A Nomogram For Predicting HCC Patients' Overall Survival Based On Double Hub Genes and Other Clinical Risk Factors.

Jianfei Chen
The Affiliated Hospital of Southwest Medical University https://orcid.org/0000-0003-2370-4222

Zhong-liang Liu
The Affiliated Hospital of Southwest Medical University

Song Su
The Affiliated Hospital of Southwest Medical University

Jun Fan
The Affiliated Hospital of Southwest Medical University

Shun-de Tan
The Affiliated Hospital of Southwest Medical University

Bo Li
The Affiliated Hospital of Southwest Medical University

Xiao-li Yang (344920646@qq.com)
The Affiliated Hospital of Southwest Medical University https://orcid.org/0000-0001-9071-0097

Research

Keywords: hepatocellular carcinoma, bioinformatics, cytoskeleton associated protein 2, forkhead box M1, score system, nomogram

Posted Date: November 9th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1035865/v1

Abstract

Background: Hepatocellular carcinoma (HCC) is one of the series malignant aggressive disease which shows elusive biological behavior and terrible prognosis. It is inadequate for the single evaluation system such as tumor-node-metastasis (TNM) staging system to predict the overall survival (OS) in HCC patients. Here, we conducted this study to identify prognosis-related genes, so as to form a reliable prognostic assessment model.

Results: From three datasets in the GEO database, we identified two hub genes cytoskeleton associated protein 2 (CKAP2) and forkhead box M1 (FOXM1) among 348 differentially expressed genes (DGEs) that were differentially expressed between HCC and normal samples. The function analysis of those DEGs are enriched in cell division process (e.g., mitotic nuclear division, nuclear chromosome segregation) and metabolic process (e.g., organic acid catabolic process). Then, we established a two-gene model, that extremely distinguished the population at risk of liver cancer to high-risk and low-risk and was even viable in the TNM stage i-ii and iii-iv, vascular invasion and non-invasion subgroups (all $P<0.05$). Next, a nomogram was set out combined with the two hub genes and clinical risk factors, and the predictive power of the nomogram performed more outstanding than the gene expression or clinical parameters alone.

Conclusions: Our two-gene-based evaluation system effectively filtered out the high-risk HCC patients, and could potentially be used for clinical decision-making and individualized management of particular HCC patients.

1. Background

Hepatocellular carcinoma (HCC), one of the most common gastrointestinal malignant cancers, holds high morbidity and mortality, which has been reviewed as a notable healthcare issue for human beings in the whole world for many years(1). Although a variety of therapies including resection, chemotherapy, immunotherapy and transcatheter arterial chemoembolization have been widely performed clinically, the prognosis of those methods is still not satisfactory. The most common reason of the poor prognosis is that the patients have lost the opportunity for radical resection because of the advanced stage at the moment of the diagnosis. Thus, it's imperative for HCC populations to execute the early assessment of prognosis. In the past decades, numerous genes and signaling pathways participating in the initiation and evolution of HCC have been extensively discovered, but lacking of effective biomarkers for early detection and prediction of prognosis still troubles clinicians. Recently, with the rapid advancement of bioinformatics technology, various public databases such as the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) could be served as Early detection powerful tools to filter differentially expressed genes (DEGs) in the carcinogenesis and progression of HCC for the genetic research(2, 3). These DEGs might become crucial targets in the diagnosis, treatment and prognosis of HCC, which can provide other researchers more predictable and accurate direction for further study in the laboratory.

In our study, we downloaded 3 microarray datasets [GSE46408, GSE62232, GSE74656] from GEO database, calculated the differential expression multiples between HCC and non-cancerous liver tissues of each gene with GEO2R online tools, then obtained the DEGs in those datasets. Next, GO and KEGG enrichment analysis were conducted to display an overview of the function of those screened DEGs. With the STRING online website and the Cytoscape software, protein-protein interaction (PPI) network was established for further analysis and filtering of the crucial genes. Last, combining some clinicopathological factors in the TCGA database, we established a two-genes clinical-related signature for assessing patients' risk. Based on the two-genes clinical-related signature and other clinicopathological factors (i.e, TNM stage and vascular invasion), a two-gene-based nomogram to evaluate the overall survival of patients with HCC was created.

2. Materials And Methods

The overall ideal and design route of this article was presented in the Figure 1.

2.1. The sources of microarray data and clinical information

GEO (http://www.ncbi.nlm.nih.gov/geo) is a public database, containing high throughout gene expression data, chips and microarrays of Hepatocellular Carcinomas (HCC). We chose three gene microarray expression datasets [GSE46408, GSE62232, GSE74656] from GEO database, respectively including 6 tumor liver tissues and 6 paired nontumourous liver tissues, 81 tumor liver tissues and 10 nontumourous liver tissues, 5 tumor liver tissues and 5 paired nontumourous liver tissues.

469 clinical features and 425 expression profile of HCC patients were downloaded from UCSC Xena(http://xena.ucsc.edu/) and TCGA (https://portal.gdc.cancer.gov/). After filtering out the patients with incomplete data, 363 patients were obtained in this research.

2.2. Data processing of clinicopathological characteristics

Some clinical factors have missing values in the 363 patients, so 14 clinicopathological characteristics (Age, Gender, Family history of cancer, Residual, Histologic grade, TNM stage, Vascular invasion, M stage, N stage, T stage, Hepatitis, Alcohol, Smoking and BMI) were selected as the clinical risk factor after removing those factors whose missing values percentages were more than 15% (Fig S1). Next, we selected 1 as the random number used multiple imputation in mice package, based on 5 replications and a chained equation approach method in the RMI procedure, to account for missing data so as to obtain a new complete dataset (entire dataset) (table 1). The clinical characteristics did not differ significantly between the two datasets (all P>0.05) (Table S1).

2.3. Selection of differentially expressed genes (DEGs)

GEO2R(http://www.ncbi.nlm.nih.gov/geo/geo2r), an online analysis tools in GEO website, was used to identify the DGEs between the tumor tissue and the nontumourous tissue in the three GEO datasets. The adjusted P-values (adj. P) and Benjamini and Hochberg false discovery rate were applied to provide a balance between discovery of statistically significant genes and limitations of false-positives. Probe sets without corresponding gene symbols or genes with more than one probe set were removed or chose the maximum, respectively. We set adj. P≤0.05 and | logFC (fold change) |≥1 as the criterion of statistically significant, then selected the common DEGs conforming the criterion in the three datasets.

2.4. KEGG and GO enrichment analyses of DEGs.

Kyoto Encyclopedia of Genes and Genomes(KEGG) is a database resource for understanding high-level functions and biological systems from large-scale molecular datasets generated by high-throughput experimental technologies. Gene ontology (GO) is a major bioinformatics tool to annotate genes and analyze biological process of these genes.

The selected DEGs can be divided into two groups, up-regulated and down-regulated. To analyze the function of DEGs, biological analyses were performed using Cytoscape(4) plug-in ClueGO(version 2.5.7)(5) and Gluepedia(version 1.5.7)(6). P<0.05 was considered statistically significant.

2.5. PPI network construction and module analysis.

Search Tool for Recurring Instances of Neighbouring Genes (STRING; http://string-db.org) (version 11.0)(7) is an online database implementing experimental and predicted PPI information. All the selected DEGs including up-regulated and down-regulated were input the online tool to establish a complicated protein to protein interaction (PPI) network. Basic setting and advanced setting of STRING database were default. After filtering out the disconnected nodes, the new PPI network was import into the Cytoscape, with Up-regulated DEGs and down-regulated DEGs respectively marked in red and blue.

We use the plug-in MCODE(8) in the Cytoscape to identify the most significant module from this PPI network. The criteria for selection were as follows: degree cut-off=2, node score cut-off=0.2, Max depth=100 and k-score=2. Module with the highest score was considered as the candidate module for the next analysis.

Biological Networks Gene Ontology (BiNGO, version 3.0.4)(9) plug-in can calculate overrepresented GO terms in the selected module and display them as a network of a significant GO terms. The setting of biological process analysis in the candidate module is as follows: Hypergeometric test as the statistical test, Benjamini & Hochberg False Discovery Rate (FDR) correction as the multiple testing correction, overrepresented categories after correction to be visualized, the whole annotation as reference set, P<0.01 as the significance level.

2.6. Identification and analysis of hub genes

The candidate genes of the most significant module from the PPI network were evaluated by the plugin cytoHubba (version 0.1)(10). The plugin cytoHubba in the Cytoscape can predict and explore important genes and subnetworks in a given PPI network with different topological algorithms. In this study, we calculated the scores of all genes in the module by the algorithms of ClusteringCoefficient. Consequently, the top 10 genes with the highest scores of the algorithm was selected as the candidate genes. For the preliminary exploration of the prognostic potential, overall survival (OS) of the 10 candidate genes was performed with the Kaplan–Meier test in 363 HCC patients.

Then, 182 patients in the whole population were randomly sampled and established as the ‘primary dataset’ (Table S2), while the ‘entire dataset’ included the complete 363 patients. Importantly, there were no significant differences in clinical characteristics between the two datasets. These two datasets were used for the subsequent model establishment and verification. During the process of grouping, we set the sample size of primary dataset as 50% in the entire dataset, while 1 was selected as the random number.

To further explore the best-fit clinical-related hub genes, we filtered the 10 candidate genes through the univariate (screening criteria: P < 0.2) and multivariate (screening criteria: P <=0.05) Cox proportional hazards regression (CPHR) analysis in the primary dataset. The hub genes with P <=0.05 in the primary dataset was chosen for assessment of the clinical-related gene signature and the development of OS-related gene signature-based prognostic nomogram.

2.8. Establish of clinical-related gene signature and clinical-related gene signature-based prognostic nomogram

During performing the multivariate CPHR in the above, we also calculated the coefficient of each gene. Therefore, a risk score formula, with the linear combination weighting of the OS-related genes’ expression values and the corresponding regression coefficients, can be established. The score formula is shown as follow:

$$\mathrm{RiskScore} = \sum_{i=1}^{n} (\mathrm{K_i} \times \mathrm{E_i})$$

As shown above, $\mathrm{E_i}$ represents the expression value of each OS-related genes, while the regression coefficient of the above genes is defined as $\mathrm{K_i}$. Next, the corresponding risk score of every patient in the two datasets can be calculated, and cut-off value of the high/low-risk group in 2 datasets was set as the median score calculated from the primary dataset. All datasets were performed Kaplan-Meier method to explore the prognosis differences in the high-risk subgroup and low-risk subgroup. Additionally, to further evaluate the predictive ability of the OS-related genes-based classifier, time-dependent ROC curves of different datasets and correspondent AUC values were conducted with one-year, three-year and five-year.

Moreover, in order to identify independent predictors of OS, we conducted 14 typical clinical characteristics and the OS-related genes signature with univariate (screening criteria: P < 0.2) and multivariate (screening criteria: P <= 0.05) CPHR analyses in the primary dataset. Furthermore, we performed a stratified analysis so as to verify whether the association of the OS-related genes signature with OS was independent of the TNM stage and vascular invasion. Using the selected clinical characteristics, we developed a prognostic nomogram with the ‘rms’ package. The predictive abilities of the established model were assessed with a C-index (non-events vs. events) and calibration curves (nomogram-predicted OS vs. actual probabilities OS) in all datasets. A bootstrap validation with 1000 resamplings was used for these activities. As for evaluating the predictive performance of our nomogram, the time-dependent ROC curves and AUC values of the nomogram based on the OS-related genes and other risk factors were the indicators considered in each dataset.

2.9. Verification of expression level and survival analysis in the public databases

First of all, two databases including GEPIA2 database (http://gepia.cancer-pku.cn/index.html)(11) and UALCAN database(http://ualcan.path.uab.edu/index.html)(12) were conducted for the 2 OS-related genes' expression difference in carcinomatous and normal samples. Meanwhile, in order to verify the expression level of CKAP2 and FOXM1 between tumor and normal tissues, we downloaded 424 sample data from HCC patient in TCGA database and analyze those data with R (version 4.0.2). After removing 5 unqualified samples(formalin-fixed paraffin-embedded tissues and recurring tumor tissues) and TMM normalization of edgeR package(13), a total of 419 samples containing 369 tumor tissues and 50 adjacent tumor tissues were taken to analyses. Then, with the aim of exploring the role of the 2 selected genes during HCC formation and progression, the mRNA expression of hub genes in different stages was compared between HCC and normal tissues by using UALCAN database.

Kaplan-Meier plotter (http://kmplot.com/analysis)(14) is an online survive analysis tool whose sources mainly coming from TCGA database. Overall survival (OS) was performed in our study to evaluate the prognostic potential of CKAP2 and FOXM1. OS analysis was also conducted in GEPIA2 database.

3. Result

3.1. Selection of DEGs in HCC

We comprised 92 Hepatocellular Carcinoma tissues and 21 normal liver tissues in the 3 databases. According to the cut-off criteria of $|\log2\ FC| \geq 1.0$ and adj P-value ≤ 0.05, we identified 2160 DEGs in GSE46408 database, 1444 DEGs in GSE62232 database and 1018 DEGs in GSE74656 database with the GEO2R online analysis tool. 348 overlapping DEGs in the 3 databases containing 194 up-regulated genes and 154 down-regulated genes was shown by a venn diagram (Fig. 2). The results of the expression level analysis of 3 the databases are presented in a volcano plot (Fig. 2).

3.2. GO and KEGG enrichment analysis of DEGs

To explore the biological classification of DEGs, functional and pathway enrichment analyses were performed using Cytoscape plug-in ClueGO. The GO function annotation can be divided into three functional groups, cell component (CC), molecular function (MF), and biological process (BP). 194 up-regulated genes were main enriched in BP (mitotic nuclear division, nuclear chromosome segregation, nuclear chromosome segregation, microtubule cytoskeleton organization involved in mitosis and mitotic sister chromatid segregation) and CC (centromeric region) (Fig. 3A and B). Changes in KEGG pathway analysis of these gene significantly enriched cell cycle, focal adhesion, ECM-receptor interaction and cellular senescence (Fig. 3C and D).

The GO function analysis of 154 down-regulated genes revealed that those biological processes including organic acid catabolic process, alpha-amino acid metabolic process, cellular amino acid catabolic process, cellular response to xenobiotic stimulus, lipid oxidation may play a critical role in the progress of HCC (Fig. 4A and B). While, the changes in KEGG pathway analysis were main enriched in fatty acid degradation, complement and coagulation cascades, chemical carcinogenesis and retinol metabolism (Fig. 4C and D).

3.3. PPI network construction and module analysis.

348 DEGs were input the STRING online website to establish a PPI network with 297 nodes and 2453 edges after removing the disconnected nodes. This network was imported into the Cytosacpe software for the visualization (Fig S2) and the next module analysis. The MCODE plug-in in Cytoscape software was used to grab the most significant module in the network. 12 clusters obtained from the MCODE plug-in were shown in the Table 2. The first cluster with the scores 50.423 was identified as the most significant module (Fig. 5A). The biological processes of the candidate module were carried out using the BiNGO plug-in (Fig. 5B).

3.4. Identification and analysis of hub genes

We implemented the ClusteringCoefficient method of cytoHubba plug-in to evaluate the significance of the genes in the module. Thereafter, 10 top genes (CDC7, CKAP2, MND1, FANCD2, EZH2, DEPDC1, ECT2, FOXM1, UBE2T and CKS2) with the highest scores were considered as the candidate genes. With the Kaplan-Meier analysis in 363 HCC patients, high expression group of every gene performed significantly worse survival event than the low-expression group ($P < 0.05$) (Fig S3).

We filtered these candidate genes through a univariate and multivariate CPHR analysis. As shown in the table (Table 3), CKAP2(HR = 1.2, CI95 = 1.02~1.41, P Value = 0.028) and FOXM1(HR = 0.88, CI95 = 0.79~0.98, P Value = 0.02) were figured out and chosen as the clinical-related genes.

3.5. Establishment and assessment of OS-related genes signature

Based on the above multivariate CPHR analysis, we got CKAP2 and FOXM1's risk-coefficients, 1.21 and 0.98 respectively. Next, we established a risk score formula:

Risk Score = (1.21 × ExpressionCKAP2)+(0.98 × ExpressionFOXM1)

Then, we calculated the two-genes-based risk score for each HCC patient in the primary dataset and entire database. Because of the non-normal distribution of risk score, we set the median risk score as the cut-off value to classify all patients to 'High-score' group and 'Low-score' group. Primary database contains 91

High-score patients and 91 Low-score patients, while 181 High-score patients and 182 Low-score patients in entire database.

In primary dataset, Kaplan-Meier curve analysis clearly demonstrated that the high-risk group (n=91) had a poorer prognosis than the low-risk group(n=91) (P= 0.0062, log-rank test) (Fig. 6A). Subsequently, we constructed a time-dependent receiver operating characteristic (ROC) curve with the primary dataset. As shown in the figure, the area under the time-dependent ROC curve (AUC) of the OS-related genes signature reached 0.68 (95% confidence interval [CI]=59.89–75.97) at one years, 0.64 (95% CI=55.2–73.61) at three years and 0.62 (95% CI=50.09–73.91). (Fig. 6B)

The performance of the OS-related genes signature for predicting survival was then validated in the entire dataset. Similar to the results of the primary dataset, a Kaplan-Meier curve analysis indicated that the survival time of HCC patients was significantly shorter in the high-risk group (n=181) than in the low-risk group (n=182) (P=0.00073, log-rank test) (Fig. 6C). The AUC of the clinical-related gene signature was 0.7 (95% CI=64.3–76.27) at one years, 0.62 (95% CI=54.97–68.99) at three years and 0.57(95% CI=48.12–65.78) in the entire dataset (Fig. 6D). Thus, the predictive performance of the clinical-related gene signature for HCC patients was great in both the primary dataset and the entire dataset.

3.6. The prognostic value of the clinical-related gene signature was independent from those of conventional clinical risk factors

Clinical risk factors such as the TNM stage and age are still vital predictors of OS in HCC patients. Therefore, we integrated these traditional risk factors with our clinical-related gene signature to develop an efficient quantitative method of predicting OS. We first evaluated the prognostic value of several clinical risk factors in univariate and multivariate CPHR analyses of the primary dataset. We found that, in addition to the clinical-related gene signature, TMN stage (iii-vi vs. i-ii) and vascular invasion (yes vs. no) were significantly associated with OS (all P<0.05) (Table 4).

Considering the number of HCC patients, we performed a risk-stratified analysis with the entire dataset. The 363 HCC patients were stratified into a stage i-ii subgroup (n=270), stage iii-iv subgroup (n=93) based on their TNM stage. Each subgroup was divided into a high-risk group and a low-risk group based on the risk scores proposed above. We found that the classification efficiency of the clinical-related gene signature was limited when it was applied to certain subgroups. As shown in the Kaplan-Meier curves, for the two subgroups, patients in the high-risk group had significantly poorer survival than those in the low-risk group (stage i-ii subgroup, P=0.049; stage iii-iv subgroup, P=0.03, log-rank test) (Fig. 7A and B). Similarly, when a stratified analysis was carried out based on vascular invasion, prognosis was worse in the high-risk group both in invasion subgroup(n=133) and non-invasion subgroup(n=230) (Fig. 7C and D).

3.7. Establishment of a nomogram containing the clinical-related gene signature with clinical risk factors

we integrated these traditional risk factors (TNM stage and vascular invasion) with our clinical-related gene signature to develop an efficient quantitative method of predicting OS. Ultimately, on the basis of clinical judgment and statistical significance, we developed a clinical-related gene signature-based nomogram, which integrated the clinical-related gene signature and two clinical risk factors (vascular invasion and TNM stage). We then used this nomogram to predict the one-year, three-year and five-year survival of HCC patients (Fig. 8A).

In the nomogram, the TNM stage contributed the most to the one-, three- and five-year OS, closely followed by the OS-related genes signature and vascular invasion. This user-friendly graphical tool allowed us to determine the one-, three- and five-year OS probability for each HCC patient easily.

We then evaluated the discrimination and calibration abilities of the prognostic nomogram by using a concordance index (C-index) and calibration plots. An internal validation using a bootstrap with 1000 resamplings revealed that the nomogram performed well for discrimination: the C-index was 0.658 (95% CI=0.633-0.683) for the entire dataset and 0.659 (95% CI=0.628-0.69) for the primary dataset. The one-year, three-year and five-year OS probabilities generated by the nomogram were plotted against the observed outcomes, as shown in Fig. 8B−G. The probabilities determined by the nomogram closely approximated the actual probabilities, especially the predicted ability of the three-year OS.

We further assessed the prognostic performance of the nomogram in a time-dependent ROC curve analysis. The AUC of the nomogram was 0.71 (95% CI=61.24-8.-0.15) at one year, 0.75(95% CI=66.05-83.8) at three years and 0.74 (95% CI=62.06-85.08) at five years in the primary dataset (Fig. 9A), while in the entire dataset (Fig. 9B), the AUC was 0.75 (95% CI=67.22-81.25) at one year, 0.7 (95% CI=62.22-76.99) at three years and 0.63 (95% CI=53.72-72.72) at five years. Combining the results of the two time-dependent ROC curves and the calibration plots, we thought the diagnostic power of the nomogram at three-yeas was more excellent than one-year and five-years. More importantly, the discrimination performance of the double-genes-based nomogram (AUC=0.7, 95% CI=62.22-76.99) at three years was superior to the performance of the TNM stage (AUC=0.63, 95% CI=57.04-69.55) and vascular invasion (AUC=0.56, 95% CI=49.29-62.77) (Fig. 9C).

3.8. Verification of expression level and survival analysis in the public databases

We next explore those genes’ expression difference in the public database. Pan-cancer research of the GEPIA database has prompted us the up-regulated of CKAP2 and FOXM1 in the HCC tissue (Fig. 10E). In GEPIA database and UALCAN database, as shown in Fig. 10A and D, the bar chart demonstrated that the expression levels of CKAP2 and FOXM1 in tumor samples are higher than those in the normal tissues. With 419 sample from TCGA database, CKAP2 and FOXM1 were overexpressed in the HCC tissues (Fig.

10C). Furthermore, according to the nomogram, we inferred that the expression level might be associated with disease progression, which was consist with our inference in the UALCAN database (Fig. 10B). All the results have significant statistical differences.

After examining the mRNA expression levels of the 2 genes in HCC, we verified the prognostic potential in the public databases. The survival significance map of Pan-cancer analysis in the GEPIA database preliminary revealed that the overexpression of those 2 genes were significant in the HCC patients (Fig. 11A). Consequently, as shown in Fig. 11B and C, high expression of CKAP2 and FOXM1 both performed poorer prognosis in the OS analysis of GEPIA database and ULACAN database.

4. Discussion

Hepatocellular Carcinomas is one of the most commonly diagnosed cancers and the fourth most common cause of cancer-related death in the world. Although the treatment of hepatocellular carcinomas has improved, the prognosis of patients is generally poor due to the lack of precise molecular targets. Therefore, better biomarkers for specific prognosis and progression of HCC are demanded. In the present experiments, bioinformatics analysis was performed to identify the potential key genes correlated with HCC and develop a reliable prognostic model for HCC patients for helping the clinical decision-making.

In our study, we compared three gene microarray expression datasets from GEO database containing tumor tissues and normal tissues. 194 up-regulated and 154 down-regulated DEGs were successfully identified, respectively enriched in cell division process (e.g., mitotic nuclear division, nuclear chromosome segregation) and metabolic process (e.g., organic acid catabolic process). Based on the degree of connectivity in PPIN, 10 selected genes from ClusteringCoefficient algorithms were identified, including CDC7, CKAP2, MND1, FANCD2, EZH2, DEPDC1, ECT2, FOXM1, UBE2T and CKS2. Through Kaplan-Meier analysis and CPHR analysis, CKAP2 and FOXM1 were considered as the crucial genes in the progress of HCC.

CKAP2 is a cytoskeleton-associated protein which was reported to colocalize with microtubules and centrosomes during interphase and with mitotic spindles during mitosis(15). The function of CKAP2 protein can stabalize microtubules and plays a role in the regulation of cell division. However, Tsuchihara et al disclosed that the overexpression of CKAP2 in the absence of p53 can induce tetraploidy and aberrant centrosome numbers(16). The CKAP2-induced cytokinesis failure may play a crucial role during the process of tumorigenesis and development. Furthermore, the overexpression of CKAP2 has been reported in cervical carcinoma(17), osteosarcoma(18), ovarian cancer(19) and prostate cancer(20). With the hepatocellular carcinomas, Hayashi et al also proved the potential predictive ability of CKAP2 in early and extensive recurrence after operative resection and the expression of CKAP2 may be associated with multiple tumors or microscopic vascular invasion(15). In our nomogram based on CKAP2 and FOXM1, high-expression of genes and vascular invasion were related with the poorer prognosis, whose results supported the present findings.

FOXM1 is one of the members of the Fork head family proteins that contains a 100 amino acid long DNA binding domain, functioning in the regulation of G1/S and G2/M transitions of cell cycle, maintenance of mitotic spindle integrity, angiogenesis, DNA damage repair, and tissue regeneration(21). On the other hand, overexpression of FOXM1 has been detected in a broad range of cancer types such as AML(22), prostate cancer(23), breast cancer(24) and lung carcinoma(25), suggesting that FOXM1 is also essential for tumorigenesis. Similar to the other cancers, up-regulated FOXM1 in hepatocellular carcinomas performed an indicator of poor survival and metastasis(26–28). Wei et al demonstrated that the elevation of P53 and down-regulation of FoxM1 in the sorafenib-treatment tissues(29), so the mechanism of P53 negative regulation might explain the overexpression of FOXM1 in HCC(30, 31).

Based on the 2 hub genes, we established a risk score system which can divided the HCC patients to high-risk subgroup and low-risk subgroup. These high-risk patients exhibited significantly shorter survival than those in the low-risk group. The present studies have indicated that TNM stage and vascular invasion are both simple but useful predicators of survival in HCC. Consequently, we combined these traditional clinical factors with molecular profiling. A two-genes clinical-related nomogram was established to quantify the individual’s probability of OS. The predictive performance of our nomogram was more accurate than those of the traditional TNM stage or vascular invasion alone. And the nomogram's predictive ability was high both in the primary and entire validation, which suggests its high application potential in clinical practice.

Meanwhile, our predicted model in the present study still has some potential limitations: First of all, we excluded participants with incomplete records for complete-case analysis to build the model, which may introduce selection bias and small sample cohorts. However, we used multiple imputations to replace missing values, and the results proved that our study participants after multiple imputations could well represent the overall population. Therefore, in the future, we can consider designing our studies or cooperating with other researchers to collect as many variables as possible to reduce missing values. Secondly, the database of TCGA lacks enough tumor recurrence information, so the accuracy of predicting the prognosis and survival possible in the HCC recurrence patients need further verification. thirdly, our used data came from an open-access published database, so our study design was retrospective. Therefore, prospective clinical studies are needed to validate our findings and to determine whether our nomogram improves patients’ satisfaction and outcomes.

5. Conclusion

In summary, the genes of CKAP2 and FOXM1 were identified as the clinical-related genes in the progression of HCC by bioinformatics analysis. We determined the altered RNA expression patterns of HCC patients and identified a clinical-related genes signature that could efficiently divide patients into different risk groups. Importantly, by combining this signature with conventional clinical risk factors (TNM stage and vascular invasion), we developed a clinical-related genes signature-based nomogram that could accurately predict the one-year, three-year and five-year OS of HCC patients. Furthermore, the prognostic performance of the nomogram was superior to those of the conventional TNM stage or vascular invasion. Therefore, we have provided a reliable, user-friendly prognostic nomogram to aid the individualized management of HCC patients.

Declarations

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and material

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. https://portal.gdc.cancer.gov/; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= [GSE46408, GSE62232 and GSE74656].

Declaration of interest statement

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China under Grant [81802778]; the Science & Technology Department of Sichuan Province under Grant [21ZDYF2161]; the Chen Xiao-Ping Foundation for the Development of Science and Technology of Hubei Province[HR20033]; Chinese Foundation for Hepatitis Prevention and Control[17002].

Authors' contributions

JFC and ZLL designed experiment together. Public data from the GEO and TCGA was downloaded and analyzed by JFC. JFC and ZLL both performed the picture processing and wrote the manuscript. The administrative and technical support were offered from XLY and BL, and they also supervised the study. SS, JF and SDT contributed to the conception and revision of the study. All authors read and approved the final manuscript and agree to be accountable for all aspects of the research.

Acknowledgements

We gratefully acknowledge the Gene Expression Omnibus and The Cancer Genome Atlas.

References

1. Sung H, Ferlay J, Siegel RL, *et al*: Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 71: 209–249, 2021.
2. Liu S, Yao X, Zhang D, *et al*: Analysis of Transcription Factor-Related Regulatory Networks Based on Bioinformatics Analysis and Validation in Hepatocellular Carcinoma. Biomed Res Int 2018: 1431396,

2018.

3. Zhu Q, Sun Y, Zhou Q, He Q and Qian H: Identification of key genes and pathways by bioinformatics analysis with TCGA RNA sequencing data in hepatocellular carcinoma. Mol Clin Oncol 9: 597–606, 2018.
4. Shannon P, Markiel A, Ozier O, *et al*: Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13: 2498–2504, 2003.
5. Bindea G, Mlecnik B, Hackl H, *et al*: ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25: 1091–1093, 2009.
6. Bindea G, Galon J and Mlecnik B: CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. Bioinformatics 29: 661–663, 2013.
7. Szklarczyk D, Gable AL, Lyon D, *et al*: STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47: D607-D613, 2019.
8. Bader GD and Hogue CW: An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 4: 2, 2003.
9. Maere S, Heymans K and Kuiper M: BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics 21: 3448–3449, 2005.
10. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT and Lin CY: cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 8 Suppl 4: S11, 2014.
11. Tang Z, Li C, Kang B, Gao G, Li C and Zhang Z: GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 45: W98-W102, 2017.
12. Chandrashekar DS, Bashel B, Balasubramanya SAH, *et al*: UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. Neoplasia 19: 649–658, 2017.
13. Robinson MD, McCarthy DJ and Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139–140, 2010.
14. Menyhart O, Nagy A and Gyorffy B: Determining consistent prognostic biomarkers of overall survival and vascular invasion in hepatocellular carcinoma. R Soc Open Sci 5: 181006, 2018.
15. Hayashi T, Ohtsuka M, Okamura D, *et al*: Cytoskeleton-associated protein 2 is a potential predictive marker for risk of early and extensive recurrence of hepatocellular carcinoma after operative resection. Surgery 155: 114–123, 2014.
16. Tsuchihara K, Lapin V, Bakal C, *et al*: Ckap2 regulates aneuploidy, cell cycling, and cell death in a p53-dependent manner. Cancer Res 65: 6685–6691, 2005.
17. Guo QS, Song Y, Hua KQ and Gao SJ: Involvement of FAK-ERK2 signaling pathway in CKAP2-induced proliferation and motility in cervical carcinoma cell lines. Sci Rep 7: 2117, 2017.
18. Zhang S, Wang Y, Chen S and Li J: Silencing of cytoskeleton-associated protein 2 represses cell proliferation and induces cell cycle arrest and cell apoptosis in osteosarcoma cells. Biomed Pharmacother 106: 1396–1403, 2018.

19. Gao Y, Liu X, Li T, *et al*: Cross-validation of genes potentially associated with overall survival and drug resistance in ovarian cancer. Oncol Rep 37: 3084–3092, 2017.
20. Xu N, Chen SH, Lin TT, *et al*: Development and validation of hub genes for lymph node metastasis in patients with prostate cancer. J Cell Mol Med 24: 4402–4414, 2020.
21. Nandi D, Cheema PS, Jaiswal N and Nag A: FoxM1: Repurposing an oncogene as a biomarker. Semin Cancer Biol 52: 74–84, 2018.
22. Sheng Y, Yu C, Liu Y, *et al*: FOXM1 regulates leukemia stem cell quiescence and survival in MLL-rearranged AML. Nat Commun 11: 928, 2020.
23. Lin JZ, Wang WW, Hu TT, *et al*: FOXM1 contributes to docetaxel resistance in castration-resistant prostate cancer by inducing AMPK/mTOR-mediated autophagy. Cancer Lett 469: 481–489, 2020.
24. Arceci A, Bonacci T, Wang X, *et al*: FOXM1 Deubiquitination by USP21 Regulates Cell Cycle Progression and Paclitaxel Sensitivity in Basal-like Breast Cancer. Cell Rep 26: 3076-3086 e3076, 2019.
25. Cheng Z, Yu C, Cui S, *et al*: circTP63 functions as a ceRNA to promote lung squamous cell carcinoma progression by upregulating FOXM1. Nat Commun 10: 3200, 2019.
26. Hu G, Yan Z, Zhang C, *et al*: FOXM1 promotes hepatocellular carcinoma progression by regulating KIF4A expression. J Exp Clin Cancer Res 38: 188, 2019.
27. Shen S, Kong J, Qiu Y, Yang X, Wang W and Yan L: Identification of core genes and outcomes in hepatocellular carcinoma by bioinformatics analysis. J Cell Biochem 120: 10069–10081, 2019.
28. Weiler SME, Pinna F, Wolf T, *et al*: Induction of Chromosome Instability by Activation of Yes-Associated Protein and Forkhead Box M1 in Liver Cancer. Gastroenterology 152: 2037-2051 e2022, 2017.
29. Wei JC, Meng FD, Qu K, *et al*: Sorafenib inhibits proliferation and invasion of human hepatocellular carcinoma cells via up-regulation of p53 and suppressing FoxM1. Acta Pharmacol Sin 36: 241–251, 2015.
30. Barsotti AM and Prives C: Pro-proliferative FoxM1 is a target of p53-mediated repression. Oncogene 28: 4295–4305, 2009.
31. Pandit B, Halasi M and Gartel AL: p53 negatively regulates expression of FoxM1. Cell Cycle 8: 3425–3427, 2009.

Tables

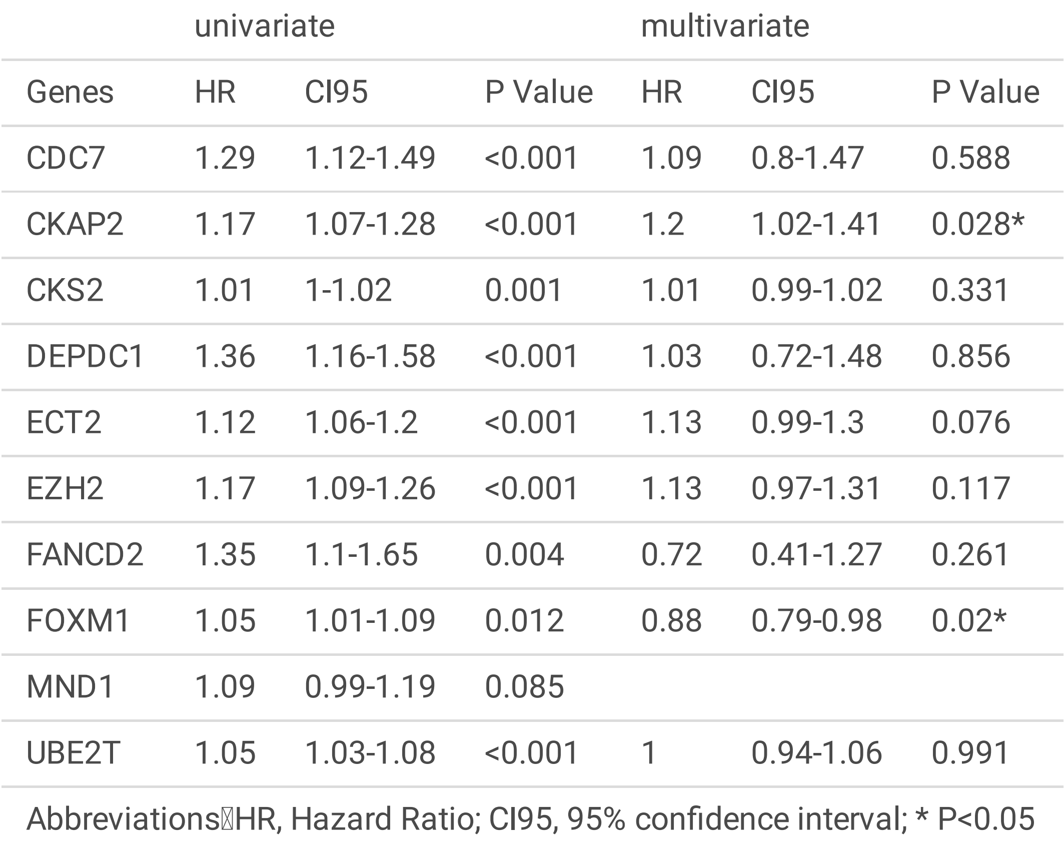
Table I:Baseline clinical characteristics of entire dataset(N=363) involved in this study

Characteristic	Freq	Characteristic	Freq
Age (%)		**M stage (%)**	
≤65	227 (63%)	M1-X	101 (28%)
>65	136 (37%)	M0	262 (72%)
Gender (%)		**N stage (%)**	
male	245 (67%)	N1-X	116 (32%)
female	118 (33%)	N0	247 (68%)
Family history of cancer (%)		**T stage (%)**	
YES	133 (37%)	T1-2	273 (75%)
NO	230 (63%)	T3-4	89 (25%)
Residual (%)		**Hepatitis (%)**	
R0	326 (90%)	YES	159 (44%)
R1-X	37 (10%)	NO	204 (56%)
Histologic grade (%)		**Alcohol (%)**	
G1–G2	234 (64%)	YES	122 (34%)
G3–G4	129 (36%)	NO	241 (66%)
TNM stage (%)		**Smoking (%)**	
stage i-ii	270 (74%)	YES	14 (4%)
stage iii-iv	93 (26%)	NO	349 (96%)
Vascular invasion (%)		**BMI (%)**	
YES	133 (37%)	≤24	186 (51%)
None	230 (63%)	<24	177 (49%)

Table II: 12 gene clusters calculated by MCODE plug-in from the PPI network

Cluster id	Score	Nodes	Edges	genes in the cluster
1	50.423	53	1311	MCM4, CKS2, RAD51AP1, CCNA2, NUF2, CDC7, PRC1, ECT2, DEPDC1, CKAP2, NDC80, SMC2, CENPE, KIAA0101, KIF20A, RACGAP1, BUB1B, ASPM, MAD2L1, NCAPG, SMC4, CEP55, MCM6, MCM7, AURKA, DLGAP5, CDKN3, KIF4A, FEN1, CCNB2, CCNB1, CDK1, TOP2A, BUB1, ZWINT, PBK, HMMR, NUSAP1, FOXM1, KIF2C, RFC4, MND1, UBE2T, DTL, BIRC5, RRM2, TPX2, KNTC1, FANCD2, FANCI, KIF11, CENPF, EZH2
2	8.235	18	70	HGFAC, KMO, HAAO, COLEC11, ALDH8A1, MASP1, ACSM3, CFP, MASP2, FTCD, FCN3, COLEC10, C8B, FCN2, MBL2, HAO2, KLKB1, F11
3	7	7	21	LAMA4, ITGA6, LAMC1, ITGA2, COL1A1, COL4A2, COL4A1
4	5.6	6	14	CYP2B6, CYP1A2, CYP2C9, CYP2C19, CYP4A11, NAT2
5	5.529	18	47	MT1X, FABP5, GCDH, GLA, GM2A, MT1F, SLC39A5, MT1H, MT1G, MT1E, ACADS, ACAA2, MT2A, ECHDC2, LPL, MT1M, ACLY, ANXA2
6	3.6	6	9	GLS2, BCKDHB, SDS, BCAT1, AGXT2, PHGDH
7	3.333	4	5	APOF, LCAT, LPA, NPC1L1
8	3.333	4	5	IGFBP3, IGFALS, GHR, PLG
9	3.2	6	8	CLIC1, PCK1, ALDOB, GCGR, GYS2, PPARGC1A
10	3	3	3	HAMP, STEAP3, FLVCR1
11	3	3	3	GPSM2, CCL20, CXCL2
12	2.889	10	13	ETFDH, ILF2, NEU1, ALDH6A1, ACADL, GNMT, PEMT, MAT1A, ACSL4, ACAA1

Table III: univariate and multivariate Cox proportional hazards regression analysis in the primary dataset



	univariate			multivariate		
Genes	HR	CI95	P Value	HR	CI95	P Value
CDC7	1.29	1.12-1.49	<0.001	1.09	0.8-1.47	0.588
CKAP2	1.17	1.07-1.28	<0.001	1.2	1.02-1.41	0.028*
CKS2	1.01	1-1.02	0.001	1.01	0.99-1.02	0.331
DEPDC1	1.36	1.16-1.58	<0.001	1.03	0.72-1.48	0.856
ECT2	1.12	1.06-1.2	<0.001	1.13	0.99-1.3	0.076
EZH2	1.17	1.09-1.26	<0.001	1.13	0.97-1.31	0.117
FANCD2	1.35	1.1-1.65	0.004	0.72	0.41-1.27	0.261
FOXM1	1.05	1.01-1.09	0.012	0.88	0.79-0.98	0.02*
MND1	1.09	0.99-1.19	0.085			
UBE2T	1.05	1.03-1.08	<0.001	1	0.94-1.06	0.991

Abbreviations：HR, Hazard Ratio; CI95, 95% confidence interval; * P<0.05

Table IV:Univariate and multivariate CPHR analysis of clinical-related gene signature and other risk factors in the primary dataset.

Characteristic	Univariate analysis		Multivariate analysis	
	HR (95%CI)	P-Value	HR (95%CI)	P-Value
Age (≥65 vs. <65)	1.04(0.67-1.6)	0.866		
Alcohol (no vs. yes)	0.74(0.47-1.16)	0.185	0.88(0.55-1.4)	0.582
BMI (<25 vs. ≥25)	0.82(0.53-1.26)	0.368		
Family history of cancer (no vs. yes)	0.95(0.61-1.47)	0.821		
Gender (male vs. female)	1.04(0.67-1.61)	0.863		
Hepatitis (YES vs. NO)	1.53(0.97-2.43)	0.068	1(0.6-1.67)	0.994
Histologic grade (G3-4 vs. G1-2)	1.14(0.73-1.77)	0.57		
M stage (M0 vs. M1-X)	0.68(0.44-1.06)	0.09	0.66(0.34-1.3)	0.231
N stage (N0 vs. N1-X)	0.74(0.47-1.15)	0.179	1.04(0.53-2.02)	0.914
Residual (R1-X vs. R0)	1.68(0.86-3.27)	0.13	1.39(0.68-2.82)	0.362
Signature (high-score vs. low-score)	1.82(1.18-2.82)	0.007	1.6(1.01-2.52)	0.045*
Smoking (YES vs. NO)	1.15(0.46-2.84)	0.767		
T stage (T3-4 vs. T1-2)	2.78(1.78-4.35)	<0.001		
TNM stage (iii-iv vs. i-ii)	2.78(1.78-4.35)	<0.001	2.44(1.49-4)	<0.001*
Vascular invasion (no vs. yes)	0.46(0.3-0.71)	0.001	0.54(0.33-0.87)	0.012*

Abbreviations: HR, hazard ratio; CI, confidence interval; * P<0.05

Figures

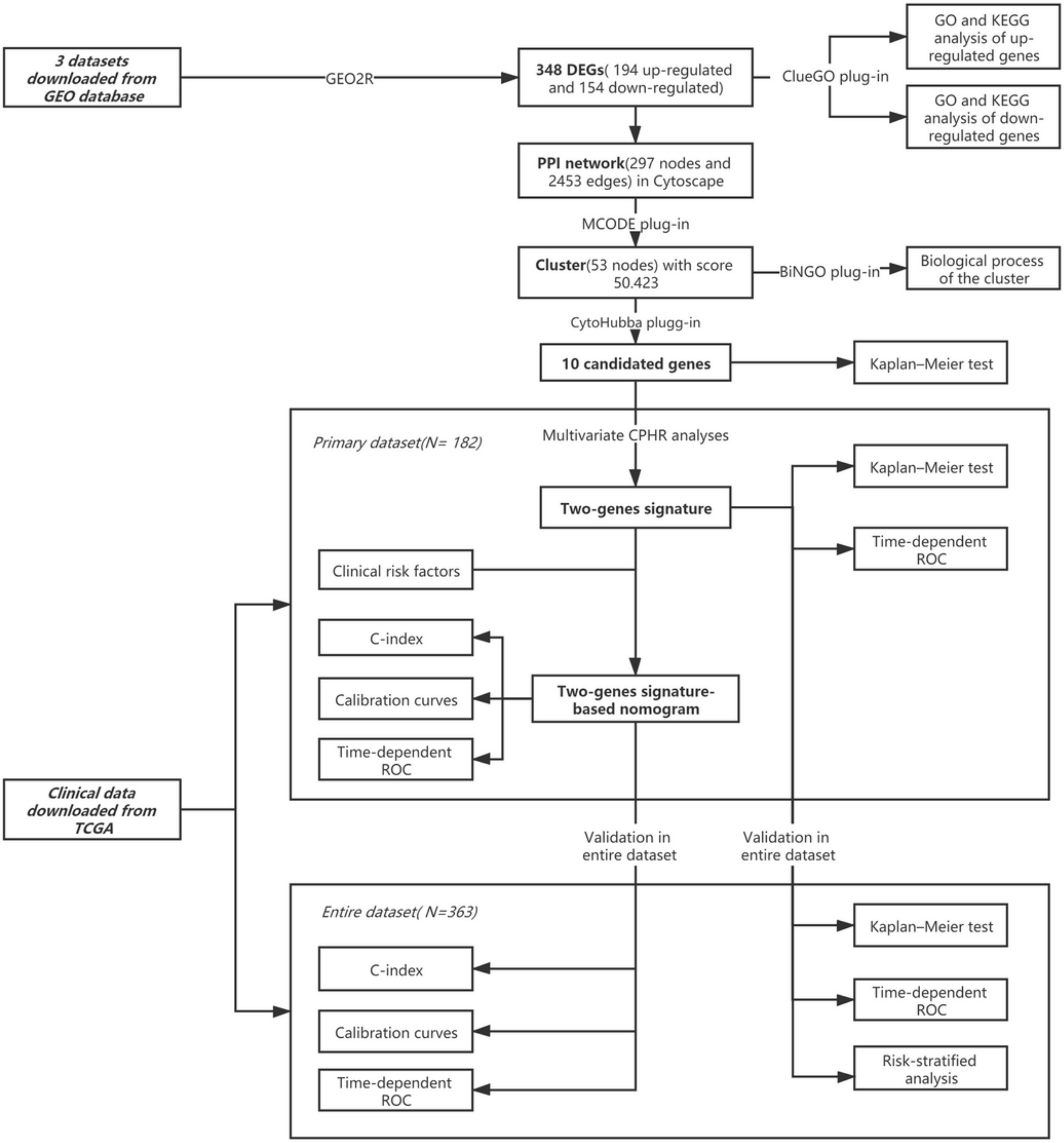

3 datasets downloaded from GEO database
GEO2R
348 DEGs(194 up-regulated and 154 down-regulated)
ClueGO plug-in
GO and KEGG analysis of up-regulated genes
GO and KEGG analysis of down-regulated genes
PPI network(297 nodes and 2453 edges) in Cytoscape
MCODE plug-in
Cluster(53 nodes) with score 50.423
BiNGO plug-in
Biological process of the cluster
CytoHubba plugg-in
10 candidated genes
Kaplan–Meier test
Primary dataset(N= 182)
Multivariate CPHR analyses
Two-genes signature
Kaplan–Meier test
Time-dependent ROC
Clinical risk factors
C-index
Calibration curves
Time-dependent ROC
Two-genes signature-based nomogram
Clinical data downloaded from TCGA
Validation in entire dataset
Validation in entire dataset
Entire dataset(N=363)
C-index
Calibration curves
Time-dependent ROC
Kaplan–Meier test
Time-dependent ROC
Risk-stratified analysis

Figure 1

Overview of article design ideas and operating procedures

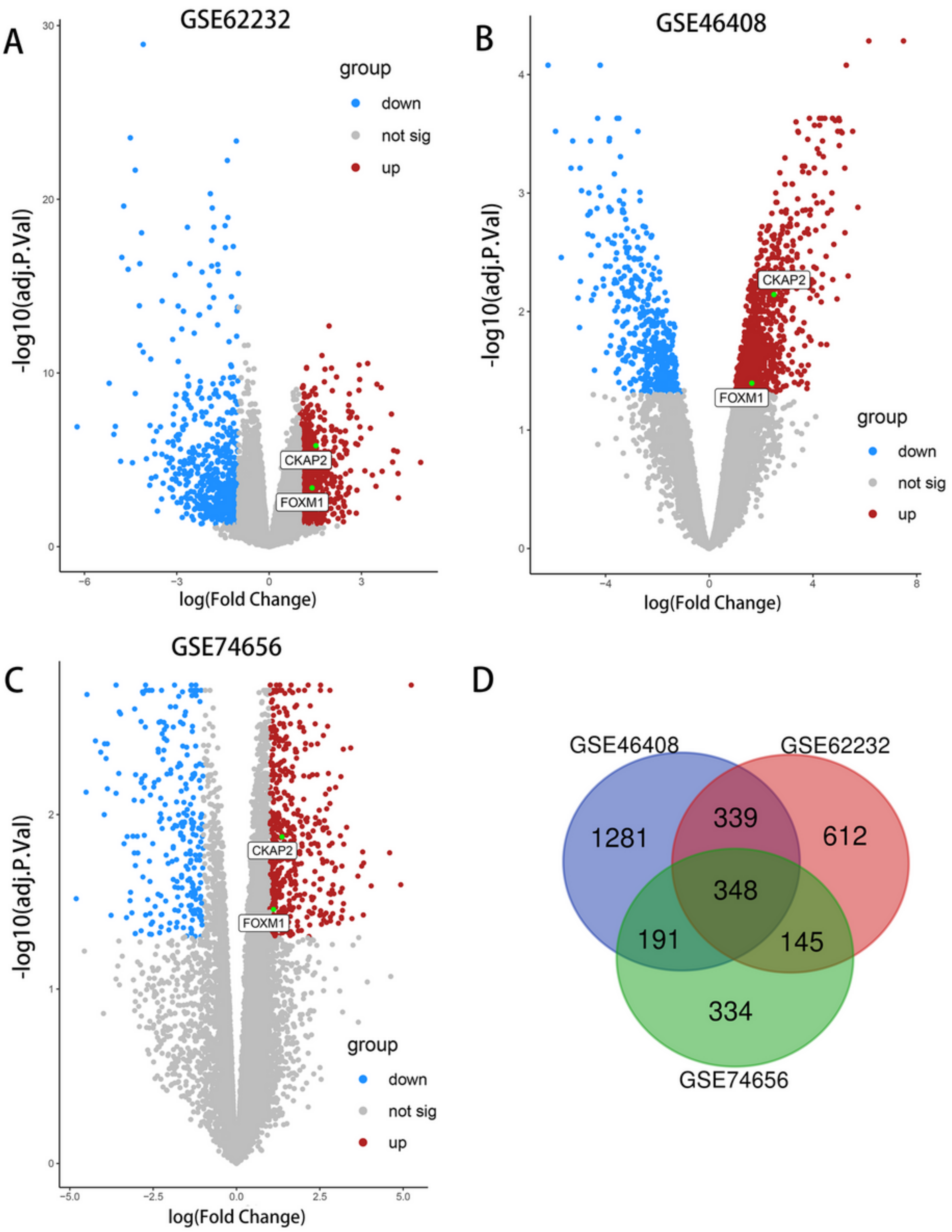


Figure 2

Distribution of DEGs in the three datasets. The volcano plot of GSE62232 (A), GSE46408 (B) and GSE74656 (C) dataset. Y axis of the volcano represents adjusted P-Value, X axis represents the fold change of those genes after log2 process. Red dots were set as up-regulated genes, while blue and grey dots were down-regulated and Insignificant genes respectively. (D) The venn plot of those three dastasets. DEGs: differentially expressed genes.

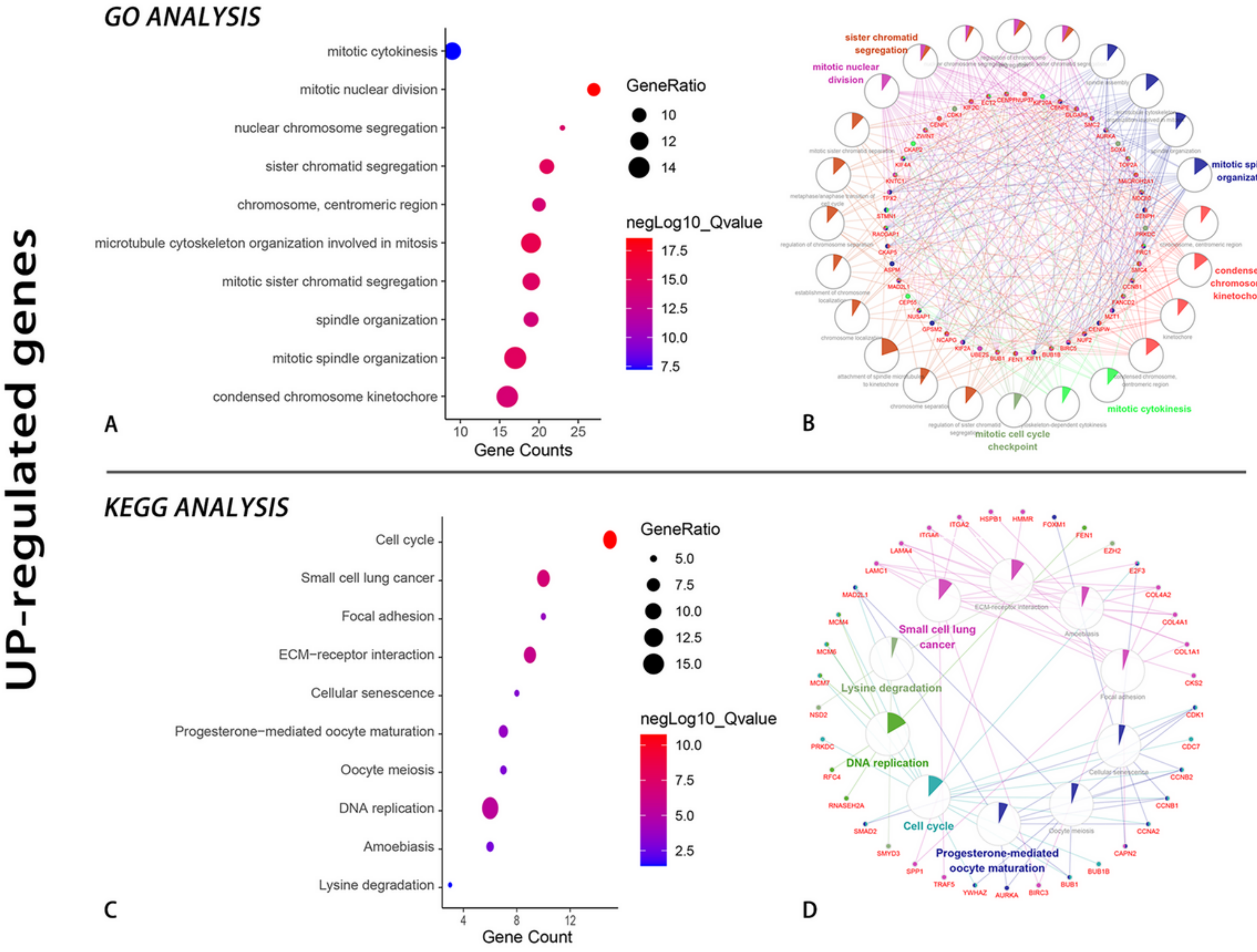


Figure 3

GO and KEGG enrichment analysis of 194 up-regulated genes in the Cytoscape software. (A) Bubble plot of top 10 enriched GO analysis. Color represents the degree of significant difference, red is defined as highly significant, while blue is low significant. Size of bubble represents the percentage of enriched genes in GO term. (B) Visualization of enriched GO results. (C) Bubble plot of top 10 enriched KEGG pathway analysis. (D) Visualization of enriched KEGG results.

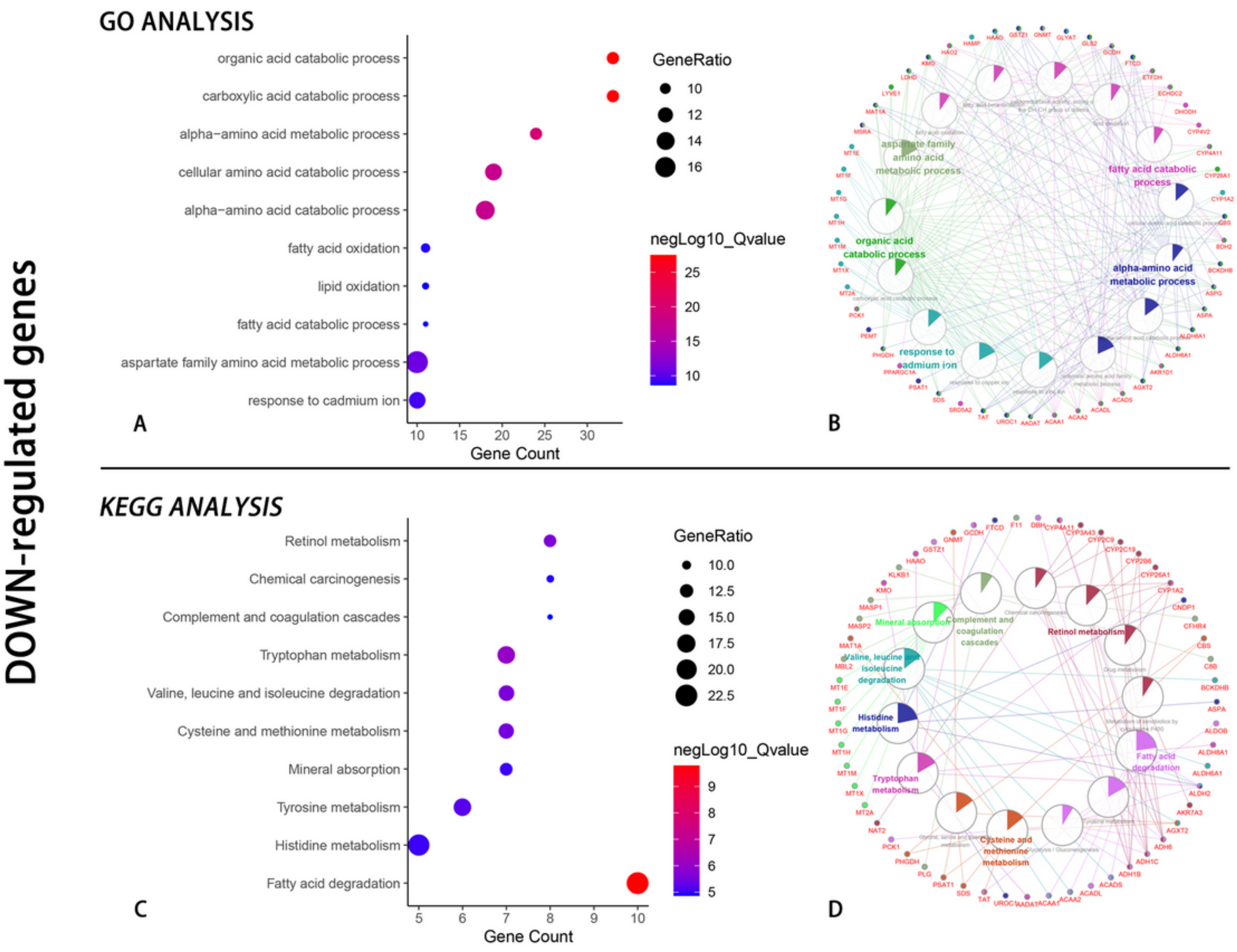


Figure 4

GO and KEGG enrichment analysis of 154 down-regulated genes in the Cytoscape software. (A) Bubble plot of top 10 enriched GO analysis. Color represents the degree of significant difference, red is defined as highly significant, while blue is low significant. Size of bubble represents the percentage of enriched genes in GO term. (B) Visualization of enriched GO results. (C) Bubble plot of top 10 enriched KEGG pathway analysis. (D) Visualization of enriched KEGG results.

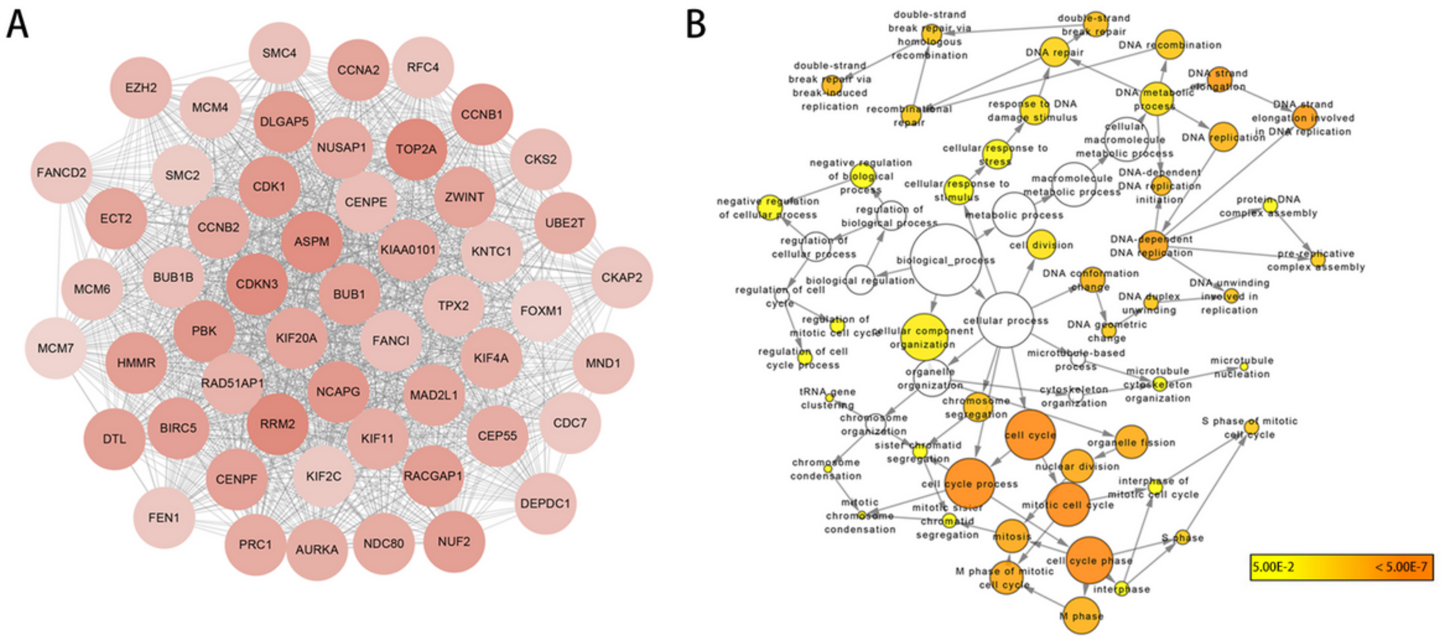

A
SMC4
CCNA2
RFC4
EZH2
MCM4
DLGAP5
CCNB1
NUSAP1
TOP2A
CKS2
FANCD2
SMC2
CDK1
ZWINT
CENPE
UBE2T
ECT2
CCNB2
ASPM
KIAA0101
KNTC1
CKAP2
MCM6
BUB1B
CDKN3
BUB1
TPX2
FOXM1
PBK
KIF20A
FANCI
MCM7
HMMR
KIF4A
MND1
RAD51AP1
NCAPG
MAD2L1
DTL
BIRC5
RRM2
KIF11
CEP55
CDC7
CENPF
KIF2C
RACGAP1
DEPDC1
FEN1
PRC1
AURKA
NDC80
NUF2
B
double-strand break repair via homologous recombination
double-strand break repair
DNA recombination
double-strand break repair via break-induced replication
DNA repair
DNA strand elongation
DNA metabolic process
DNA strand elongation involved in DNA replication
recombinational repair
response to DNA damage stimulus
cellular macromolecule metabolic process
DNA replication
cellular response to stress
negative regulation of biological process
macromolecule metabolic process
DNA-dependent DNA replication initiation
cellular response to stimulus
protein-DNA complex assembly
negative regulation of cellular process
regulation of biological process
metabolic process
regulation of cellular process
cell division
DNA-dependent DNA replication
pre-replicative complex assembly
biological_process
biological regulation
DNA conformation change
DNA unwinding involved in replication
regulation of cell cycle
DNA duplex unwinding
regulation of mitotic cell cycle
cellular process
cellular component organization
DNA geometric change
regulation of cell cycle process
microtubule-based process
microtubule nucleation
organelle organization
microtubule cytoskeleton organization
tRNA gene clustering
cytoskeleton organization
chromosome segregation
chromosome organization
cell cycle
S phase of mitotic cell cycle
organelle fission
sister chromatid segregation
chromosome condensation
nuclear division
cell cycle process
interphase of mitotic cell cycle
mitotic cell cycle
mitotic chromosome condensation
mitotic sister chromatid segregation
mitosis
S phase
cell cycle phase
M phase of mitotic cell cycle
interphase
M phase
5.00E-2
< 5.00E-7

Figure 5

Module analysis of 53 most closely connected genes in the PPI network. (A) Overview of cluster with the highest score. Each dot represents a kind of protein, the shade of color represents the expression of the protein, and the edges indicate the relationship between the 2 gene that are linked. (B) the biological processes of this cluster. The size of node is positively related with the number of the enriched genes in the GO terms; while the color scale shows the change of P-value, the more prominent, the darker the color. PPI: Protein to Protein Interaction.

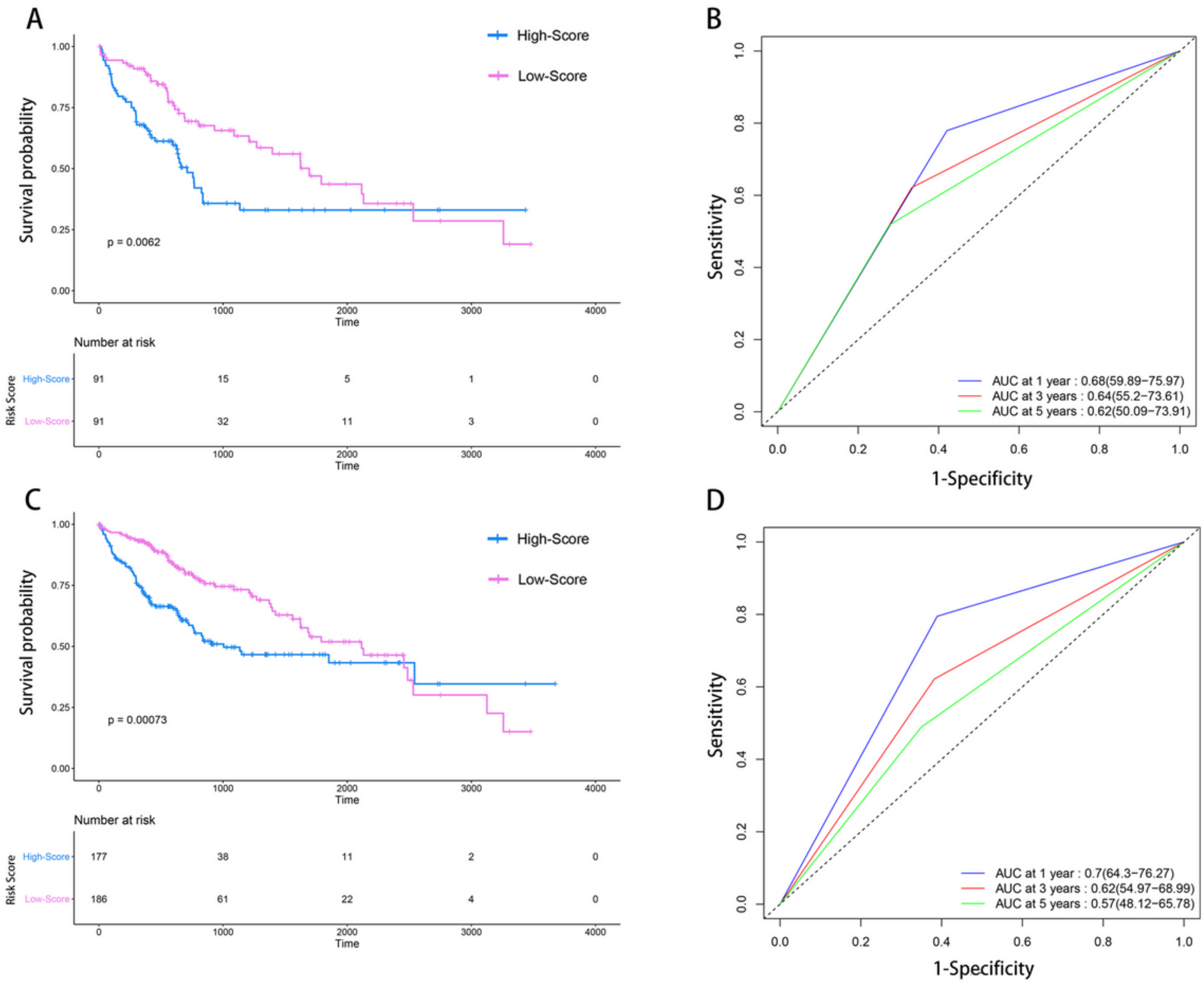


Figure 6

Predicting performance of the OS-related genes signature in the two datasets. Kaplan-Meier curves (A) and time-dependent ROC curve (B) for overall survival based on the OS-related genes signature in the primary dataset. Kaplan-Meier curves (C) and time-dependent ROC curve (D) of OS-related genes signature was also performed in the entire dataset. The number of patients at risk is listed below the KM curve. ROC curve: Receiver Operating Characteristic curve, KM: Kaplan-Meier.

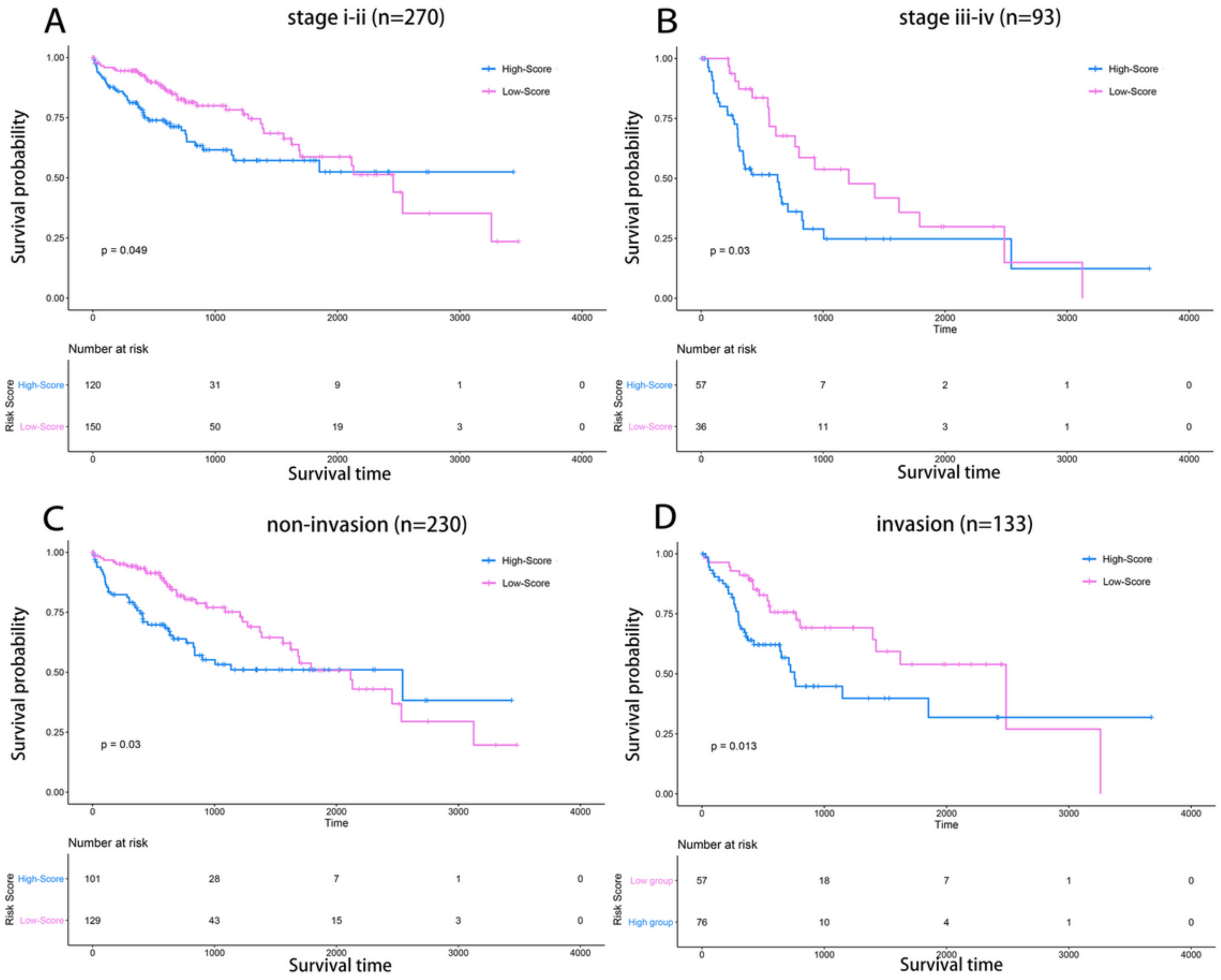


Figure 7

Risk-stratified analysis of the OS-related genes signature for HCC patients. Kaplan-Meier analysis were performed in HCC patients of TNM stage i-ii subgroup (A), TNM stage iii-iv subgroup (B), vascular non-invasion subgroup (C) and vascular invasion subgroup (D). The digital below the KM curve represents the number of under risking patients.

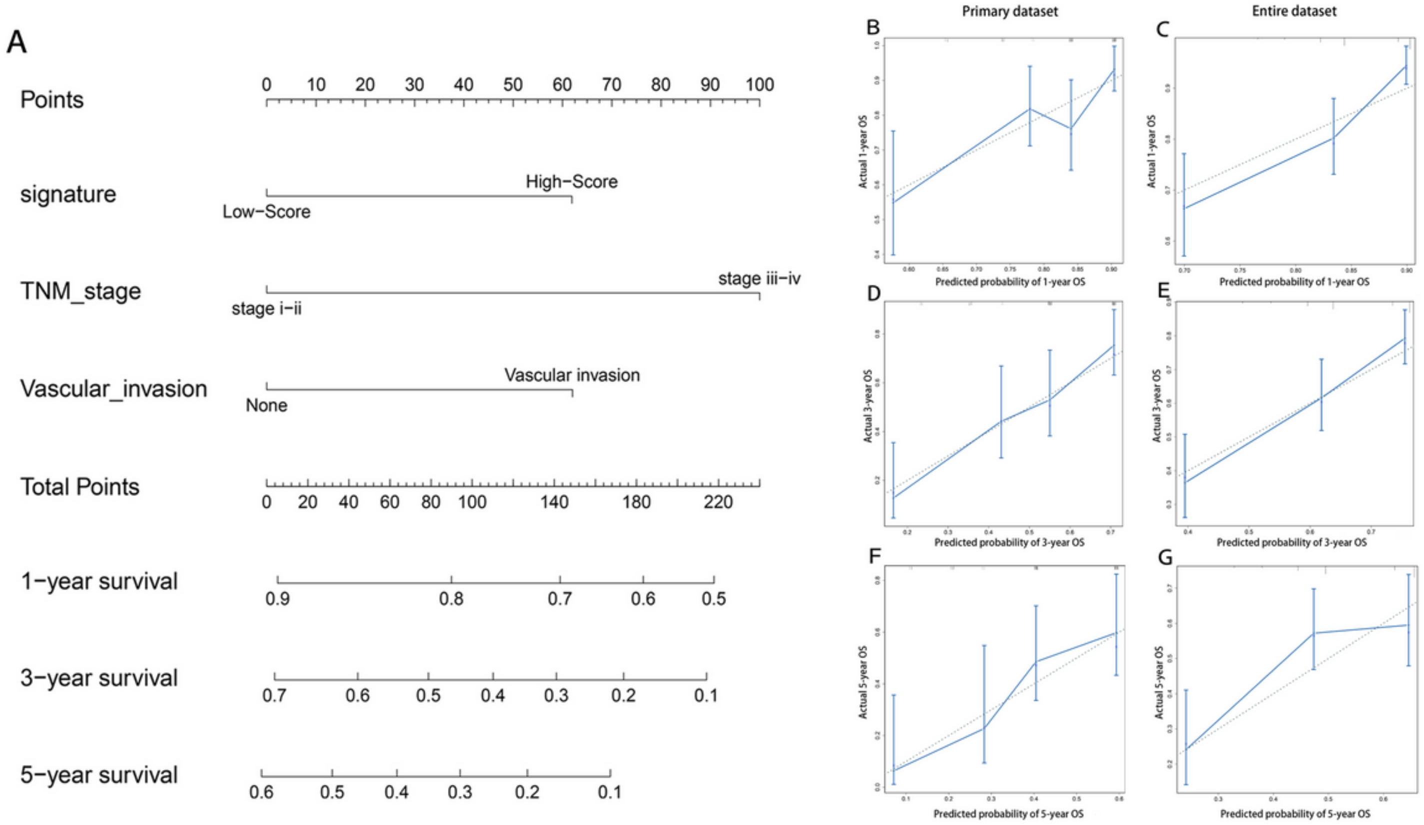


Figure 8

Nomogram based on the OS-related genes signature and TNM stage and vascular invasion. (A) Nomogram used for predicting overall survival in HCC patients. Explanation: the scale of signature, TNM stage and vascular invasion vertical upward correspond to their respective scores. Then, taking the total points after summing the three scores, and the scale of the total points vertical downward to the one-, three- or five-year OS for a specific liver cancer patient. Calibration plots of the nomogram for predicting OS at one years (B) and three years (D) and five years (F) in the primary dataset, and at one years (C), three years (E) and five years (G) in the entire dataset. X axis of the calibration plot represents the predictive probability, while Y axis of the plot indicates the actual survival probability. The 45-degree dotted line represents a perfect prediction, and the solid line represents the predictive performance of the nomogram.

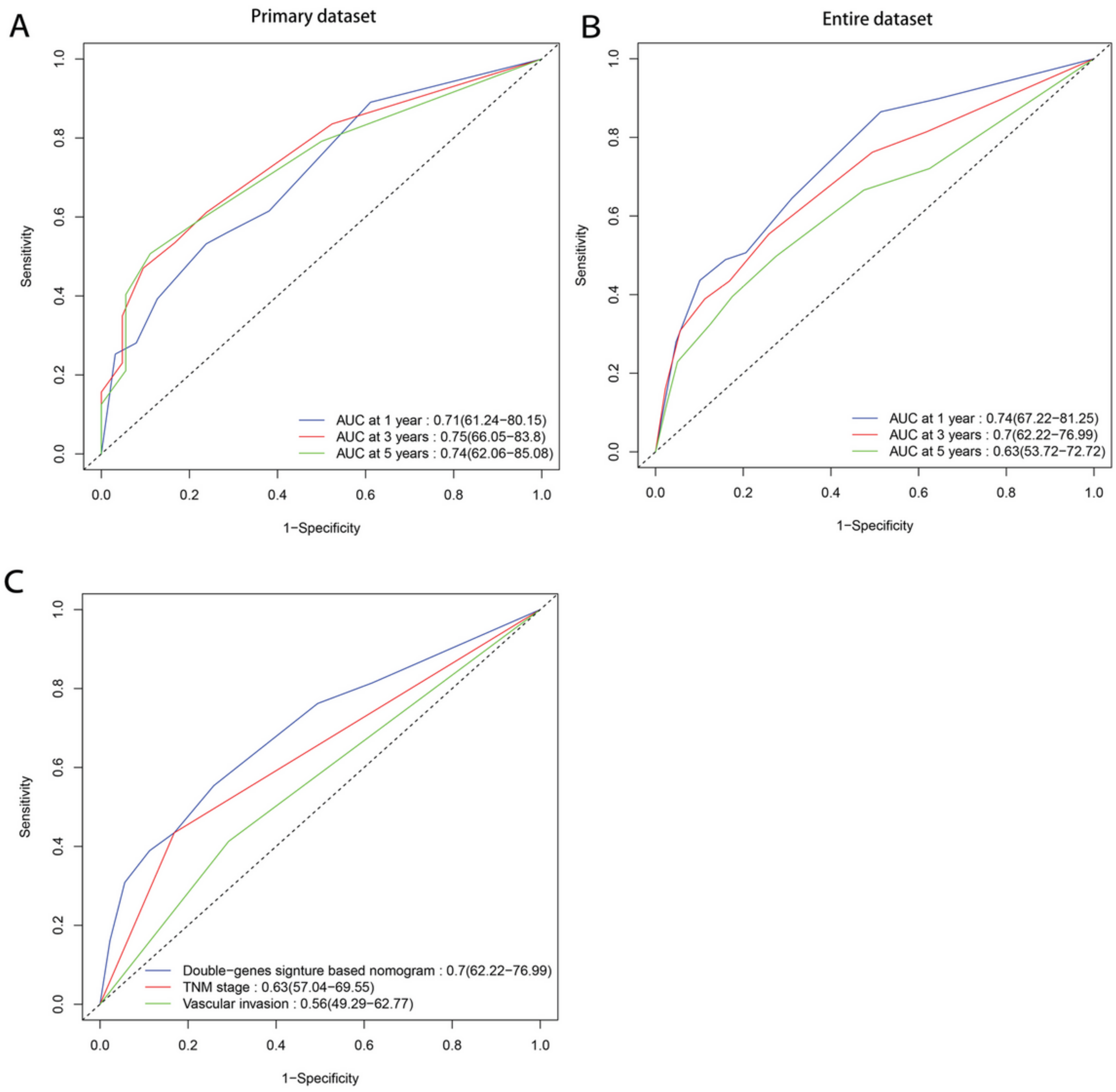


Figure 9

The predictive ability of nomogram in comparison with other traditional clinical factors. The time-dependent ROC curves of the nomogram for predicting OS in the entire dataset (A) and the primary dataset (B). (C) The prognostic accuracy of the two-genes-based nomogram compared with those of the TNM stage and vascular invasion.

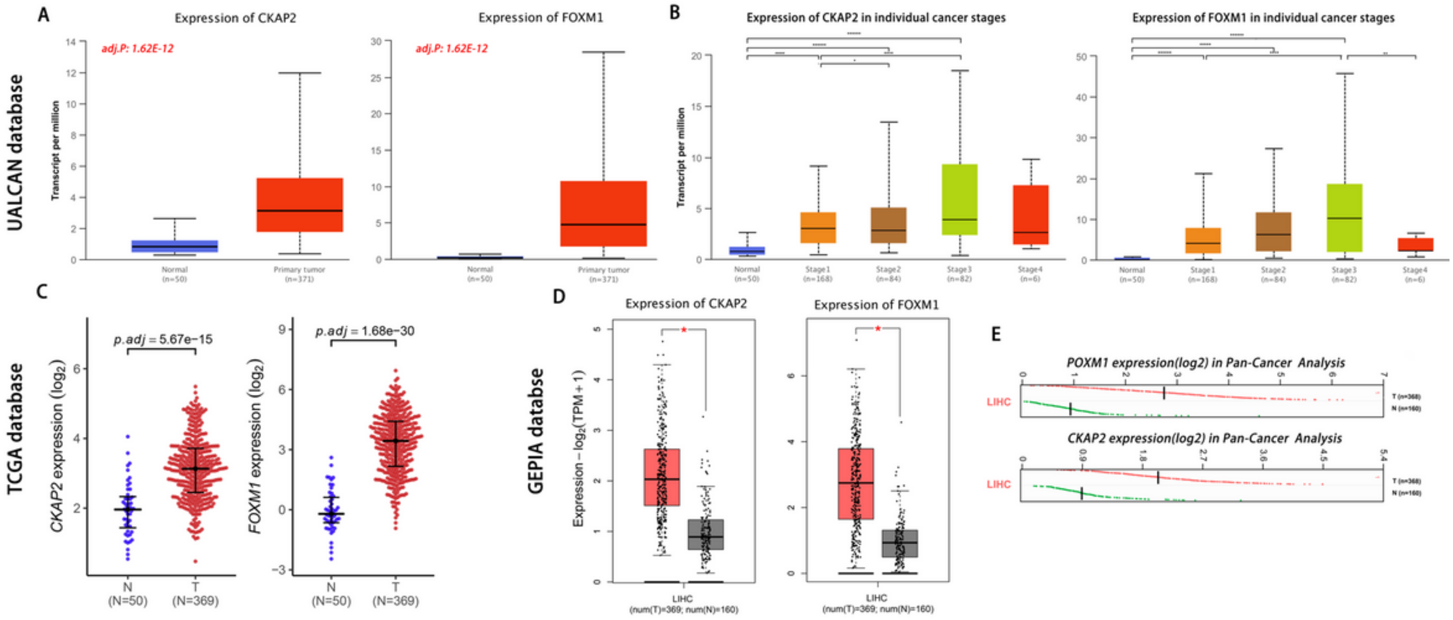


Figure 10

The gene of CKAP2 and FOXM1’s transcriptional expression was validated in the public database. The expression level of CKAP2 and FOXM1 in the UALCAN database (A) and GEPIA database (D). The expression change of CKAP2 and FOXM1 between the normal tissues and tumor tissues was also manifested with R (C). The relationship between tumor stage and CKAP2 and FOXM1’s expression change in the UALCAN database (B). Pan-cancer analysis of CKAP2 and FOXM1 in the GEPIA database (E). * P<0.05, ** P<0.01, *** P<10-3, **** p<10-4, ***** p<10-5.

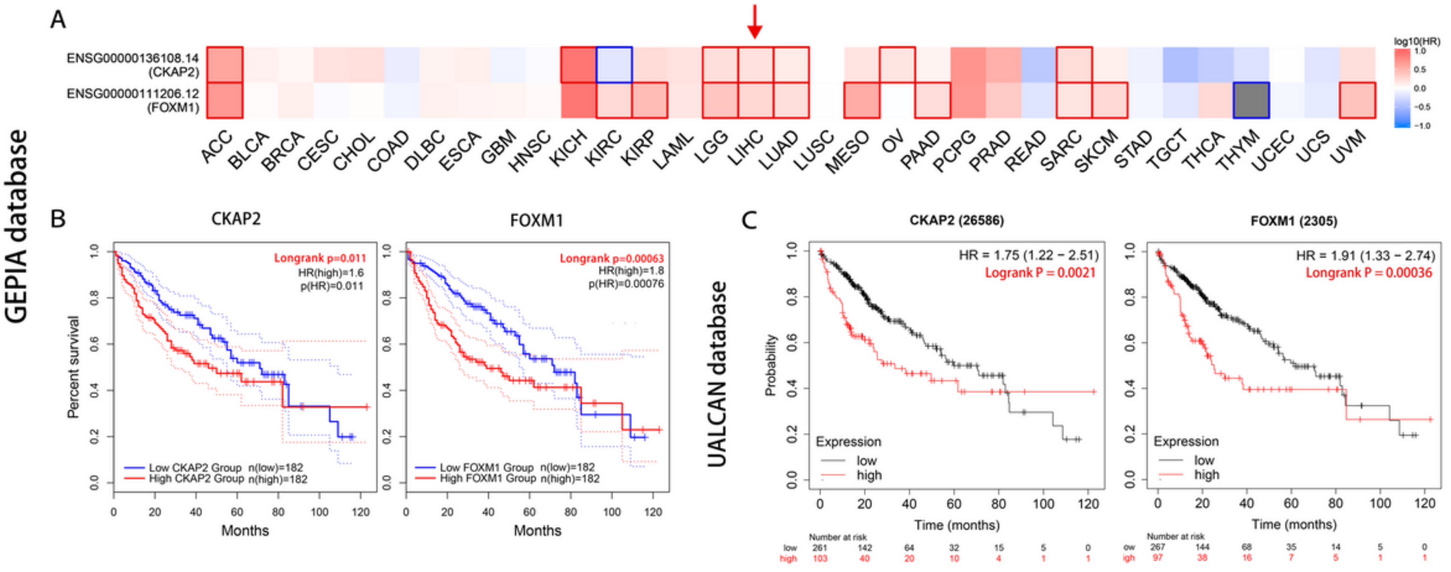


Figure 11

Pan-cancer survival analysis and Kaplan-Meier plot of CKAP2 and FOXM1 were validated in the public database. (A) the pan-cancer survival heatmap of CKAP2 and FOXM1 in the GEPIA database. Risk factor

(HR>1) was colored as red, while protective factor (HR<1) as blue. Red frame represents the statistically difference. Kaplan-Meier plot of CKAP2 and FOXM1 between the high/low expression subgroup in the GEPIA database (B) and UALCAN database (C). LIHC: Liver hepatocellular carcinoma.

Supplementary Files

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

- FigureS1Missingvalue.tif
- FigureS2PPInetwork.tif
- FigureS310geneKM.tif
- TableS1.docx
- TableS2.docx