecular functions (MF), and such analyses are commonly used to understand the biological functions, pathways, or localization of DEGs. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis database surveys as a valuable resource for assessing how particular DEGs may be involved in or influenced by specific signaling pathways and disease states. We used DAVID (https://david.ncifcrf.gov/) for functional enrichment analyses, with a $P < 0.05$ cutoff for significance.

PPI network analysis

The STRING database (v10.5; https://stringdb.org/) was used to generate a DEG PPI, with a minimum interaction score cutoff of 0.4. Cytoscape (v 3.4.0, https://cytoscape.org/) was used for network visualization, and the CytoHubba plug-in was used to identify hub genes with the criteria of filtering degree≥10. The MCODE plug-in was used to construct key clustering modules (MCODE score>10, degree cut-off=2, node score cut-off=0.2, Max depth=100 and k-score= 2).

Validation of core genes

Hub genes among overall DEGs and the most critical clustering module were identified through the CytoHubba plug-in, and intersecting core genes were identified. The UCSC Cancer Genomics Browser (https://genome-cancer.ucsc.edu/) was then used for hierarchical clustering of these core genes. In addition, the expression profiles of these core genes in 421 TCGA liver cancer tissues, including 50 solid normal tissues and 371 primary tumors, were determined by analyzing available datasets. A core gene PPI network was constructed with the cBioportal online database (http://www.cbioportal.org/). Furthermore, correlations between the expression of FEN1 and MCM2, RFC4, or BIRC5 in TCGA liver cancer tissues were investigated. Finally, we used the ONCOMINE (https://www.oncomine.org/resource/login.html) and FIREBROWSE online database (http://firebrowse.org/) to investigate the FEN1 expression in various cancers including HCC.

Survival analysis

The GEPIA database (http://gepia.cancer-pku.cn/) was used to conduct survival analyses based on core gene expression, with hazard ratios (HRs) and 95% confidence intervals being calculated, and logrank P value < 0.05 being the threshold of statistical significance.

Clinical samples

We obtained 34 paired HCC tumor tissues as well as corresponding adjacent non-cancerous tissues from our Hospital's Department of Hepatobiliary Surgery, with all patients providing informed consent. The content of the informed consent includes research purposes, risks and discomfort, benefits, privacy issues, etc. The study was examined and approved by the Second affiliated Hospital of Chongqing Medical University Ethics Committee.

Cell culture

Seven human liver cancer cell lines (SMMC-7721, BEL-7404, HCCLM3, HepG2, MHCC97-H, SK-HEP-1, and Huh-7) as well as normal human liver HL7702 cells were donated by the Institute for Viral Hepatitis, Chongqing Medical University. All cells were grown in high glucose DMEM (Gibco, USA) containing 10% FBS (Corning, USA) at 37°C in a 5% CO2 incubator.

Hematoxylin and eosin (H&E) staining

Paraffin-embedded tissues were dewaxed in xylene I, II, and III for 20 minutes each, and then dehydrated in an ethanol gradient (100%, 95%, 90%, 80%, and 70%), 3 minutes per step. Sections were then rinsed with distilled water for 5 minutes, and the nucleus was counterstained with hematoxylin for 3 minutes. Sections were washed again in water, followed by differentiation for 30 seconds in a 75% hydrochloric acid alcohol solution, and blue color was returned by washing with distilled water for 5 minutes. A red dye was then used for counterstaining for 5 minutes, after which samples were dehydrated in 70%, 80%, 95%, and 100% ethanol, 1 minute per concentration. Sections were then cleared with xylene and sealed using neutral gum.

IHC analysis

IHC staining was conducted as previously described (WG et al. 2016). Briefly, paraffin-embedded sections were incubated at 56 °C for 2 h, and then 3% hydrogen peroxide was used for antigen retrieval. Afterwards, the sections were incubated with rabbit anti-human FEN1 (1:100, A1175, ABclonal, China) at 4 °C overnight. Then, sections were probed for 1 h using HRP-conjugated secondary antibodies at 37 °C, after which a DAB substrate kit was utilized, and hematoxylin was used for nuclear staining. The Image-Pro Plus (IPP) software (Media Cybernetics, Rockville, MD, USA) was used to quantify staining intensity. The frequency of positively-stained cells was determined on a 0 – 100 scale, while staining intensity was scored as follows: 0 = negative; 1 = weak; 2 = moderate; 3 = strong. These two scores were then multiplied together to yield an IHC score between 0 and 300. The final scores were assigned by two independent pathologists. The mean IHC score was used as a cutoff value to separate patients into low- and high-expression groups.

RT-qPCR analysis

TRIzol (ThermoFisher Scientific, USA) was used for total RNA extraction, and RNA was then reverse transcribed with the PrimeScript RT-PCR kit (Takara Bio, Dalian, China) based on provided protocols. The SYBR Premix Ex Taq II (Takara, Japan) kit was used to conduct RT-qPCR analysis on a Bio-Rad CFX96 Real-Time System (Bio-Rad, Hercules, CA). mRNA levels were determined via the $2^{-\Delta\Delta Ct}$ method, with GAPDH used for normalization. Primers used were: GAPDH F：5'- GGTGGTCTCCTCTGACTTCAACA -3' and R：5'- GTTGCTGTAGCCAAATTCGTTGT-3', FEN1 F: 5'- CTGTGGACCTCATCCAGAAGCA -3' and R: 5'- CCAGCACCTCAGGTTCCAAGA -3'.

Statistical analysis

Statistical analyses and graphing were performed with SPSS v19.0 (SPSS Inc., USA) and Graph Pad Prism v8.0 (Graph Pad Software, USA), respectively. Data are means ± standard deviation (SD). Student's t-tests were used to compare groups. Fisher's exact test was used to assess correlations between the expression of FEN1 and HCC patient clinicopathological features.

Spearman's correlation analyses were used to compare the expression of pairs of genes in TCGA liver cancer tissues. P<0.05 was the significance threshold (*P<0.05, **P<0.01).

RESULTS

HCC-associated DEG identification

In this study, we screened 1088 total DEGs (505 and 583 up- and down-regulated, respectively) in GSE14520, 1449 total DEGs (837 and 612 up- and down-regulated, respectively) in GSE29721, 1604 total DEGs (713 and 891 up- and down-regulated, respectively) in GSE45267, and 1533 total DEGs (792 and 741 up- and down-regulated, respectively) in GSE60502. Based on these datasets, a total of 409 overlapping DEGs were identified among these four datasets (142 and 267 up- and down-regulated, respectively), as visualized with the Funrich software (Figure 1). DEGs are listed in Table S1.

DEG functional enrichment analyses

To explore the biological activities of these DEGs, we used the DAVID database to conduct GO and KEGG enrichment analysis. With respect to BPs, up-regulated DEGs were primarily enriched in processes such as mitotic nuclear division, cell division, cell cycle, DNA replication, and mitotic sister chromatid segregation (Figure S1A), while down-regulated DEGs were primarily enriched in processes such as redox process, the cytochrome 450 pathway, drug metabolism, and negative regulation of growth (Figure S2A). With respect to CCs, up-regulated DEGs were primarily enriched in the nucleoplasm, nucleus, cytoplasm, spindle, and cellular intermediates (Figure S1B), while down-regulated DEGs were mostly enriched in extracellular exosomes, organelle membranes, blood microparticles, extracellular regions, and the mitochondrial matrix (Figure S2B). With respect to MFs, up-regulated DEGs were primarily enriched in functions such as protein binding, ATP binding, DNA helicase activity, protein kinase binding, single-stranded DNA binding, and chromatin binding (Figure S1C), whereas down-regulated DEGs were primarily associated with iron ion binding, oxidoreductase activity, heme binding, monooxygenase activity, and oxygen binding (Figure S2C).

A KEGG analysis revealed that up-regulated DEGs were particularly enriched in pathways such as the cell cycle, DNA replication, P53 signaling, and tumor pathways (Figure S1D). However, down-regulated DEGs were mostly associated with metabolic pathways, fatty acid degradation, chemical carcinogenesis, and PPAR signaling (Figure S2D).

PPI network and module analyses

To better understand interactions among DEGs, the STRING online database was used to generate a PPI network consisting of 403 nodes and 3502 edges, which was visualized using Cytoscape. Six of the 409 DEGs were not included in this network (Figure 2A). This network was then analyzed using the MCODE plug-in, and three clustering modules were filtered out according to the chosen screening conditions. Clustering module 1 scored 58.492 with 62 nodes and 1784 edges (Figure 2B), clustering module 2 scored 11.529 with 18 nodes and 98 edges

(Figure 2C), and clustering module 3 scored 10.917 with 25 nodes and 131 edges (Figure 2D). The genes in clustering module 1 were up-regulated DEGs, whereas those in the other two modules were primarily down-regulated DEGs.

Functional enrichment analysis of key clustering modules

The DAVID database was next used to explore the biological functions of genes in these key clustering modules (Tables S2-4). With respect to BPs, clustering module 1 was primarily enriched in cell differentiation, mitotic nuclear division, DNA replication, and DNA helicase activity, while clustering module 2 was primarily enriched in plasminogen activation, coagulation, cytolysis, and complement activation regulation, and clustering module 3 was primarily enriched in steroid metabolism, heterogeneous biomass metabolism, and exogenous drug catabolism. With respect to CCs, clustering module 1 was primarily enriched for the nucleoplasm, nucleus, intermediate, spindle, cytoplasm, nuclear chromosome, while clustering module 2 was primarily enriched for exosomes, extracellular regions, membrane attack complexes, and extracellular vesicles, and clustering module 3 was primarily enriched for organelle membranes, the endoplasmic reticulum membrane, and high-density lipoproteins. With respect to MFs, clustering module 1 was primarily enriched for protein binding, ATP binding, DNA helicase activity, protein kinase binding, and DNA binding, while clustering module 2 was primarily enriched for endopeptidase activity, transcription factor binding, steroid binding, RNA polymerase II transcription factor activity, and enzyme binding activity, and clustering module 3 was primarily enriched for aerobic binding, iron ion binding and heme binding.

A KEGG analysis revealed that clustering module 1 was primarily enriched in the cell cycle, oocyte meiosis, DNA replication, and p53 signaling, while clustering module 2 was primarily enriched in the complement system, prion disease, and systemic lupus erythematosus, and clustering module 3 was primarily enriched in chemical carcinogenesis, retinol metabolism, P450 drug metabolism, and metabolic pathways.

Identification of core genes and analysis of their clinical significance

We next sought to identify core genes involved in HCC based on their levels of interaction via analyzing our PPI network using the Cytoscape program. Based on our clustering module analysis, we found that clustering module 1 included 62 genes and was closely related to the progression of HCC. We then identified nine total genes (BIRC5, DLGAP5, DTL, FEN1, KIAA0101, KIF4A, MCM2, MKI67, and RFC4) based on the intersecting genes among the top 40 genes derived from 12 different algorithms by the cytoHubba plug-in (Table S5). In addition, we identified 185 hub genes across 403 nodes in our PPI network based on the filtering degree $\geqslant$ 10 criteria (Table S6). We found that the nine identified core genes belonged to this larger subset of hub genes (Figure 3A).

To investigate core gene expression in HCC, a hierarchical clustering analysis was performed using the UCSC Cancer Genomics Browser, revealing that these nine core genes were highly expressed in most liver cancer samples (Figure 3C). Next, the expression profiles of these

nine core genes in 421 TCGA liver cancer tissues, including 50 solid normal tissues and 371 primary tumors, were downloaded and analyzed, revealing that the expression of these core genes was significantly elevated in HCC (Figure 3B). We then further assessed correlations between core gene expression levels and patient prognosis in 182 total HCC samples, revealing that BIRC5 expression (HR=2, logrank P= 6.7e-05) was correlated with worse overall survival (OS) for HCC patients, as was that of DLGAP5 (HR=1.9, logrank P= 0.00039), DTL (HR=1.7, logrank P= 0.0049), FEN1 (HR=1.5, logrank P= 0.022), KIAA0101 (HR=1.7, logrank P= 0.002), KIF4A (HR=1.8, logrank P= 0.001), MCM2 (HR=1.7, logrank P= 0.0022), MKI67 (HR=1.9, logrank P= 0.00045), and RFC4 (HR=1.7, logrank P= 0.004) (Figure 4). We also found that the expression of BIRC5 (HR=1.6, logrank P= 0.002) was correlated with decreased disease-free survival (DFS) for HCC patients, as was that of DLGAP5 (HR=1.6, logrank P= 0.0033), DTL (HR=1.6, logrank P= 0.0016), FEN1 (HR=1.5, logrank P= 0.0075), KIAA0101 (HR=1.6, logrank P= 0.0022), KIF4A (HR=1.6, logrank P= 0.0011), MCM2 (HR=1.6, logrank P= 0.0034), MKI67 (HR=1.9, logrank P= 4.2e-05), and RFC4 (HR=1.5, logrank P= 0.011) (Figure 5). High expression of these nine core genes was associated with significantly reduced survival among HCC patients.

FEN1 may be a key candidate gene in HCC

To clarify the PPI network for these nine core genes, we next explored the cBioportal online database and identified an interaction between FEN1, MCM2, BIRC5 and RFC4 (Figure 6A). Previous studies have confirmed that MCM2, BIRC5, and RFC4 are abnormally highly expressed in HCC, and that they participate in the regulation of HCC tumor biology. However, the expression of FEN1 and its clinical significance in HCC is unclear. We therefore further analyzed the correlations between the expression of FEN1 and these three other genes. The expression of FEN1 in TCGA liver cancer tissues was positively correlated with that of MCM2 (r= 0.853, P= 0.000), BIRC5 (r= 0.809, P= 0.000), and RFC4 (r= 0.852, P= 0.000) (Figure 6B-D), suggesting that FEN1 may play as important a role in the progression of liver cancer as do MCM2, BIRC5, and RFC4. We further utilized the ONCOMINE and FIREBROWSE databases to investigate the expression of FEN1 in various cancers, including HCC. The results revealed that FEN1 expression was clearly elevated in most cancers, including bladder, breast, colorectal, esophageal, lung, and liver cancer (Figure 7A, B). In addition, we found that FEN1 was overexpressed in three HCC-related datasets (Figure 7C).

Experimental validation

In this study, 34 paired HCC and adjacent control tissues were used to verify the expression of FEN1 in HCC. The pathologic diagnosis of HCC and matched adjacent samples were confirmed by H&E staining (Figure 8A, C). FEN1 staining was localized to the nucleus and cytoplasm, with stronger expression in the HCC samples relative to the adjacent controls (Figure 8B, D). Furthermore, IHC analysis indicated that FEN1 levels were significantly higher in HCC relative

to adjacent tissues (Figure 8E), with IHC scores being significantly higher in HCC samples (Figure 8F).

The correlation between FEN1 expression and HCC patient clinicopathological features was investigated via Fisher's exact test. As shown in Table 2, we found significant correlations between FEN1 expression and tumor size ($P= 0.047 <0.05$) as well as metastasis ($P=0.013 <0.05$). FEN1 expression did not significantly correlated with gender, age, tumor multiplicity, TNM stage, pathological grade, HBsAg, or liver cirrhosis ($P>0.05$). We further found that FEN1 mRNA levels were significantly elevated in six human hepatoma cell lines relative to that in the normal human liver cell line HL7702 (Figure 8G).

DISCUSSION

The occurrence and progression of HCC are complex, with multiple cumulative genetic changes ultimately culminating in progressive disease. High-throughput technologies such as gene chips have been widely employed to elucidate the underlying mechanisms, providing an innovative and effective approach to the diagnosis, prevention, and treatment of HCC. Numerous studies have been conducted to clarify genetic changes underlying HCC development, but results to date remain inconclusive of incomplete. As such, there is further need to investigate the molecular mechanisms governing HCC.

In this study, we identified 409 total DEGs (142 and 267 up- and down-regulated, respectively) shared among four HCC datasets, and we used these genes for functional enrichment analyses. A GO analysis revealed the up-regulated DEGs to be primarily linked with cell division, the cell cycle, and DNA replication, whereas down-regulated DEGs were mostly associated with redox reactions, cytochrome 450 functionality, and negative growth regulation. A KEGG pathway analysis revealed up-regulated DEGs to mostly be associated with signaling relating to the cell cycle, DNA replication, p53 signaling, and tumor pathways, whereas down-regulated DEGs were mostly linked to metabolic pathways such as fatty acid degradation. These results suggested that the up-regulated DEGs may affect HCC progression via regulating DNA replication and the cell cycle, whereas down-regulated DEGs may be linked to HCC progression through metabolic pathways. Previous studies have indicated that the dysregulation of the cell cycle is a key hallmark of many cancer types (P et al., 2017). As one of the most important tumor suppressor genes, p53 is closely related to tumorigenesis, with at least 50% of cancer patients exhibiting p53 mutations or loss of function (C et al. 2013; G & M 2019; MP et al. 2015). P53 signaling dysregulation has repeatedly been confirmed to be linked with cancer development (M & J 2016). There is also increasing evidence that metabolism regulates cancer growth and proliferation (J & CB 2019; T et al. 2016). We further used our DEGs to generate a PPI, which consisted of three key clustering modules screened using the MCODE plug-in. Functional enrichment analyses revealed that clustering module 1 was closely related to gene mutations in the progression of HCC, and was enriched in genes linked to the cell cycle, DNA replication, and the p53 signaling pathway. Then nine core genes (BIRC5, DLGAP5, DTL, FEN1, KIAA0101, KIF4A, MCM2, MKI67 and RFC4) in clustering module 1 were screened, as these genes were

also hub genes in the overall PPI network. The hierarchical clustering and expression profiles of core genes in TCGA liver cancer tissues were analyzed using the UCSC Cancer Genomics Browser, and survival analyses suggested that aberrantly high expression of these core genes was predictive of a poor HCC patient prognosis, suggesting these core genes may be key molecular biomarkers for HCC diagnosis and treatment.

BIRC5, also known as survivin, is an anti-apoptotic protein reported to function as a potential oncogene in the context of many cancers (MJ et al. 2007). Studies have shown that BIRC5 is highly expressed in the vast majority of tumors, including HCC (C 2016). In addition, elevated BIRC5 levels have been found to be associated with histological grade, tumor size, and TNM stage in HCC patients (W et al. 2019). DLGAP5, also known as HURP, is a cell cycle regulatory gene that has been found to be highly expressed in liver cancer (Chang et al. 2011; Tsou et al. 2003). Weijia Liao (Liao et al. 2013) reported the abnormally high expression of DLGAP5 in HCC, and found that this was related to promoter methylation level, with the silencing of DLGAP5 significantly inhibiting cell proliferation, migration, and colony formation in vitro. DLGAP5 knockdown inhibited the proliferation of hepatoma cells by reducing P53 accumulation (Kuo et al. 2012). DTL is a substrate receptor for the CRL4 ubiquitin ligase, serving as a key regulator of the cell cycle and genomic stability. DTL upregulation in invasive HCC has been found to be positively correlated with tumor grade and patient survival (Chen et al. 2018b; Pan et al. 2006). FEN1, a structurally specific metal nuclease, plays a vital role in DNA damage repair (Lieber 1997) and maintenance of genomic stability (Becker et al. 2018). FEN1 has been found to be expressed at high levels in several cancer types, including those of the lung (He et al. 2017), breast (Abdel-Fatah et al. 2014), gastric (Wang et al. 2014), prostate (Lam et al. 2006) and pancreatic (Isohookana et al. 2018). FEN1 mutations have been associated with the occurrence of gastrointestinal tumors, including HCC (Liu et al. 2012b), suggesting that FEN1 may be important in the development of gastrointestinal tumors. Nevertheless, the expression and biological function of FEN1 in HCC remains unclear. KIAA0101 is a nuclear antigen-associated factor that is present in proliferating cells and which regulates proliferative processes. KIAA0101 has been found to be highly expressed in HCC, representing a potential biomarker for this tumor type that is relevant to disease treatment and patient prognosis (Yuan et al. 2007). KIAA0101 variant 1 can promote the survival of liver cancer cells by regulating p53 functionality, indicating that inhibition of KIAA0101 variant 1 may be a promising therapeutic strategy (Liu et al. 2012a). Numerous studies have revealed that KIF4A is an oncogene in the context of lung, oral, and breast cancer. High KIF4A expression was significantly correlated with tumor stage, tumor differentiation, and metastasis, and may be a biomarker for poor HCC prognosis (Hou et al. 2017). In addition, KIF4A promoted the proliferation and invasion of hepatoma cells via p53 and Akt signaling (Hou et al. 2017; Huang et al. 2018). MCM2 has been shown to be critical for chromatin composition and for the focal formation of p53 binding protein 1 in HepG2 cells (Chen et al. 2018a). High MCM2 expression has been associated with poor prognosis in HCC patients (Liu et al. 2018). MKI67 encodes a nuclear antigen expressed during the G1, S and G2-M phases in proliferating cells, and its expression levels are closely

related to tumor growth rate (T et al. 1995), histological stage (IO et al. 1995), and tumor recurrence for HCC (K et al. 1995). Higher MKI67 expression was associated with both faster HCC progression and a poorer patient prognosis (Y et al. 2015). Consistent with our results, RFC4 has been found to be overexpressed in HCC (M et al. 2009; S et al. 2019). The silencing of RFC4 expression was able to reduce HepG2 cell proliferation and promote apoptosis, and was also associated with the increased sensitivity of cells to doxorubicin and camptothecin, suggesting RFC4 may be a novel target for liver cancer treatment.

As these past results show, all of these core genes except for FEN1 have been reported to play a role in HCC progression. To further explore interactions among these nine core genes, a PPI network was generated revealing interactions between FEN1, MCM2, RFC4, and BIRC5. As expected, FEN1 expression was significantly positively correlated with that of MCM2, RFC4, and BIRC5 in TCGA liver cancer tissues, suggesting that FEN1 may be as involved in the development of HCC as are MCM2, RFC4, and BIRC5. We additionally found that FEN1 was highly expressed in multiple tumor tissues including those of HCC patients using the ONCOMINE and FIREBROWSE online databases, consistent with previous reports. We then verified this in HCC tissues from 34 patients via IHC staining. Interestingly, a correlation analysis revealed that high expression of FEN1 was significantly correlated with tumor size and metastasis, suggesting that FEN1 may be involved in the regulation of liver cancer proliferation and metastasis. In addition, our results showed that FEN1 was elevated in six human hepatoma cell lines relative to a control cell line via RT-qPCR analysis. However, the biological function of FEN1 in HCC remains to be further investigated.

CONCLUSIONS

In summary, we screened 409 HCC-associated DEGs, of which 142 and 267 were up- and down-regulated, respectively. We identified nine core genes, including BIRC5, DLGAP5, DTL, FEN1, KIAA0101, KIF4A, MCM2, MKI67, and RFC4, that were up-regulated in HCC and that may play important roles in the development or progression of this cancer. We further confirmed the expression of FEN1 in HCC, suggesting that FEN1 may have potential as a new biomarker for HCC diagnosis, treatment, or prognosis determination. Our results offer significant improvements to current understanding of HCC pathogenesis. Additional studies will be needed to validate our findings, and to confirm whether and how FEN1 regulates HCC.

ACKNOWLEDGEMENTS

None.

REFERENCES

Abdel-Fatah TM, Russell R, Albarakati N, Maloney DJ, Dorjsuren D, Rueda OM, Moseley P, Mohan V, Sun H, Abbotts R, Mukherjee A, Agarwal D, Illuzzi JL, Jadhav A, Simeonov A, Ball G, Chan S, Caldas C, Ellis IO, Wilson DM, 3rd, and Madhusudan S. 2014. Genomic and protein expression analysis reveals flap endonuclease 1 (FEN1) as a key

438 biomarker in breast and ovarian cancer. Mol Oncol 8:1326-1338.
439 10.1016/j.molonc.2014.04.009
440 B S, J H, B B, M S, M H, ZG H, and M S. 2011. Definition of the landscape of promoter DNA
441 hypomethylation in liver cancer. Cancer research 71:5891-5903.
442 Becker JR, Gallo D, Leung W, Croissant T, Thu YM, Nguyen HD, Starr TK, Brown GW, and
443 Bielinsky AK. 2018. Flap endonuclease overexpression drives genome instability and
444 DNA damage hypersensitivity in a PCNA-dependent manner. Nucleic Acids Res
445 46:5634-5650. 10.1093/nar/gky313
446 C K, MD M, F V, K Y, B N, C L, M X, Q Z, JF M, MA W, MDM L, CA M, JS W, MJ W, MC
447 W, TJ L, RK W, BJ R, and L D. 2013. Mutational landscape and significance across 12
448 major cancer types. Nature 502:333-339.
449 C S. 2016. Survivin in survival of hepatocellular carcinoma. Cancer letters 379:184-190.
450 C T, A S, and RS H. 2018. Cancer genetics, precision prevention and a call to action. Nature
451 genetics 50:1212-1218.
452 Chang ML, Lin SM, and Yeh CT. 2011. HURP expression-assisted risk scores identify prognosis
453 distinguishable subgroups in early stage liver cancer. PLoS One 6:e26323.
454 10.1371/journal.pone.0026323
455 Chen CF, Hsu EC, Lin KT, Tu PH, Chang HW, Lin CH, Chen YJ, Gu DL, Lin CH, Wu JY,
456 Chen YT, Hsu MT, and Jou YS. 2010. Overlapping high-resolution copy number
457 alterations in cancer genomes identified putative cancer genes in hepatocellular
458 carcinoma. Hepatology 52:1690-1701. 10.1002/hep.23847
459 Chen Y, Weng C, Zhang H, Sun J, and Yuan Y. 2018a. A Direct Interaction Between P53-
460 Binding Protein 1 and Minichromosome Maintenance Complex in Hepg2 Cells. Cell
461 Physiol Biochem 47:2350-2359. 10.1159/000491607
462 Chen YC, Chen IS, Huang GJ, Kang CH, Wang KC, Tsao MJ, and Pan HW. 2018b. Targeting
463 DTL induces cell cycle arrest and senescence and suppresses cell growth and colony
464 formation through TPX2 inhibition in human hepatocellular carcinoma cells. Onco
465 Targets Ther 11:1601-1616. 10.2147/ott.s147453
466 CL C, YS T, YH H, YJ L, YY S, CW S, GY C, YC Y, YS C, JT H, and JC W. 2018. Lymphoid
467 Enhancer Factor 1 Contributes to Hepatocellular Carcinoma Progression Through
468 Transcriptional Regulation of Epithelial-Mesenchymal Transition Regulators and
469 Stemness Genes. Hepatology communications 2:1392-1407.
470 Forner A, Reig M, and Bruix J. 2018. Hepatocellular carcinoma. Lancet 391:1301-1314.
471 10.1016/s0140-6736(18)30010-2
472 G DO, and M C. 2019. Mutant p53 and Cellular Stress Pathways: A Criminal Alliance That
473 Promotes Cancer Progression. Cancers 11:undefined.
474 He L, Luo L, Zhu H, Yang H, Zhang Y, Wu H, Sun H, Jiang F, Kathera CS, Liu L, Zhuang Z,
475 Chen H, Pan F, Hu Z, Zhang J, and Guo Z. 2017. FEN1 promotes tumor progression and
476 confers cisplatin resistance in non-small-cell lung cancer. Mol Oncol 11:640-654.
477 10.1002/1878-0261.12058
478 Hou G, Dong C, Dong Z, Liu G, Xu H, Chen L, Liu L, Wang H, and Zhou W. 2017. Upregulate
479 KIF4A Enhances Proliferation, Invasion of Hepatocellular Carcinoma and Indicates poor
480 prognosis Across Human Cancer Types. Sci Rep 7:4148. 10.1038/s41598-017-04176-9
481 Huang Y, Wang H, Lian Y, Wu X, Zhou L, Wang J, Deng M, and Huang Y. 2018. Upregulation
482 of kinesin family member 4A enhanced cell proliferation via activation of Akt signaling
483 and predicted a poor prognosis in hepatocellular carcinoma. Cell Death Dis 9:141.

10.1038/s41419-017-0114-4
IO N, J N, EC L, ST F, and M N. 1995. Ki-67 antigen expression in hepatocellular carcinoma using monoclonal antibody MIB1. A comparison with proliferating cell nuclear antigen. American journal of clinical pathology 104:313-318.
IR K, O F, G H, E vM, and MV K. 2017. What is precision medicine? Eur Respir J 50:undefined.
Isohookana J, Haapasaari KM, Soini Y, Leppanen J, and Karihtala P. 2018. Proteins of the retinoblastoma pathway, FEN1 and MGMT are novel potential prognostic biomarkers in pancreatic adenocarcinoma. Pathol Res Pract 214:840-847. 10.1016/j.prp.2018.04.016
J Z, and CB T. 2019. Metabolic regulation of cell growth and proliferation. Nat Rev Mol Cell Biol undefined:undefined.
K S, M K, S T, T M, T M, and K S. 1995. A long-term survivor of ruptured hepatocellular carcinoma after hepatic resection. Journal of gastroenterology and hepatology 10:351-354.
Kuo TC, Chang PY, Huang SF, Chou CK, and Chao CC. 2012. Knockdown of HURP inhibits the proliferation of hepacellular carcinoma cells via downregulation of gankyrin and accumulation of p53. Biochem Pharmacol 83:758-768. 10.1016/j.bcp.2011.12.034
Lam JS, Seligson DB, Yu H, Li A, Eeva M, Pantuck AJ, Zeng G, Horvath S, and Belldegrun AS. 2006. Flap endonuclease 1 is overexpressed in prostate cancer and is associated with a high Gleason score. BJU Int 98:445-451. 10.1111/j.1464-410X.2006.06224.x
Liao W, Liu W, Yuan Q, Liu X, Ou Y, He S, Yuan S, Qin L, Chen Q, Nong K, Mei M, and Huang J. 2013. Silencing of DLGAP5 by siRNA significantly inhibits the proliferation and invasion of hepatocellular carcinoma cells. PLoS One 8:e80789. 10.1371/journal.pone.0080789
Lieber MR. 1997. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. Bioessays 19:233-240. 10.1002/bies.950190309
Liu L, Chen X, Xie S, Zhang C, Qiu Z, and Zhu F. 2012a. Variant 1 of KIAA0101, overexpressed in hepatocellular carcinoma, prevents doxorubicin-induced apoptosis by inhibiting p53 activation. Hepatology 56:1760-1769. 10.1002/hep.25834
Liu L, Zhou C, Zhou L, Peng L, Li D, Zhang X, Zhou M, Kuang P, Yuan Q, Song X, and Yang M. 2012b. Functional FEN1 genetic variants contribute to risk of hepatocellular carcinoma, esophageal cancer, gastric cancer and colorectal cancer. Carcinogenesis 33:119-123. 10.1093/carcin/bgr250
Liu Z, Li J, Chen J, Shan Q, Dai H, Xie H, Zhou L, Xu X, and Zheng S. 2018. MCM family in HCC: MCM6 indicates adverse tumor features and poor outcomes and promotes S/G2 cell cycle progression. BMC Cancer 18:200. 10.1186/s12885-018-4056-8
M A, N K, N I, A H, K H, O M, and M Y. 2009. The knockdown of endogenous replication factor C4 decreases the growth and enhances the chemosensitivity of hepatocellular carcinoma cells. Liver international : official journal of the International Association for the Study of the Liver 29:55-62.
M L, and J Z-R. 2016. Mechanisms of HBV-induced hepatocellular carcinoma. Journal of hepatology 64:S84-S101.
MJ D, N OD, DJ B, WM G, and BM R. 2007. Survivin: a promising tumor biomarker. Cancer letters 249:49-60.
MP K, Y Z, and G L. 2015. Mutant p53: Multiple Mechanisms Define Biologic Activity in Cancer. Frontiers in oncology 5:249.

Nikolova T, Christmann M, and Kaina B. 2009. FEN1 is overexpressed in testis, lung and brain tumors. Anticancer Res 29:2453-2459.

P H, Y W, A F, X L, V M, T O, YJ C, KE S, JM S, H Y, RL B, S G, H Y, KJ K, J Y, C J, Y L, D F, R S, K N, E S, L A, JJ Z, K P, DG A, C L, and P S. 2017. Cell-Cycle-Targeting MicroRNAs as Therapeutic Tools against Refractory Cancers. Cancer cell 31:576-590.e578.

Pan HW, Chou HY, Liu SH, Peng SY, Liu CL, and Hsu HC. 2006. Role of L2DTL, cell cycle-regulated nuclear and centrosome protein, in aggressive hepatocellular carcinoma. Cell Cycle 5:2676-2687. 10.4161/cc.5.22.3500

PN R. 2012. Deep phenotyping for precision medicine. Human mutation 33:777-780.

R N, H J, CJ T, and F C. 2019. Review: Precision medicine and driver mutations: Computational methods, functional assays and conformational principles for interpreting cancer drivers. PLoS computational biology 15:e1006658.

S R, EL L, A B, Y C, X Z, J J, R W, HL J, QH Y, LX Q, ZY T, P H, KW H, SS T, PS M, and XW W. 2012. Integrative genomic identification of genes on 8p associated with hepatocellular carcinoma progression and patient survival. Gastroenterology 142:957-966.e912.

S R, HL J, A B, M F, QH Y, JS L, SS T, Z S, ZY T, LX Q, and XW W. 2010. A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. Cancer research 70:10202-10212.

S S, J K, Y Q, X Y, W W, and L Y. 2019. Identification of core genes and outcomes in hepatocellular carcinoma by bioinformatics analysis. Journal of cellular biochemistry 120:10069-10081.

T M, K T, E A, T M, K S, H H, K S, and M T. 1995. Small hepatocellular carcinoma of single nodular type: a specific reference to its surrounding cancerous area undetected radiologically and macroscopically. Journal of surgical oncology 60:75-79.

T S, Y I, K T, T S, T M, K T, Y H, M N, T I, T F, M O, K E, T U, Y N, S O, M O, T K, R I, Y T, R O, D Y, K T, T F, J P, MS L, T U, T O, T M, T W, H K, T O, T N, H M, S W, T S, M Y, K T, and M K. 2016. p62/Sqstm1 promotes malignancy of HCV-positive hepatocellular carcinoma through Nrf2-dependent metabolic reprogramming. Nature communications 7:12030.

Tsou AP, Yang CW, Huang CY, Yu RC, Lee YC, Chang CW, Chen BR, Chung YF, Fann MJ, Chi CW, Chiu JH, and Chou CK. 2003. Identification of a novel cell cycle regulated gene, HURP, overexpressed in human hepatocellular carcinoma. Oncogene 22:298-307. 10.1038/sj.onc.1206129

W C, W D, J W, Z W, and X H. 2019. Elevated Expressions of Survivin and Endoglin in Patients with Hepatic Carcinoma. Cancer biotherapy & radiopharmaceuticals 34:7-12.

Wang K, Xie C, and Chen D. 2014. Flap endonuclease 1 is a promising candidate biomarker in gastric cancer and is involved in cell proliferation and apoptosis. Int J Mol Med 33:1268-1274. 10.3892/ijmm.2014.1682

WG Z, CF L, M L, XF C, K S, X K, L L, and ZC M. 2016. Aquaporin 9 is down-regulated in hepatocellular carcinoma and its over-expression suppresses hepatoma cell invasion through inhibiting epithelial-to-mesenchymal transition. Cancer letters 378:111-119.

Y L, F R, Y L, Z S, Z T, H X, Y D, and G C. 2015. Clinicopathological and prognostic significance of high Ki-67 labeling index in hepatocellular carcinoma patients: a meta-analysis. International journal of clinical and experimental medicine 8:10235-10247.

YH W, TY C, TY C, KM C, VP C, and KJ K. 2014. Plasmalemmal Vesicle Associated Protein (PLVAP) as a therapeutic target for treatment of hepatocellular carcinoma. BMC Cancer 14:815.

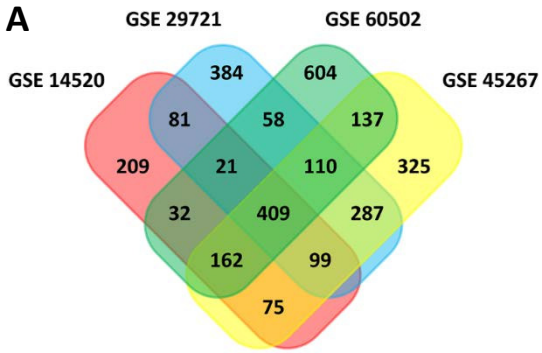
Yuan RH, Jeng YM, Pan HW, Hu FC, Lai PL, Lee PH, and Hsu HC. 2007. Overexpression of KIAA0101 predicts high stage, early tumor recurrence, and poor prognosis of hepatocellular carcinoma. Clin Cancer Res 13:5368-5376. 10.1158/1078-0432.ccr-07-1113

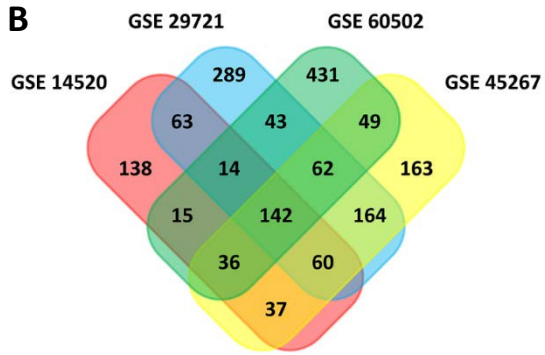
Zhang K, Keymeulen S, Nelson R, Tong TR, Yuan YC, Yun X, Liu Z, Lopez J, Raz DJ, and Kim JY. 2018. Overexpression of Flap Endonuclease 1 Correlates with Enhanced Proliferation and Poor Prognosis of Non-Small-Cell Lung Cancer. Am J Pathol 188:242-251. 10.1016/j.ajpath.2017.09.011

Figure 1(on next page)

Venn diagram.

(A) 409 DEGs were identified in 4 datasets (GSE14520, GSE29721, GSE45267, and GSE60502) via FUNRICH. These included 142 upregulated genes. (B) and 267 downregulated genes. (C) Colors correspond to specific datasets, with intersecting areas indicating overlapping gene sets. DEG identification criteria were adj. $P < 0.05$ and $|\log FC \text{ (fold change)}| \geq 1$.


A
GSE 29721
GSE 60502
GSE 14520
GSE 45267
384
604
81
58
137
209
21
110
325
32
409
287
162
99
75


B
GSE 29721
GSE 60502
GSE 14520
GSE 45267
289
431
63
43
49
138
14
62
163
15
142
164
36
60
37

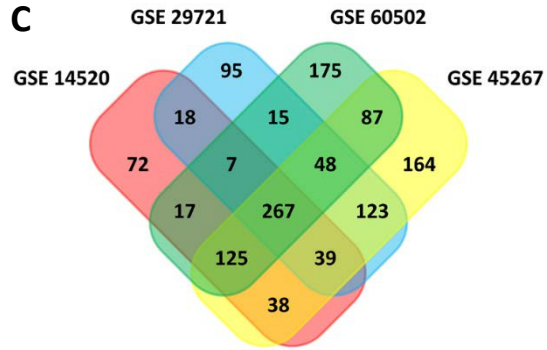

C
GSE 29721
GSE 60502
GSE 14520
GSE 45267
95
175
18
15
87
72
7
48
164
17
267
123
125
39
38

Figure 2(on next page)

DEG PPI network and modular analysis.

(A) STRING was employed to create a PPI network of 403 nodes and 3502 edges, visualized using Cytoscape software. Genes that are upregulated are shown by red nodes, while those that are downregulated are blue. The MCODE plug-in was used to analyze highlighted regions. (B) Clustering module 1 scored 58.492 with 62 nodes and 1784 edges. (C) Clustering module 2 scored 11.529 with 18 nodes and 98 edges. (D) Clustering module 3 scored 10.917 with 25 nodes and 131 edges.

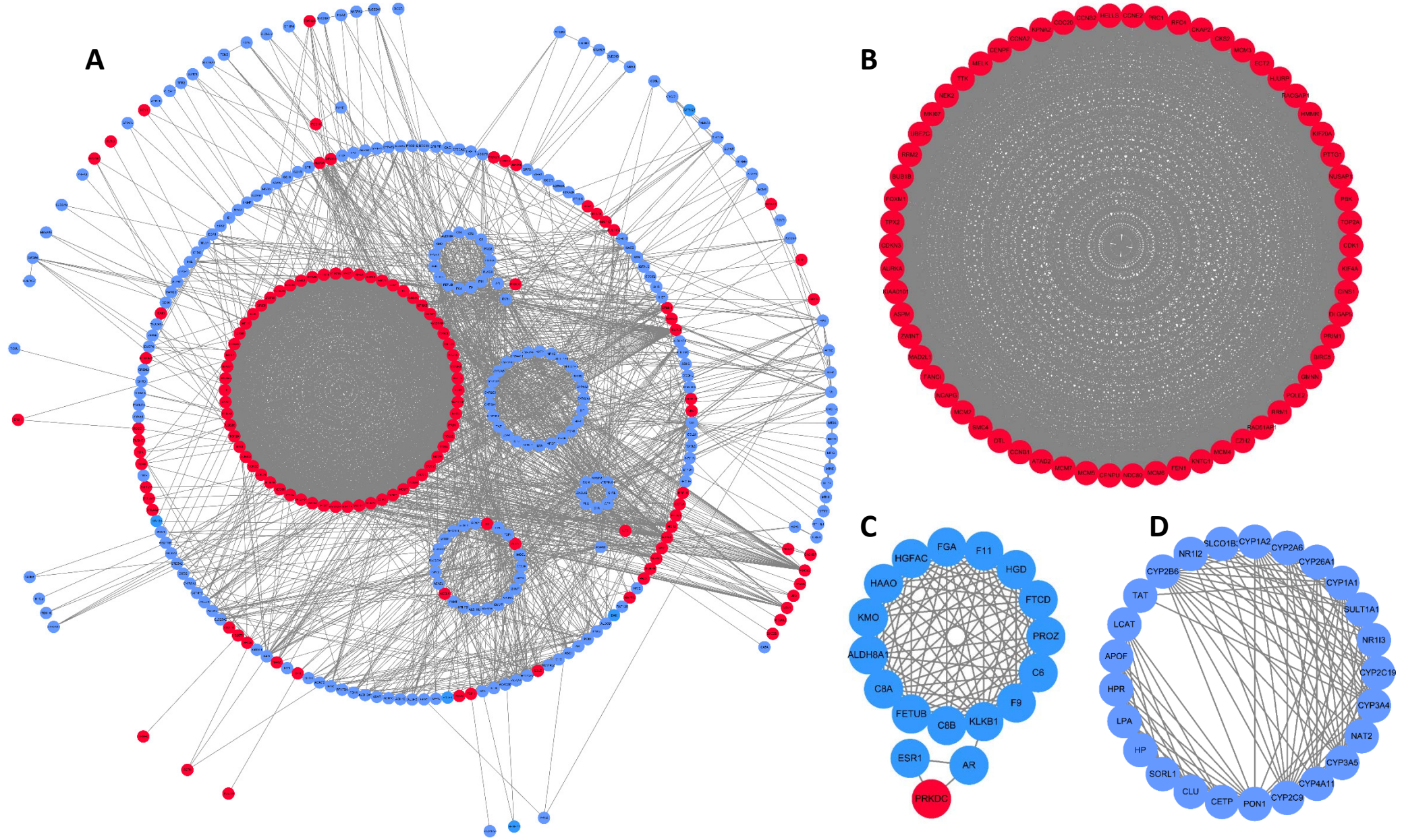

A
B
C
D
FGA
F11
HGFAC
HGD
HAAO
FTCD
KMO
PROZ
ALDH8A1
C6
C8A
F9
FETUB
C8B
KLKB1
ESR1
AR
PRKDC
SLCO1B1
CYP1A2
CYP2A6
NR1I2
CYP26A1
CYP2B6
CYP1A1
TAT
SULT1A1
LCAT
NR1I3
APOF
CYP2C19
HPR
CYP3A4
LPA
NAT2
HP
CYP3A5
SORL1
CYP4A11
CLU
CETP
PON1
CYP2C9

Figure 3(on next page)

Identification of nine core HCC-associated genes.

(A) Nine core genes in clustering module 1 screened by 12 algorithms using the cytoHubba plug-in intersected with Hub genes. (B) Core gene expression profiles in 421 TCGA liver cancer tissues, including 50 solid normal tissues and 371 primary tumors, **P < 0.01. (C) The UCSC databased was used for core gene hierarchical clustering.

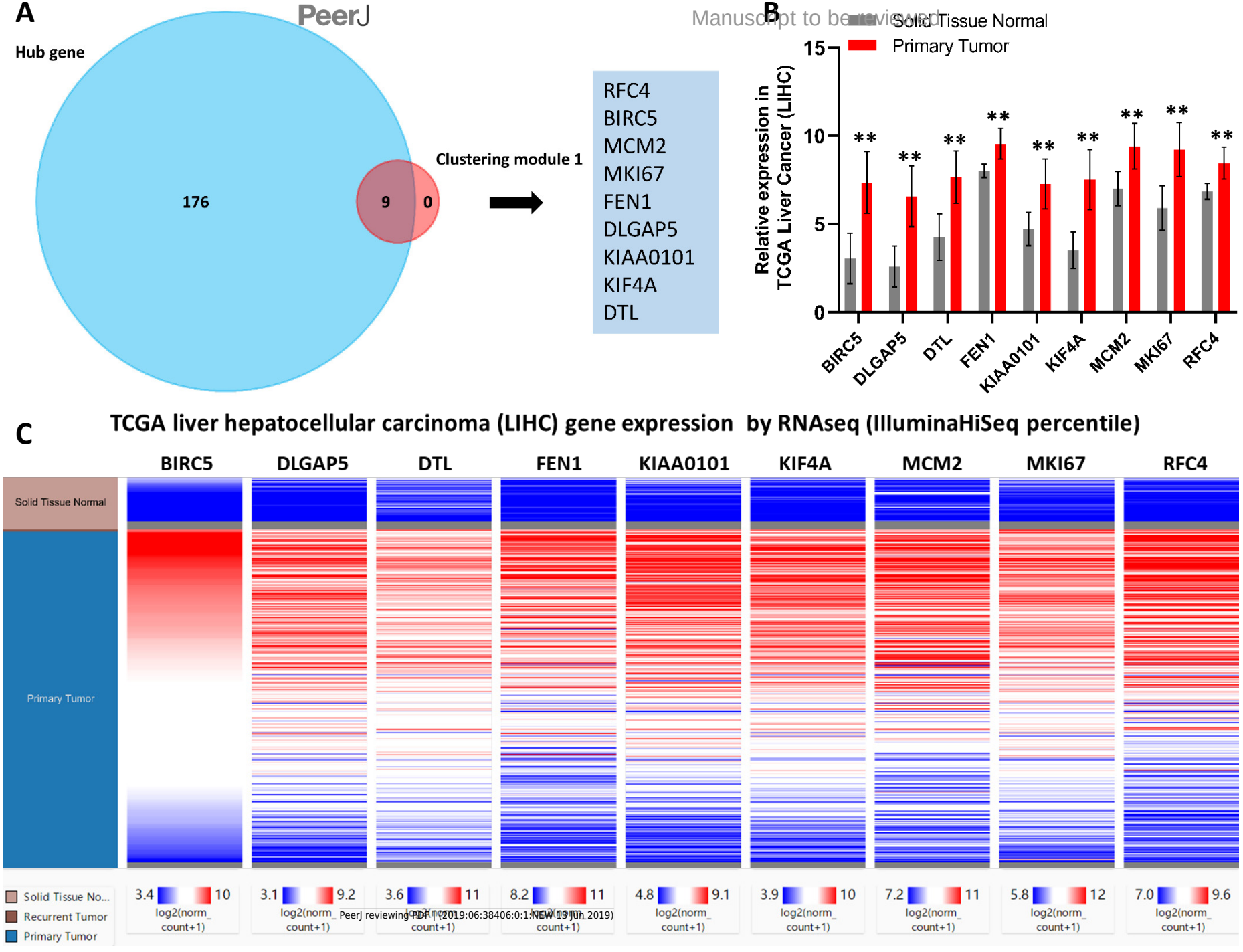


Figure 4(on next page)

Association between core genes and overall survival.

Association between core genes and overall survival (A, BIRC5; B, DLGAP5; C, DTL; D, FEN1; E, KIAA0101; F, KIF4A; G, MCM2; H, MKI67 and I, RFC4) in those with HCC. CI, confidence interval; HR, hazard ratio. High- and low-risk groups are shown in red and blue, respectively. $P < 0.05$ was the significance threshold.

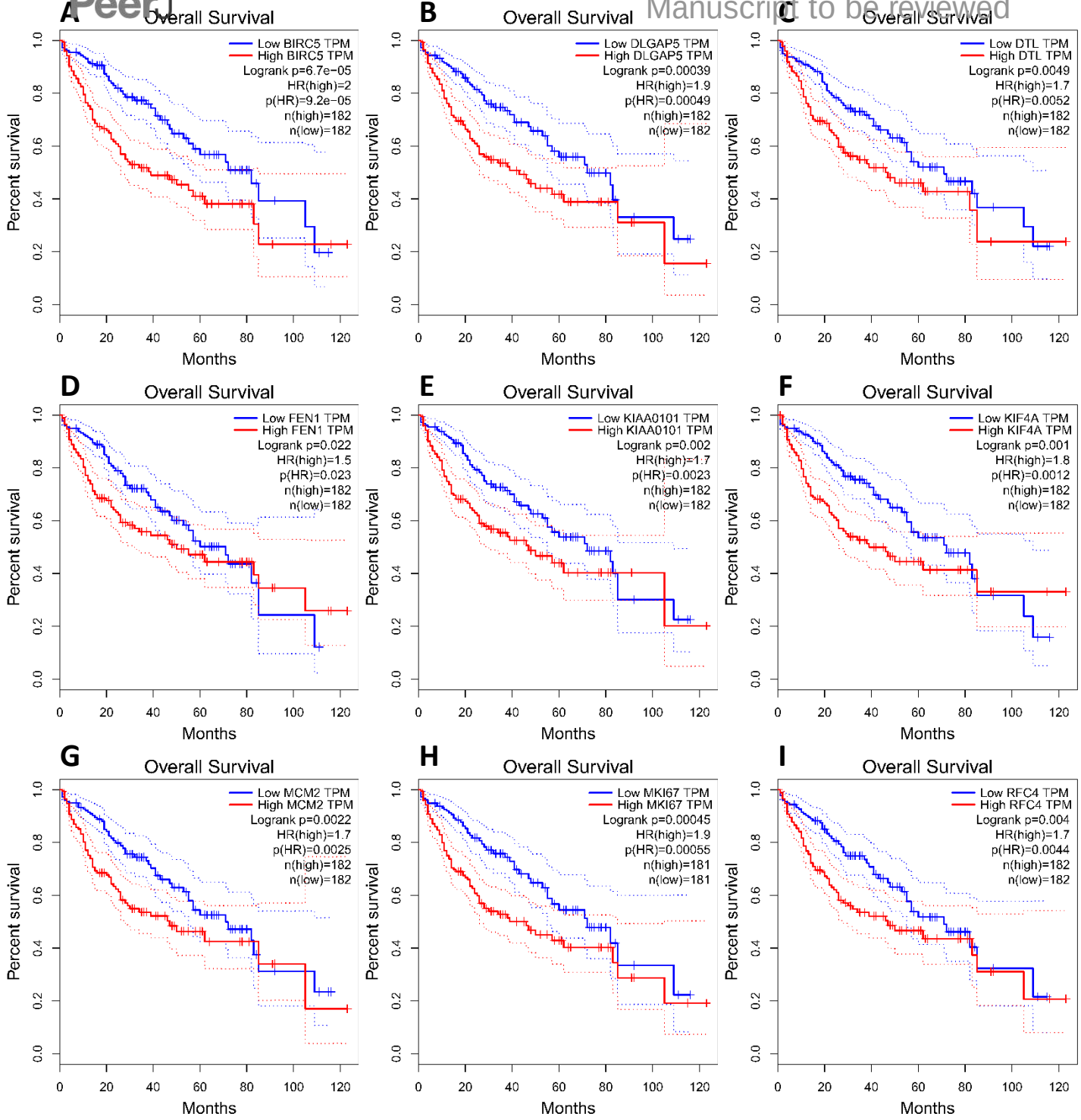


Figure 5(on next page)

Association between core genes and disease-free survival.

Association between core genes and disease-free survival (A, BIRC5; B, DLGAP5; C, DTL; D, FEN1; E, KIAA0101; F, KIF4A; G, MCM2; H, MKI67 and I, RFC4) in those with HCC. CI, confidence interval; HR, hazard ratio. High- and low-risk groups are shown in red and blue, respectively. $P < 0.05$ was the significance threshold.

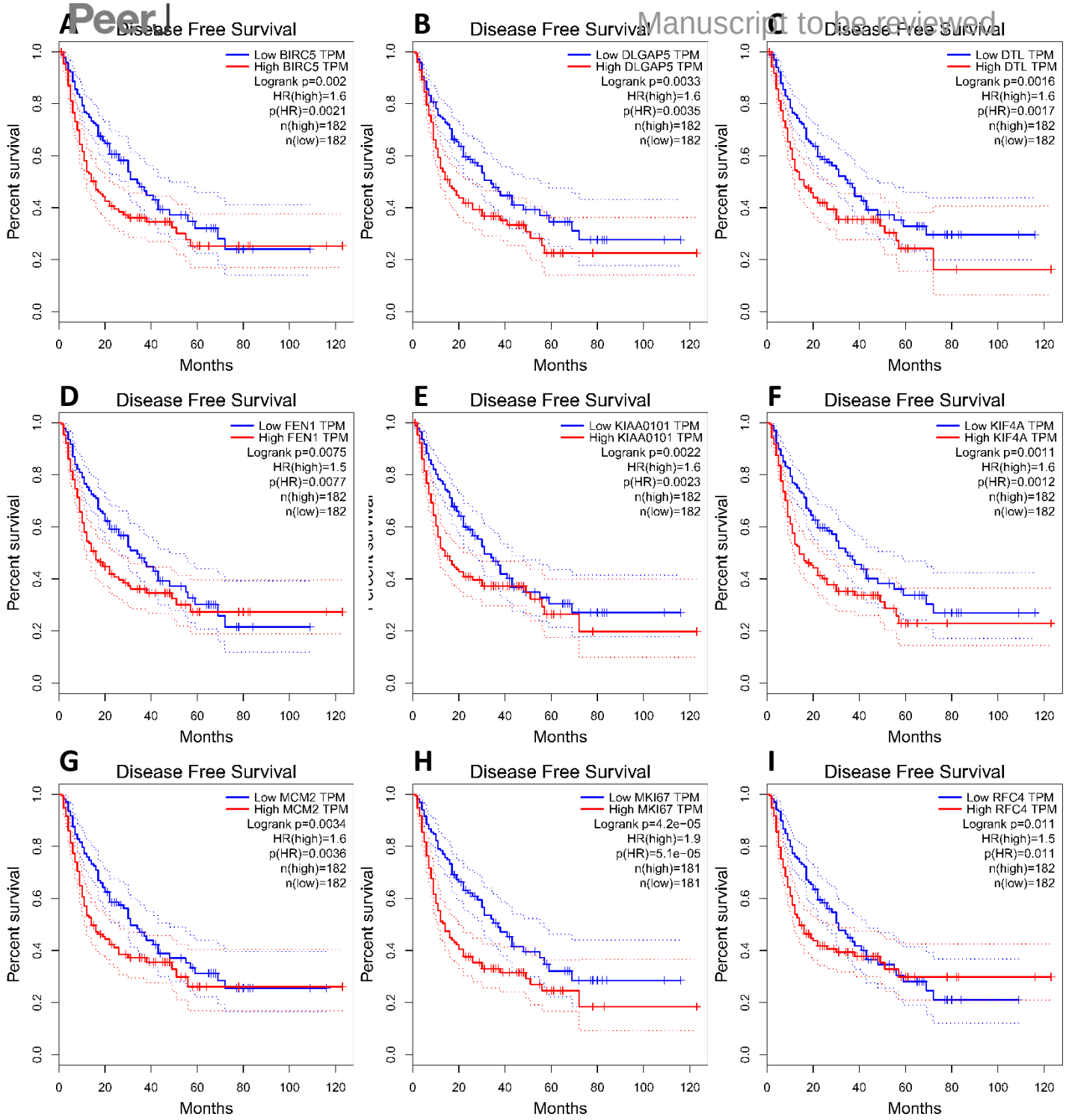
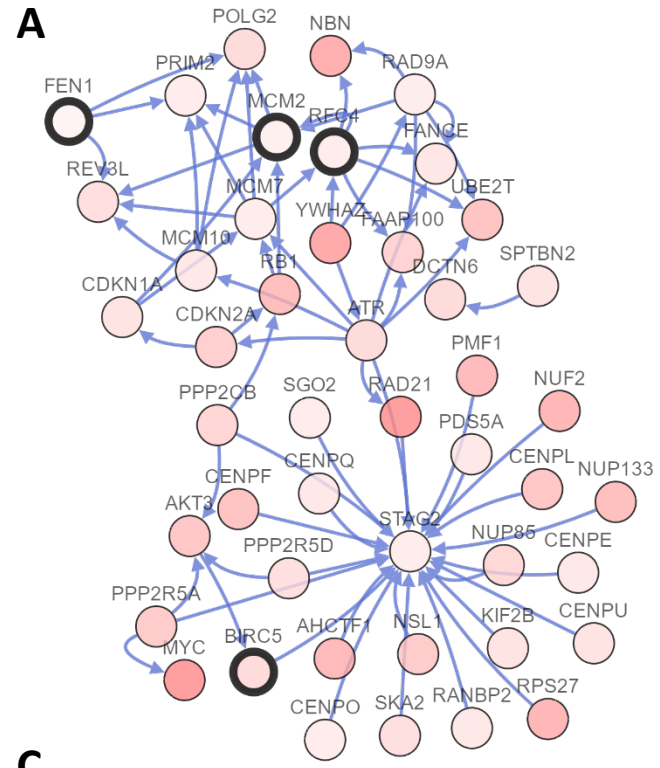
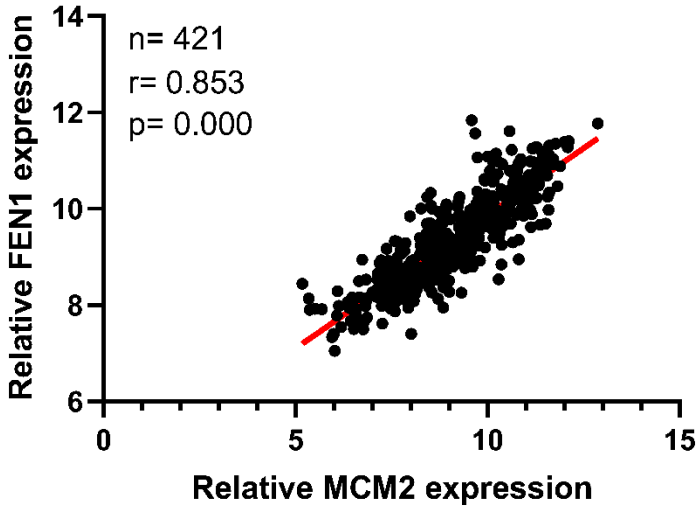


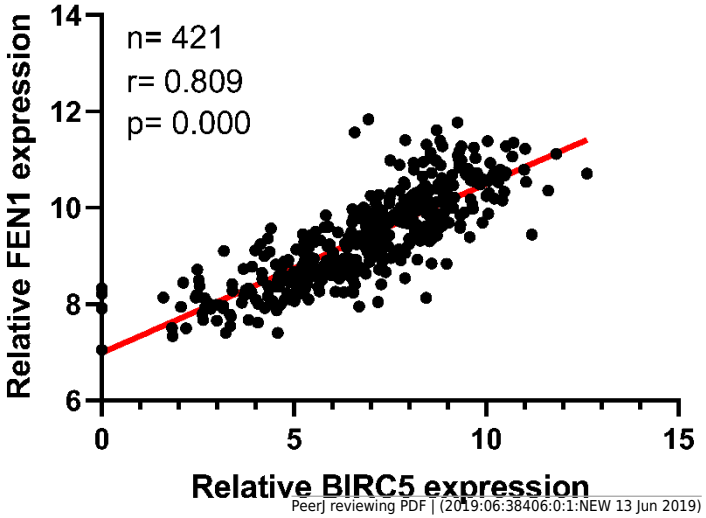
Figure 6(on next page)

Correlations between the expression of FEN1 and MCM2, BIRC5, and RFC4.

(A) A PPI network for the nine core genes generated with the Cbioportal database; (B) The correlation between FEN1 and MCM2 in 421 TCGA liver cancer tissues including 50 solid normal tissues and 371 primary tumors; (C) The correlation between FEN1 and BIRC5 in 421 TCGA liver cancer tissues including 50 solid normal tissues and 371 primary tumors; (D) The correlation between FEN1 and RFC4 in 421 TCGA liver cancer tissues including 50 solid normal tissues and 371 primary tumors. $P < 0.05$ was the significance threshold.


A
POLG2
NBN
PRIM2
RAD9A
FEN1
MCM2
RFC4
FANCE
REV3L
MCM7
UBE2T
YWHAZ
FAAP100
MCM10
RB1
DCTN6
SPTBN2
CDKN1A
CDKN2A
ATR
PMF1
PPP2CB
SGO2
RAD21
NUF2
PDS5A
CENPQ
CENPL
NUP133
CENPF
AKT3
STAG2
NUP85
CENPE
PPP2R5D
PPP2R5A
KIF2B
CENPU
BIRC5
AHCTF1
NSL1
MYC
CENPO
SKA2
RANBP2
RPS27


B
n= 421
r= 0.853
p= 0.000
Relative FEN1 expression
Relative MCM2 expression
14
12
10
8
6
0
5
10
15


C
n= 421
r= 0.809
p= 0.000
Relative FEN1 expression
Relative BIRC5 expression
14
12
10
8
6
0
5
10
15

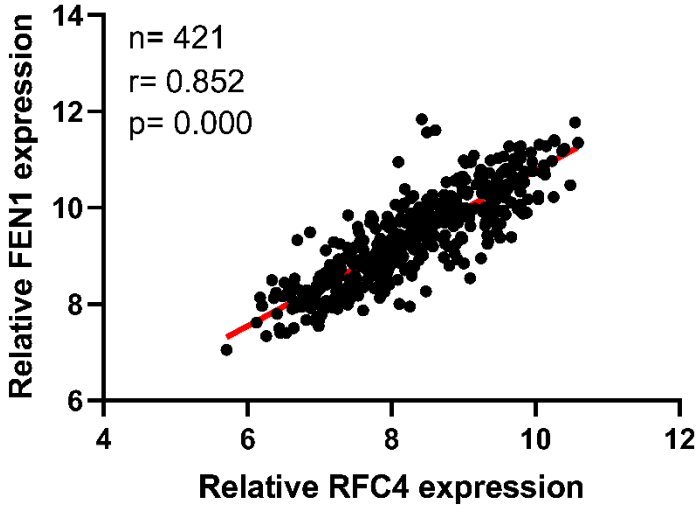

D
n= 421
r= 0.852
p= 0.000
Relative FEN1 expression
Relative RFC4 expression
14
12
10
8
6
4
6
8
10
12

Figure 7(on next page)

The expression of FEN1 in cancers including HCC.

The expression of FEN1 in various tumor tissue types was analyzed in the ONCOMINE (A) and FIREBROWSE databases (B); (C) The expression of FEN1 in three different HCC-related chip data sets was analyzed in the ONCOMINE database.

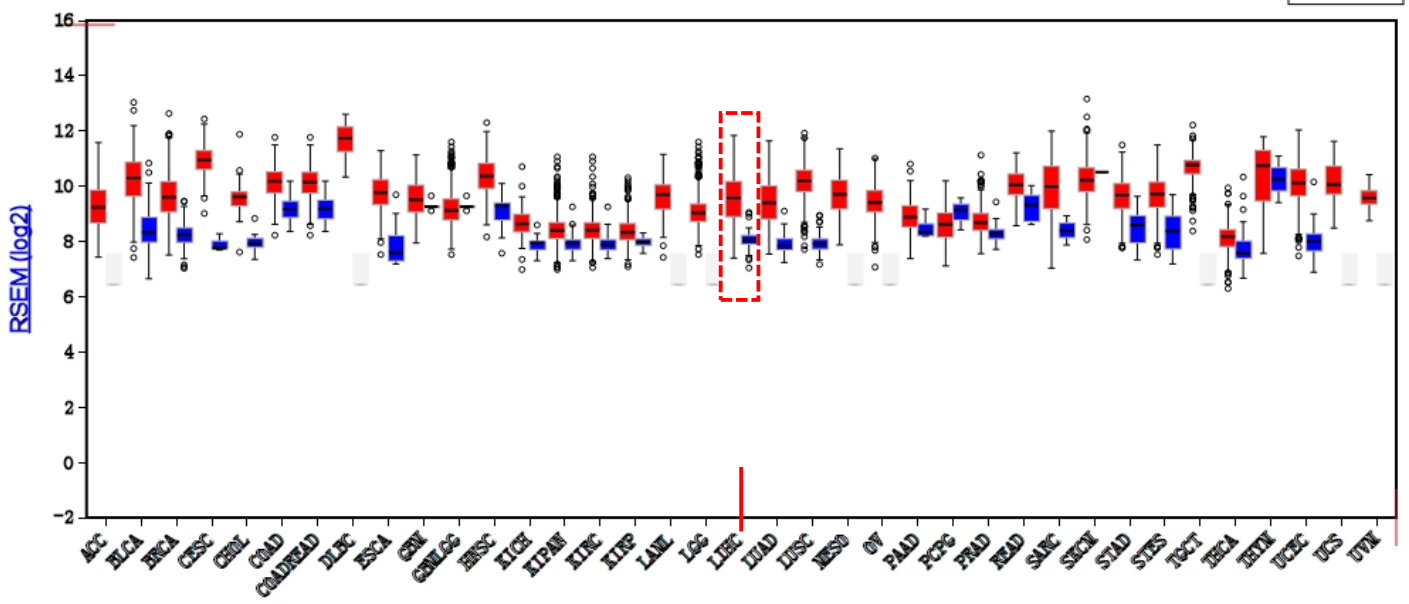
A Disease Summary for FEN1

Analysis Type by Cancer	Cancer vs. Normal	
Bladder Cancer	4	
Brain and CNS Cancer		1
Breast Cancer	8	1
Cervical Cancer	3	
Colorectal Cancer	7	
Esophageal Cancer	2	
Gastric Cancer	1	
Head and Neck Cancer	1	
Kidney Cancer	2	
Leukemia		1
Liver Cancer	3	
Lung Cancer	10	
Lymphoma	5	
Melanoma		
Myeloma	1	
Other Cancer	3	
Ovarian Cancer	1	
Pancreatic Cancer	1	
Prostate Cancer		
Sarcoma	7	
Significant Unique Analyses	58	3
Total Unique Analyses	384	

Cell color is determined by the best gene rank percentile for the analyses within the cell.

NOTE: An analysis may be counted in more than one cancer type.

FEN1 differential plot



C

Comparison of FEN1 Across 3 Analyses

Over-expression

Median Rank	p-Value	Gene	1	2	3
108.0	2.02E-18	FEN1			

Legend

1. Hepatocellular Carcinoma vs. Normal
   Chen Liver, Mol Biol Cell, 2002
2. Hepatocellular Carcinoma vs. Normal
   Roessler Liver, Cancer Res, 2010
3. Hepatocellular Carcinoma vs. Normal
   Roessler Liver 2, Cancer Res, 2010

Not measured

The rank for a gene is the median rank for that gene across each of the analyses.

The p-Value for a gene is its p-Value for the median-ranked analysis.

Figure 8(on next page)

FEN1 was up-regulated in HCC tissues and hepatoma cell lines.

H&E staining of adjacent tissue (A) and HCC tissue (C). (B) FEN1 IHC in adjacent tissue. (D) FEN1 IHC in HCC tissue. (E) FEN1 IHC staining quantification; n=34/group. (F) FEN1 IHC scores; n=34/group. (G) FEN1 expression in control HL7702 liver cells and in hepatoma cell lines (as indicated); *P < 0.05, **P < 0.01.

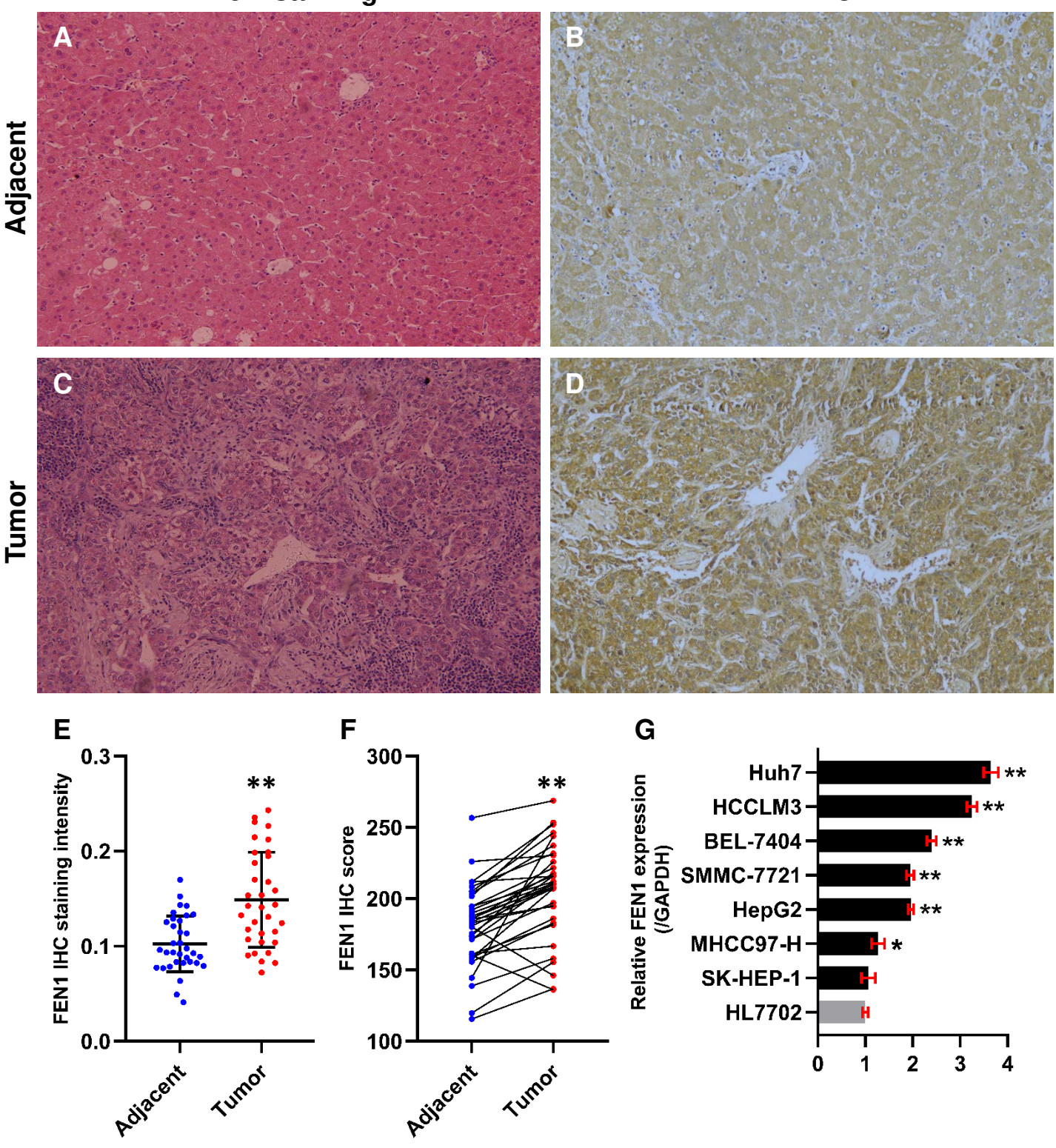

H&E staining
FEN1 IHC
Adjacent
Tumor
A
B
C
D
E
FEN1 IHC staining intensity
0.3
0.2
0.1
0.0
**
Adjacent
Tumor
F
FEN1 IHC score
300
250
200
150
100
**
Adjacent
Tumor
G
Relative FEN1 expression (/GAPDH)
Huh7
HCCLM3
BEL-7404
SMMC-7721
HepG2
MHCC97-H
SK-HEP-1
HL7702
**
**
**
**
**
*
0
1
2
3
4

Table 1(on next page)

Detailed information of the GEO datasets in this study.

1 **Table.1** Detailed information of the GEO datasets in this study.

Series accession	Species	Type	Platform
GSE14520	Homo sapiens	Expression profiling by array	GPL3921 Affymetrix HT Human Genome U133A Array
GSE29721	Homo sapiens	Expression profiling by array	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
GSE45267	Homo sapiens	Expression profiling by array	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
GSE60502	Homo sapiens	Expression profiling by array	GPL96 [HG-U133A] Affymetrix Human Genome U133A Array

2 GEO, Gene Expression Omnibus.

3

Table 2(on next page)

Correlation between FEN1 expression and clinicopathological features in 34 paired HCC patients.

Table 2 Correlation between FEN1 expression and clinicopathological features in 34 paired HCC patients

Clinicopathological features	Cases (n=34)	FEN1 expression		*P* value
		High (%)	Low (%)	
All case (n=34)	34	25	9	
Gender				1.000
Male	29	21(61.8%)	8(23.5%)	
Female	5	4(11.8%)	1(2.9%)	
Age(y)[a]				0.697
<52	15	12(35.3%)	3(8.8%)	
≥52	19	13(38.2%)	6(17.6%)	
Tumor size				0.047
<5cm	19	11(32.4%)	8(23.5%)	
≥5cm	15	14(41.2%)	1(2.9%)	
Tumor multiplicity				0.348
Single	27	21(61.8%)	6(17.6%)	
Multiple	7	4(11.8%)	3(8.8%)	
TNM stage				1.000
I ~ II	23	17(50.0%)	6(17.6%)	
III ~ IV	11	8(23.5%)	3(8.8%)	
Pathological grade				1.000
Well/moderate	27	20(58.8%)	7(20.6%)	
Poor	7	5(14.7%)	2(5.9%)	
Metastasis				0.013
With	12	12(35.3%)	0(0.0%)	
Without	22	13(38.2%)	9(26.5%)	
HBsAg				1.000
Positive	30	22(64.7%)	8(23.5%)	
Negative	4	3(8.8%)	1(2.9%)	
Liver cirrhosis				0.687
With	12	8(23.5%)	4(11.8%)	
Without	22	17(50.0%)	5(14.7%)	

a: patients were divided according to the median age.