

A prognostic model for hepatocellular carcinoma based on apoptosis-related genes

Renjie Liu

Yangzhou University Affiliated Northern Jiangsu People's Hospital: Northern Jiangsu People's Hospital

Guifu Wang

Yangzhou University Affiliated Northern Jiangsu People's Hospital: Northern Jiangsu People's Hospital

Chi Zhang

Yangzhou University Affiliated Northern Jiangsu People's Hospital: Northern Jiangsu People's Hospital

Dousheng Bai (✉ drbaidousheng@163.com)

Yangzhou University Affiliated Northern Jiangsu People's Hospital: Northern Jiangsu People's Hospital

<https://orcid.org/0000-0001-8032-7959>

Research

Keywords: Apoptosis-related genes, hepatocellular carcinoma, prognostic model, The Cancer Genome Atlas, nomogram

Posted Date: February 17th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-141169/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on March 12th, 2021. See the published version at <https://doi.org/10.1186/s12957-021-02175-9>.

Abstract

Background: Dysregulation of the balance between proliferation and apoptosis is the basis for human hepatocarcinogenesis. In many malignant tumors, such as hepatocellular carcinoma (HCC), there is a correlation between apoptotic dysregulation and poor prognosis. However, the prognostic values of apoptosis-related genes (ARGs) in HCC have not been elucidated.

Methods: To screen for differentially expressed ARGs, the expression levels of 161 ARGs from The Cancer Genome Atlas (TCGA) database(<https://cancergenome.nih.gov/>) were analyzed. Gene Ontology (GO) enrichment and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were performed to evaluate the underlying molecular mechanisms of differentially expressed ARGs in HCC. The prognostic values of ARGs were established using Cox regression, and subsequently, a prognostic risk model for scoring patients was developed. Kaplan-Meier (K-M) and receiver operating characteristic (ROC) curves were plotted to determine the prognostic value of the model.

Results: Compared to normal tissues, 43 highly up-regulated and 8 down-regulated ARGs in HCC tissues were screened. GO analysis results revealed that these 51 genes are indeed related to the apoptosis function. KEGG analysis revealed that these 51 genes were correlated with MAPK, P53, TNF, and PI3K-AKT signaling pathways, while Cox regression revealed that 5 ARGs (PPP2R5B, SQSTM1, TOP2A, BMF, and LGALS3) were associated with prognosis and were, therefore, obtained to develop the prognostic model. Based on the median risk scores, patients were categorized into high-risk and low-risk groups. Patients in the low-risk groups exhibited significantly elevated two-year or five-year survival probabilities ($p < 0.0001$). The risk model had a better clinical potency than the other clinical characteristics, with the area under the ROC curve (AUC = 0.741). The prognosis of HCC patients was established from a plotted nomogram.

Conclusion: Based on the differential expression of ARGs, we established a novel risk model for predicting HCC prognosis. This model can also be used to inform the individualized treatment of HCC patients.

Introduction

Globally, liver cancer is the sixth most common tumor and the fourth leading cause of cancer-associated mortalities [1]. Among the frequent primary liver cancers, hepatocellular carcinoma (HCC) accounts for approximately 75% of all cases [1]. Despite advances in diagnostic techniques and treatment, HCC is still associated with poor survival outcomes due to the high rate of recurrence and metastasis [2-4]. The TNM staging system is a traditional method for prognostic prediction, however, it lacks performance accuracy due to substantive variations in HCC clinical outcomes [5]. Over the past decades, serum alpha-fetoprotein (AFP) has been the only biomarker for detecting and predicting the prognosis of HCC, however, its low sensitivity limits its clinical utility [6]. Therefore, the identification of a novel prognostic

biomarker and establishment of an advanced prognostic model for HCC patients is of paramount importance.

Bioinformatics analysis is important in elucidating the functions of numerous differentially expressed genes as well as evaluating the complexity of HCC occurrence and development [7, 8]. Meng et al. used a series of bioinformatics analyses to identify hub genes and pathways associated with HCC pathogenesis and prognosis [9]. However, these studies usually ignore clinical information such as sex, age, grade, and stage of tumors. It may be very innovative, and informative to construct a prognostic model that combines patient's gene expression level and clinical information. Hepatocarcinogenesis develops following an imbalance between proliferation and apoptosis [10]. It has been documented that the overexpression of spindle and kinetochore-related complex subunit 3 (SKAT3) in HCC inhibits P53 activation by binding cyclin-dependent kinase 2 (CDK2), before impeding cell apoptosis, and thereby promoting cancer cell proliferation [11]. Besides, several biomolecules may influence HCC prognosis by regulating apoptosis-related genes (ARGs) or apoptosis-related pathways [12-14]. Yu et al. evaluated the association between the haplotype of the apoptosis-related gene cyclin-dependent kinase inhibitor 1B (CDKN1B) and the prognosis of HCC patients who were subjected to surgical resection [15]. The CCT/ACT haplotype patients were found to exhibit lower overall survival rates than those with the more common ACT/CCT haplotype. Therefore, ARGs can potentially be used to assess HCC prognosis.

Based on the gene expression and clinical characteristics data obtained from the Cancer Genome Atlas (TCGA) database, we established the ARGs associated with HCC prognosis and developed a prognostic prediction model. The model calculates the risk score to predict and evaluate the HCC prognosis.

Materials And Methods

Data collection

The mRNA expression data and clinical information of HCC patients were obtained from the TCGA database (<https://cancergenome.nih.gov/>). The obtained clinical information included age, gender, TNM stage, T stage, N stage, M stage, and histological grade. A total of 161 ARGs were acquired from the gene set "HALLMARK_APOPTOSIS" in the Molecular Signatures Database v7.1 in GSEA [16].

Gene Set Enrichment Analysis and Differentially Expressed ARGs

Gene sets with significant differences between HCC and normal samples were evaluated by GSEA. Subsequently, using the mRNA expression profiles, the limma package and the Wilcoxon signed-rank test in R software 3.6.2 ($|\log_2FC| > 1$, $FDR < 0.05$) were used to show the significantly differently expressed ARGs in the HCC cohort. The pheatmap and ggpubr packages in R software were used to develop volcano plots, heatmaps, and box plots.

Functional enrichment, KEGG pathway and PPI network analysis

Gene ontology (GO) enrichment and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were performed to evaluate the potential biological functions of ARGs, after which they were visualized through R software packages such as ggplot2, DOSE, clusterProfiler, enrichplot, GOplot, digest, etc. Interactions among the selected ARGs were determined through protein-protein interaction (PPI) networks from the STRING database (<http://www.string-db.org/>) [17] and visualized by Cytoscape software [18].

Establishment of a Prognostic risk model based on ARGs

Univariate and multivariate Cox proportional hazard regression analyses were performed to identify prognosis-associated ARGs in HCC. Then, a prognosis-associated prediction formula, acquired from multivariate Cox regression analysis, was used to construct a prognostic model using the “glmnet” package in R. Using the prognostic model, Kaplan-Meier (K-M) analysis was performed to evaluate the survival rates of the high- and low-risk groups. Subsequently, the area under the receiver operating characteristic (ROC) curve (AUC), and KEGG enrichment analysis were used to assess the predictive value of the prognostic model. Finally, the R package (rms) was used to develop a risk model-based nomogram for predicting the prognosis of HCC patients.

Statistical Analysis

The gene expression data were normalized by log2 transformation. Thereafter, statistical analyses were performed while plots were constructed using the R software 3.6.2 (<https://www.r-project.org/>) and Perl language packages. $p \leq 0.05$ was considered statistically significant.

Results

Acquisition of apoptosis-related genes set and GSE analysis

Following the search in the Molecular Signatures Database v7.1 in GSEA, 161 ARGs were identified from the gene set “HALLMARK_APOPTOSIS”. Thereafter, the 161 ARGs in this gene set were further assessed through GSEA analysis to ascertain their biological significance in HCC. Fig. 1a shows that ARG set was significantly differentially expressed between HCC and normal samples.

Identification of differentially expressed ARGs

The mRNA sequence data for 374 HCC tissue samples and 50 samples of normal tissue were obtained from the TCGA database. The limma package and the Wilcoxon signed-rank test in R ($|\log_2FC| > 1$, $FDR < 0.05$) were used to screen differentially expressed ARGs in HCC and non-tumor samples. In this study, 43 and 8 ARGs were found to be significantly upregulated and downregulated, respectively. These findings are presented in a volcano plot, heatmap, and box plot (Fig. 1(b-d)).

Functional enrichment and PPI network analysis of differentially expressed ARGs

To establish the biological functions and significant pathways of the 51 identified genes, GO enrichment and KEGG pathway enrichment analysis were performed (Fig. 2). The 51 identified ARGs were found to be involved in the pathways associated with cellular apoptosis (Fig. 2a and 2b). Furthermore, KEGG enrichment analysis revealed that these ARGs are involved in platinum drug resistance, transcriptional dysregulation in cancer and some oncogenic pathways such as MAPK, P53, TNF, and PI3K-AKT signaling pathways (Fig. 2c). The STRING online tool was used to establish a PPI network to evaluate the interactions among ARGs coded proteins. Results were observed using the Cytoscape software (Fig. 2d).

Association between ARGs and HCC patient's survival and prognosis

Univariate Cox regression analyses were performed on both mRNA expression and the corresponding clinical data for the 51 selected ARGs to identify their prognosis-associated ARGs in HCC (Fig. 3a). There were 20 upregulated ARGs and 1 downregulated ARG which were statistically significant. These 21 genes were subsequently analyzed by multivariate Cox regression ($p < 0.05$) to determine their association with the prognosis of HCC patients and to acquire the corresponding regression coefficients. Five prognosis-related ARGs: Protein phosphatase 2 regulatory subunit B'beta (PPP2R5B), Sequestosome 1 (SQSTM1), DNA topoisomerase II alpha (TOP2A), Bcl2 modifying factor (BMF), and Galectin 3 (LGALS3) were screened. Expression levels in normal and HCC samples were further compared to establish a prognostic value involving these 5 genes. Compared to normal samples, the 5 identified ARGs in HCC samples exhibited significantly elevated expression levels than normal specimens as indicated by heatmap and boxplot (Fig. 3b and 3c). Besides, the K-M curve was constructed by utilizing the survival rate differences in the high- and low-expressed groups of the identified ARGs. Interestingly, the high expression levels of the 5 identified ARGs indicated a low survival rate (Fig. 3d).

Furthermore, mutations in these 5 HCC genes were analyzed through the cBioPortal database (<http://cbioportal.org>). Data from 353 HCC patients in this database revealed that 22 patients (6.3%) had mutations. Among the 22 patients with the mutation, 0.85% had missense mutations, 0.56% had amplifications, while 0.56% had deep deletions in the PPP2R5B gene; 0.28% had missense mutations, 0.85% had amplifications while 0.28% had deep deletions in the SQSTM1 gene; while 0.85% had missense mutations, 0.28% had truncating mutations, 0.56% had amplifications and 0.28% had deep deletions in the TOP2A gene. Moreover, 0.28% had missense mutations, and 0.28% had deep deletions in the BMF gene whereas 0.28% had amplifications in the LGALS3 gene (Fig. 3e).

Immunohistochemical analysis of the Human Protein Atlas (HPA) database revealed that these 5 genes were significantly upregulated in HCC (Fig. 4).

Establishment of a prognostic risk signature based on ARGs

We combined the expression levels of ARGs and the regression coefficients of multiple Cox regression analyses to establish a risk scoring formula. Risk score = (0.385327 * expression value of PPP2R5B) + (0.2787 * expression value of SQSYM1) + (0.152062 * expression value of TOP2A) + (0.172177 * expression value of BMF) + (0.110211 * expression value of LGALS3). All genes had a positive

coefficient, implying that elevated expression levels of the identified genes were negatively correlated with prognosis. After calculating the risk score in HCC patients, the median risk score was used as a cutoff, and these patients were assigned into high- and low-risk groups (Fig. 5). The plotted heatmap of the 5 ARGs expression levels showed that patients in the same group had distinct expression levels (Fig. 5a). Patient's scores were ranked in ascending order (Fig. 5b) and their survival time presented as a scatterplot (Fig. 5c). Low-risk patients had a longer survival time and higher survival rates than high-risk patients.

Survival analyses of high- and low-risk groups were performed to show the correlations between risk scores and the prognosis of patients. First, we evaluated the prognostic significance of the risk scores using K-M curves and the log-rank method as shown in (Fig. 6). Compared to the high-risk group, patients in the low-risk group had a significantly high two-year or five-year survival probability ($p < 0.0001$; Fig. 6) revealing a worse prognosis for the high-risk group.

Independent prediction of HCC prognosis by risk score

Both univariate and multivariate Cox regression analyses were used to compare the predictive value of the prognostic model and other clinical prognostic variables. The risk score was established as an independent prognostic predictor for HCC patients (Fig. 7a and 7b). However, clinical variables were found not to be independent prognostic predictors ($p > 0.05$; Fig. 7b). Moreover, the ROC curve was plotted to validate the predictive value of the risk score (Fig. 7c). The AUC of the risk score was 0.741 as the highest value, compared to that of the clinical variables, implying that the risk score had a higher predictive value than the clinical variables.

KEGG enrichment analysis between different risk patients

KEGG analysis was used to evaluate to explore the potential biological pathways that were associated with 5 ARGs. The 5 majorly enriched pathways were apoptosis, cell cycle, mechanistic target of rapamycin kinase (mTOR) signaling, WNT signaling, and phosphatidylinositol signaling systems (Fig. 7d).

Construction of a nomogram to predict the prognosis of HCC patients

To apply the risk score in predicting the prognosis of HCC patients, we combined the risk score with the corresponding clinical variables to construct a nomogram to predict the OS of patients at 1, 2, and 3 years. Based on the risk scores and clinical variables, an average point for a patient can be established and extrapolated to determine the 1-, 2- and 3-year OS (Fig. 8).

The correlations between risk scores/ 5 ARGs and clinical variables

Based on the gene expression and corresponding clinical data obtained from the TCGA database, we analyzed the correlations between these clinical variables and risk scores of the 5 ARGs. Risk scores were correlated with tumor stage; BMF was associated with age, gender, grade, and stage; PPP2R5B and

LGALS3 were correlated with stage; SQSTM1 was associated with age and gender; while TOP2A was correlated with grade and stage (Fig. 9).

Discussion

HCC is the most common malignancy that is associated with a high recurrence rate and poor prognosis [19, 20]. Several limitations have been attributed to the late diagnosis and treatment of HCC [21]. Therefore, identification of prognostic-related biomarkers and the development of a prognostic model is of critical importance for HCC patients. Currently, bioinformatics analyses have been shown to play an increasingly significant role in identifying therapeutic targets for the diagnosis, treatment, and prognosis of numerous tumors [22]. Apoptosis plays a crucial role in liver tumor development and regeneration [12, 23]. Dysregulated apoptosis leads to the occurrence and development of liver tumors [24].

In this study, we developed a novel prognostic-predictive model for HCC based on ARGs expression. Based on the 161 ARGs list from the GSEA and the data from the TCGA database, 43 significantly upregulated ARGs and 8 significantly downregulated ARGs in HCC samples were screened out. GO enrichment analysis revealed that these 51 genes were enriched in the pathways associated with cellular apoptosis, while KEGG analysis revealed that the 51 genes were enriched in MAPK, P53, TNF, and PI3K-Akt signaling pathways. Among the 51 genes we screened, several genes have been reported to be related to the prognosis of HCC, such as apoptosis regulator BCL2 associated X (BAX) [25], SQSTM1 [26] and CDK2 [27]. Therefore, it is feasible to evaluate a prognostic model with these genes. Besides, 5 prognosis-related ARGs (PPP2R5B, SQSTM1, TOP2A, BMF, and LGALS3) were identified. Elevated expression levels of these 5 ARGs were negatively correlated with prognosis. Based on this, we designed a risk model and it was considered as an independent prognostic model of HCC according to the results of multivariate regression analysis. The predictive value of the model was confirmed by the K-M and ROC curves, which showed the highest prognostic predictive power of the model when combined with other clinical features of HCC patients. Therefore, by combining risk scores and clinical features as the basis for developing the nomogram, its efficiency in predicting the prognosis of HCC patients was enhanced.

The potential impact of these 5 hub ARGs on the progression of HCC were established. For example, TOP2A was identified as a potential biomarker for cancer therapy in ovarian cancer, colon cancer, pancreatic adenocarcinoma, and HCC [9, 22, 28, 29]. However, TOP2A was up-regulated in HCC [30]. Overexpressed TOP2A in HCC leads to a worse prognosis and its inhibitors have a potential therapeutic effect in HCC patients [31]. Autophagy was demonstrated to suppress spontaneous tumorigenesis in the liver [32]. SQSTM1, an autophagy-related protein, participates in cell survival, growth, and death through several pathways, and degrades during autophagy [33, 34]. Overexpression of SQSTM1 gene, or abnormal aggregation and phosphorylation of SQSTM1 lead to disorder of glucose and glutamine metabolism and promote tumor development in HCV-positive HCC through persistent activation of nuclear factor erythroid 2-related factor 2 (Nrf2) [35]. Moreover, animal experiments showed that SQSTM1 was necessary for hepatocarcinogenesis in mice [26]. BMF, one of the Bcl-2 family members, promotes apoptosis by inactivating pro-survival Bcl-2-like proteins through the BH3 domain following its activation

by stress signals [36]. Gramantieri et al. and Xie et al. [37, 38] revealed a close relationship between BMF and activated caspase-3. Moreover, they found that miR-221 inhibits apoptosis by targeting BMF in HCC and ovarian cancer cells. Besides, BMF inhibition promotes survival outcomes in multiple myeloma patients [39]. LGALS3 plays a significant role in the progression and metastasis of colon cancer, acute myeloid leukemia, melanoma, and pituitary tumors [40-43]. In addition, LGALS3 enhances HCC cell tumorigenesis and metastasis through the β -catenin signaling [44]. A series of studies have demonstrated that Galectin-3 can serve as a biomarker for prognosis predicting of HCC patients which are the same as our results [45-47]. PPP2R5B (B56 β) as the regulatory subunit of PP2A is involved in cell growth, survival, and metabolism [48]. It has been reported that deleted PPP2R5B gene induced sensitivity to paclitaxel in cervical cancer and this sensitivity change was supposed to be associated with apoptosis [49]. PPP2R5B mutations have been postulated to cause human overgrowth [50], though its role in HCC progression has not been elucidated.

In summary, we established a 5-gene risk model and constructed a nomogram for predicting HCC outcomes in clinical practice. Moreover, we present an advanced survival prediction tool for HCC patients, and reveal the association between ARGs and HCC that can further be confirmed by corresponding experimental studies.

Declarations

Acknowledgements

Not applicable.

Authors' contributions

RL and GW designed the study; RL collected data and wrote the manuscript; GW performed the data analysis and drew the figures as well as tables; RL, GW, CZ, and DB reviewed and revised the manuscript. Besides, the organization, revision, and submission of this manuscript were completed by DB. The final manuscript draft was read and approved by all authors.

Funding

This project was funded by The National Natural Science Foundation of China (No. 81871909), "13th five-year Plan" Science and Education strong Health Project leading personnel of Yangzhou (LJRC20181; YZCXTD201801), Provincial-level discipline leader of the NJPH (DTRC201809), Research Funds for Tian Qing Liver Diseases (TQGB20200180).

Availability of data and materials

The mRNA expression data and clinical information of HCC patients were obtained from the TCGA database (<https://cancergenome.nih.gov/>).

Ethical approval and consent participate

There were no cell, tissue, or animal studies. No ethical requirements are involved.

Consent for publication

Not applicable.

Competing interests

The authors declare that there is no conflict of interest.

References

1. Feng RM, Zong YN, Cao SM, Xu RH: Current cancer situation in China: good or bad news from the 2018 Global Cancer Statistics? *Cancer communications (London, England)* 2019, 39(1):22.
2. Ryu T, Takami Y, Wada Y, Hara T, Sasaki S, Saitsu H: Actual 10-Year Survival After Surgical Microwave Ablation for Hepatocellular Carcinoma: A Single-Center Experience in Japan. *Annals of surgical oncology* 2019, 26(12):4126-4133.
3. Zheng J, Kuk D, Gönen M, Balachandran VP, Kingham TP, Allen PJ, D'Angelica MI, Jarnagin WR, DeMatteo RP: Actual 10-Year Survivors After Resection of Hepatocellular Carcinoma. *Annals of surgical oncology* 2017, 24(5):1358-1366.
4. Wen T, Jin C, Facciorusso A, Donadon M, Han HS, Mao Y, Dai C, Cheng S, Zhang B, Peng B et al: Multidisciplinary management of recurrent and metastatic hepatocellular carcinoma after resection: an international expert consensus. *Hepatobiliary Surg Nutr* 2018, 7(5):353-371.
5. Tellapuri S, Sutphin PD, Beg MS, Singal AG, Kalva SP: Staging systems of hepatocellular carcinoma: A review. *Indian J Gastroenterol* 2018, 37(6):481-491.
6. Johnson PJ: Role of alpha-fetoprotein in the diagnosis and management of hepatocellular carcinoma. *J Gastroenterol Hepatol* 1999, 14 Suppl:S32-36.
7. Liu S, Yao X, Zhang D, Sheng J, Wen X, Wang Q, Chen G, Li Z, Du Z, Zhang X: Analysis of Transcription Factor-Related Regulatory Networks Based on Bioinformatics Analysis and Validation in Hepatocellular Carcinoma. *Biomed Res Int* 2018, 2018:1431396.
8. Zhou L, Du Y, Kong L, Zhang X, Chen Q: Identification of molecular target genes and key pathways in hepatocellular carcinoma by bioinformatics analysis. *OncoTargets and therapy* 2018, 11:1861-1869.
9. Meng Z, Wu J, Liu X, Zhou W, Ni M, Liu S, Guo S, Jia S, Zhang J: Identification of potential hub genes associated with the pathogenesis and prognosis of hepatocellular carcinoma via integrated bioinformatics analysis. *The Journal of international medical research* 2020, 48(7):300060520910019.
10. Kountouras J, Zavos C, Chatzopoulos D: Apoptosis in hepatocellular carcinoma. *Hepato-gastroenterology* 2003, 50(49):242-249.

11. Hou Y, Wang Z, Huang S, Sun C, Zhao J, Shi J, Li Z, Wang Z, He X, Tam NL et al: SKA3 Promotes tumor growth by regulating CDK2/P53 phosphorylation in hepatocellular carcinoma. *Cell death & disease* 2019, 10(12):929.
12. Li J, Sun RR, Yu ZJ, Liang H, Shen S, Kan Q: Galectin-1 Modulates the Survival and Tumor Necrosis Factor-Related Apoptosis-Inducing Ligand (TRAIL) Sensitivity in Human Hepatocellular Carcinoma Cells. *Cancer biotherapy & radiopharmaceuticals* 2015, 30(8):336-341.
13. Jin L, He Y, Tang S, Huang S: LncRNA GHET1 predicts poor prognosis in hepatocellular carcinoma and promotes cell proliferation by silencing KLF2. *J Cell Physiol* 2018, 233(6):4726-4734.
14. Zhang X, Lin C, Song J, Chen H, Chen X, Ren L, Zhou Z, Pan J, Yang Z, Bao W et al: Parkin facilitates proteasome inhibitor-induced apoptosis via suppression of NF- κ B activity in hepatocellular carcinoma. *Cell death & disease* 2019, 10(10):719.
15. Yu GP, Xiao QY, Shi ZQ, Tang LS, Ma XP, Zhang LY, Chen HT, Wang WJ, Zhang PY, Ding DL et al: Genetic polymorphisms in apoptosis-related genes and the prognosis of hepatocellular carcinoma. *American journal of cancer research* 2015, 5(10):3249-3259.
16. Subramanian A, Tamayo P, Mootha V, Mukherjee S, Ebert B, Gillette M, Paulovich A, Pomeroy S, Golub T, Lander E et al: Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* 2005, 102(43):15545-15550.
17. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P et al: STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019, 47(D1):D607-D613.
18. Shannon P: Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research* 2003, 13(11):2498-2504.
19. Hua S, Lei L, Deng L, Weng X, Liu C, Qi X, Wang S, Zhang D, Zou X, Cao C et al: miR-139-5p inhibits aerobic glycolysis, cell proliferation, migration, and invasion in hepatocellular carcinoma via a reciprocal regulatory interaction with ETS1. *Oncogene* 2018, 37(12):1624-1636.
20. Tong H, Liu X, Li T, Qiu W, Peng C, Shen B, Zhu Z: NR1D2 Accelerates Hepatocellular Carcinoma Progression by Driving the Epithelial-to-Mesenchymal Transition. *OncoTargets and therapy* 2020, 13:3931-3942.
21. Mou T, Zhu D, Wei X, Li T, Zheng D, Pu J, Guo Z, Wu Z: Identification and interaction analysis of key genes and microRNAs in hepatocellular carcinoma by bioinformatics analysis. *World journal of surgical oncology* 2017, 15(1):63.
22. Shi LE, Shang X, Nie KC, Xu Q, Chen NB, Zhu ZZ: Identification of potential crucial genes associated with the pathogenesis and prognosis of pancreatic adenocarcinoma. *Oncology letters* 2020, 20(4):60.
23. Guicciardi ME, Gores GJ: Apoptosis: a mechanism of acute and chronic liver injury. *Gut* 2005, 54(7):1024-1033.

24. Fabregat I, Roncero C, Fernández M: Survival and apoptosis: a dysregulated balance in liver cancer. *Liver Int* 2007, 27(2):155-162.
25. Lo SJ, Fan LC, Tsai YF, Lin KY, Huang HL, Wang TH, Liu H, Chen TC, Huang SF, Chang CJ et al: A novel interaction of nucleophosmin with BCL2-associated X protein regulating death evasion and drug sensitivity in human hepatoma cells. *Hepatology* 2013, 57(5):1893-1905.
26. Umemura A, He F, Taniguchi K, Nakagawa H, Yamachika S, Font-Burgada J, Zhong Z, Subramaniam S, Raghunandan S, Duran A et al: p62, Upregulated during Preneoplasia, Induces Hepatocellular Carcinogenesis by Maintaining Survival of Stressed HCC-Initiating Cells. *Cancer Cell* 2016, 29(6):935-948.
27. Wang Y, Xie BH, Lin WH, Huang YH, Ni JY, Hu J, Cui W, Zhou J, Shen L, Xu LF et al: Amplification of SMYD3 promotes tumorigenicity and intrahepatic metastasis of hepatocellular carcinoma via upregulation of CDK2 and MMP2. *Oncogene* 2019, 38(25):4948-4961.
28. Ghisoni E, Maggiorotto F, Borella F, Mittica G, Genta S, Giannone G, Katsaros D, Sciarrillo A, Ferrero A, Sarotto I et al: TOP2A as marker of response to pegylated liposomal doxorubicin (PLD) in epithelial ovarian cancers. *Journal of ovarian research* 2019, 12(1):17.
29. Zhang R, Xu J, Zhao J, Bai JH: Proliferation and invasion of colon cancer cells are suppressed by knockdown of TOP2A. *Journal of cellular biochemistry* 2018, 119(9):7256-7263.
30. Panvichian R, Tantiwettrueangdet A, Angkathunyakul N, Leelaudomlipi S: TOP2A amplification and overexpression in hepatocellular carcinoma tissues. *Biomed Res Int* 2015, 2015:381602.
31. Wong N, Yeo W, Wong WL, Wong NL, Chan KY, Mo FK, Koh J, Chan SL, Chan AT, Lai PB et al: TOP2A overexpression in hepatocellular carcinoma correlates with early age onset, shorter patients survival and chemoresistance. *Int J Cancer* 2009, 124(3):644-652.
32. Madrigal-Matute J, Cuervo AM: Regulation of Liver Metabolism by Autophagy. *Gastroenterology* 2016, 150(2):328-339.
33. Moscat J, Diaz-Meco MT: p62 at the crossroads of autophagy, apoptosis, and cancer. *Cell* 2009, 137(6):1001-1004.
34. Moscat J, Diaz-Meco MT: Feedback on fat: p62-mTORC1-autophagy connections. *Cell* 2011, 147(4):724-727.
35. Saito T, Ichimura Y, Taguchi K, Suzuki T, Mizushima T, Takagi K, Hirose Y, Nagahashi M, Iso T, Fukutomi T et al: p62/Sqstm1 promotes malignancy of HCV-positive hepatocellular carcinoma through Nrf2-dependent metabolic reprogramming. *Nat Commun* 2016, 7:12030.
36. Willis SN, Adams JM: Life in the balance: how BH3-only proteins induce apoptosis. *Current opinion in cell biology* 2005, 17(6):617-625.
37. Gramantieri L, Fornari F, Ferracin M, Veronese A, Sabbioni S, Calin GA, Grazi GL, Croce CM, Bolondi L, Negrini M: MicroRNA-221 targets Bmf in hepatocellular carcinoma and correlates with tumor multifocality. *Clinical cancer research: an official journal of the American Association for Cancer Research* 2009, 15(16):5073-5081.

38. Xie X, Huang Y, Chen L, Wang J: miR-221 regulates proliferation and apoptosis of ovarian cancer cells by targeting BMF. *Oncology letters* 2018, 16(5):6697-6704.
39. Fedele P, Liao Y, Gong J, Yao Y, van Delft M, Low M, Tai L, Herold M, Jackson J, Teh C et al: The transcription factor IRF4 represses proapoptotic BMF and BIM to licence multiple myeloma survival. *Leukemia* 2020.
40. Han L, Wu Z, Zhao Q: Revealing the molecular mechanism of colorectal cancer by establishing LGALS3-related protein-protein interaction network and identifying signaling pathways. *International journal of molecular medicine* 2014, 33(3):581-588.
41. Cheng CL, Hou HA, Lee MC, Liu CY, Jhuang JY, Lai YJ, Lin CW, Chen HY, Liu FT, Chou WC et al: Higher bone marrow LGALS3 expression is an independent unfavorable prognostic factor for overall survival in patients with acute myeloid leukemia. *Blood* 2013, 121(16):3172-3180.
42. Braeuer RR, Zigler M, Kamiya T, Dobroff AS, Huang L, Choi W, McConkey DJ, Shoshan E, Mobley AK, Song R et al: Galectin-3 contributes to melanoma growth and metastasis via regulation of NFAT1 and autotaxin. *Cancer research* 2012, 72(22):5757-5766.
43. Ruebel KH, Jin L, Qian X, Scheithauer BW, Kovacs K, Nakamura N, Zhang H, Raz A, Lloyd RV: Effects of DNA methylation on galectin-3 expression in pituitary tumors. *Cancer research* 2005, 65(4):1136-1140.
44. Song M, Pan Q, Yang J, He J, Zeng J, Cheng S, Huang Y, Zhou ZQ, Zhu Q, Yang C et al: Galectin-3 favours tumour metastasis via the activation of β -catenin signalling in hepatocellular carcinoma. *Br J Cancer* 2020.
45. Kong F, Jin M, Cao D, Jia Z, Liu Y, Jiang J: Galectin-3 not Galectin-9 as a candidate prognosis marker for hepatocellular carcinoma. *PeerJ* 2020, 8:e9949.
46. Zhang FP, Huang YP, Luo WX, Deng WY, Liu CQ, Xu LB, Liu C: Construction of a risk score prognosis model based on hepatocellular carcinoma microenvironment. *World J Gastroenterol* 2020, 26(2):134-153.
47. Wang X, Yuegao, Bai L, Ibrahim MM, Ma W, Zhang J, Huang Y, Wang B, Song L, Tang J: Evaluation of Annexin A7, Galectin-3 and Gelsolin as possible biomarkers of hepatocarcinoma lymphatic metastasis. *Biomed Pharmacother* 2014, 68(3):259-265.
48. Seshacharyulu P, Pandey P, Datta K, Batra SK: Phosphatase: PP2A structural importance, regulation and its aberrant expression in cancer. *Cancer letters* 2013, 335(1):9-18.
49. Wei W, He Y, Wu Y: Identification of genes associated with SiHa cell sensitivity to paclitaxel by CRISPR-Cas9 knockout screening. *International journal of clinical and experimental pathology* 2018, 11(4):1972-1978.
50. Loveday C, Tatton-Brown K, Clarke M, Westwood I, Renwick A, Ramsay E, Nemeth A, Campbell J, Joss S, Gardner M et al: Mutations in the PP2A regulatory subunit B family genes PPP2R5B, PPP2R5C and PPP2R5D cause human overgrowth. *Hum Mol Genet* 2015, 24(17):4775-4779.

Tables

Table 1
Multivariate Cox regression results of prognosis-related ARGs in HCC.

Gene id	Coefficient	HR	HR.95L	HR.95H	P value
PPP2R5B	0.3853266	1.4700944	1.0225253	2.1135688	0.0375054
SQSTM1	0.2786996	1.3214104	1.1163784	1.5640981	0.0011966
TOP2A	0.1520620	1.1642324	1.0033461	1.3509168	0.0450713
BMF	0.1721765	1.1878875	0.9781174	1.4426455	0.0824262
ARGs, apoptosis-related genes; HCC, hepatocellular carcinoma.					

Figures

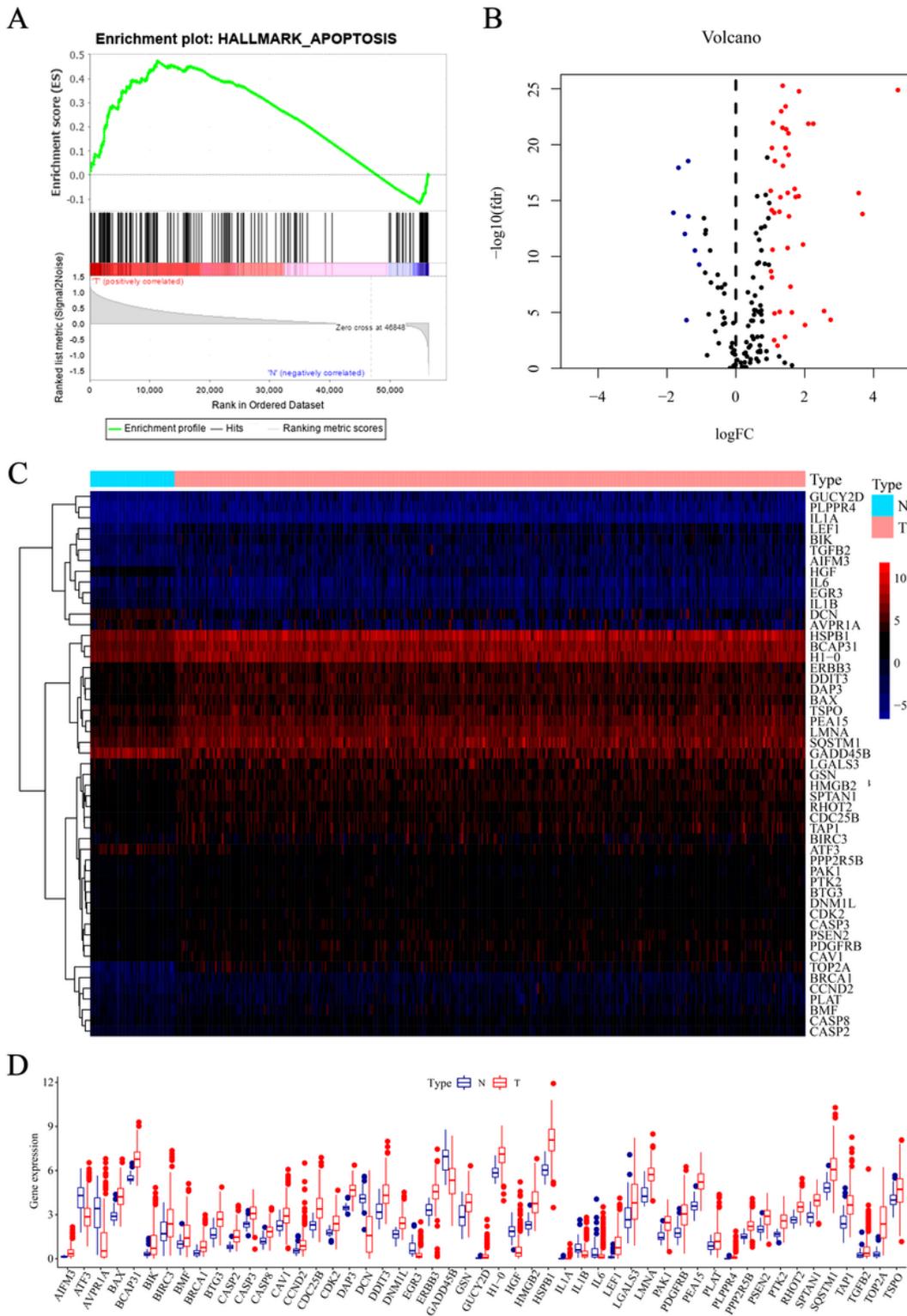


Figure 1

Identification of differentially expressed 51 apoptosis-related genes (ARGs) in HCC. a GSE analysis of 161 ARGs. b Volcano plot of differentially expressed ARGs. Red represents high expression, blue represents low expression, black represents no difference between HCC and normal tissues. c The heatmap of 51 identified ARGs. d The boxplot of 51 identified ARGs. Red represents HCC tissues while

blue represents normal tissues, respectively. ARGs; apoptosis-related genes. HCC; hepatocellular carcinoma.

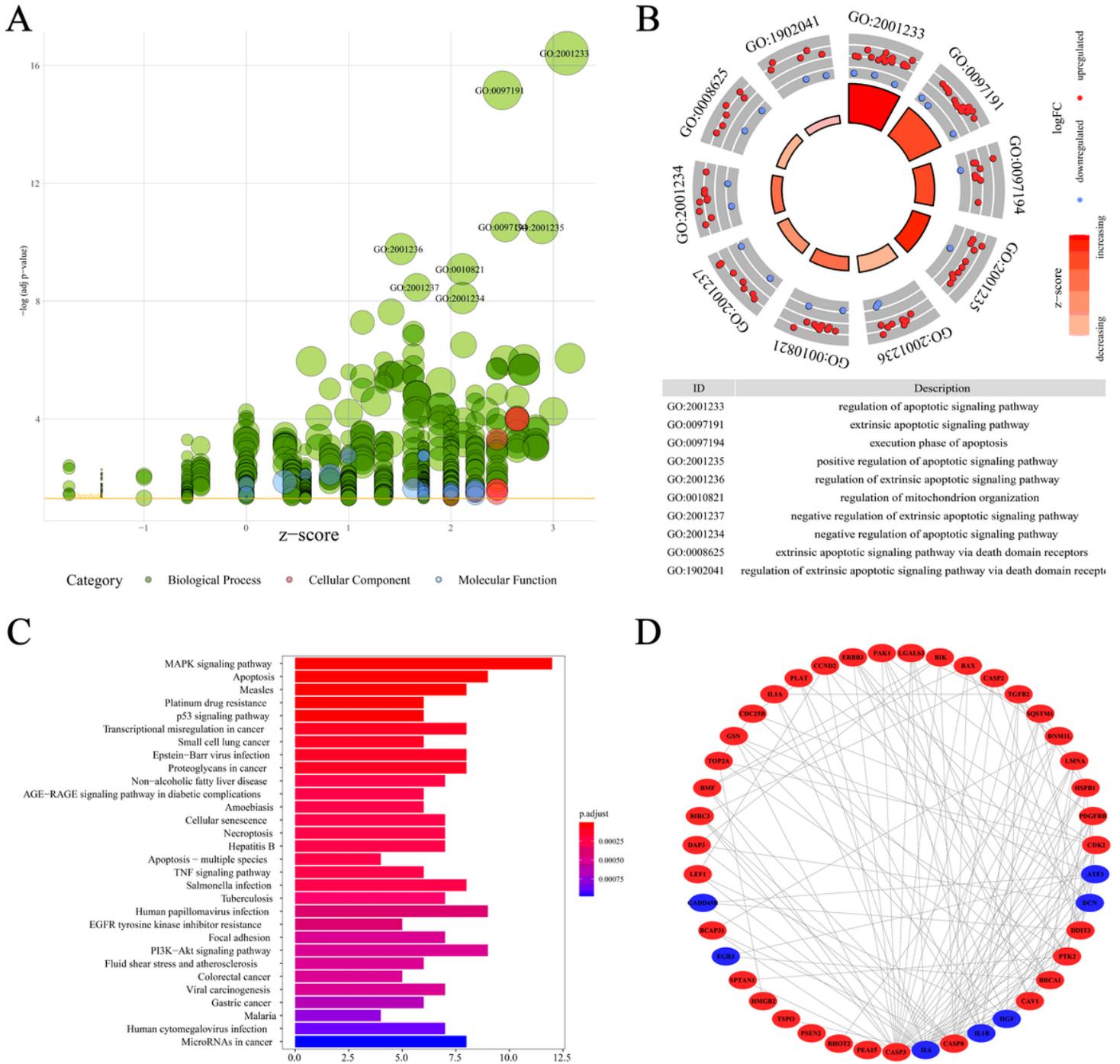


Figure 2

GO enrichment, KEGG pathway and PPI analysis of the 51 identified ARGs. Findings of the GO analysis are presented in the bubble chart (a) and the circle plot (b). c Results from KEGG analysis are presented in the bar plot. d The PPI network of the 51 identified ARGs.

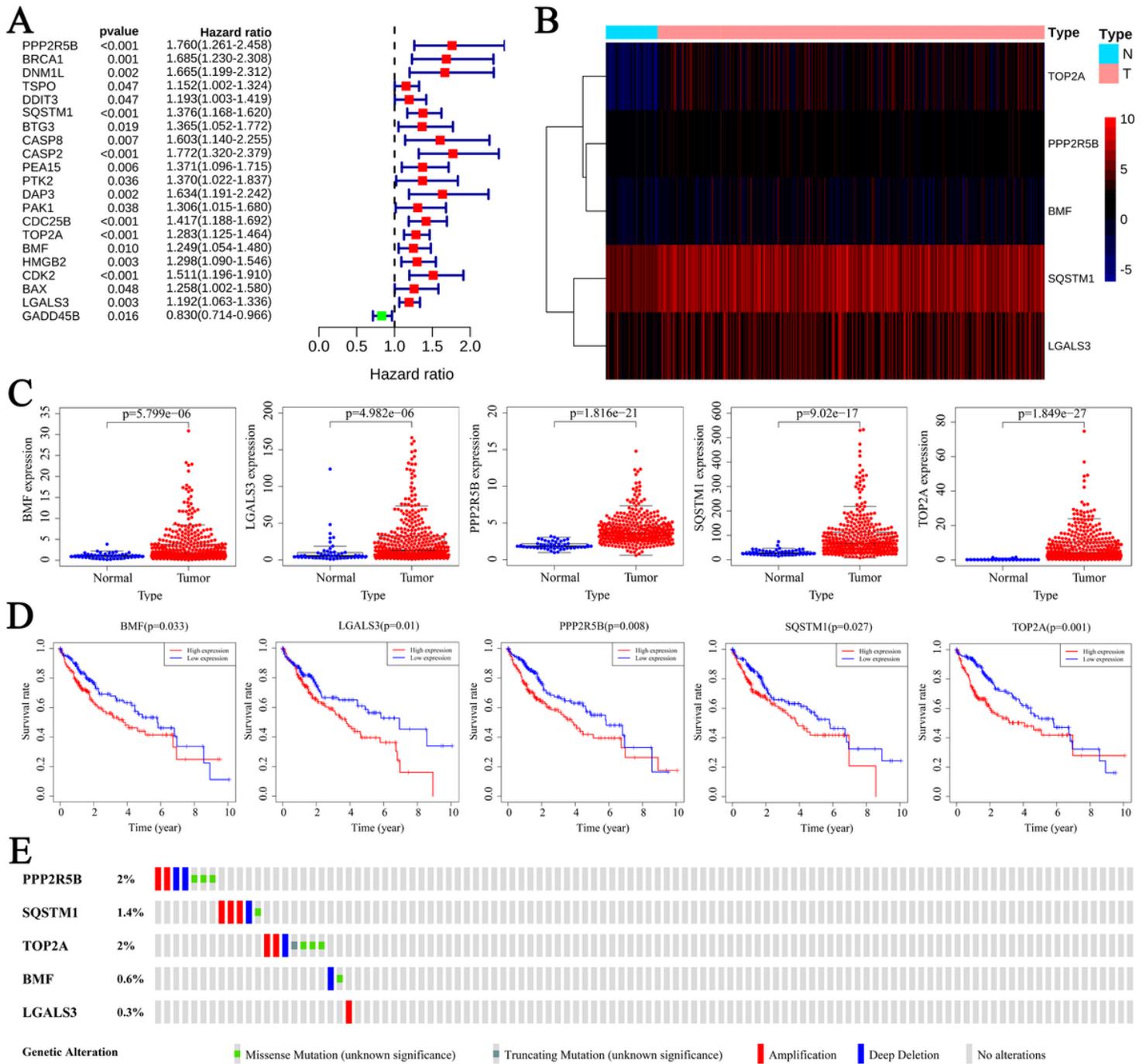


Figure 3

Identification of ARGs associated with the prognosis of HCC. a The forest plot of univariate Cox regression results. b The heatmap of 5 identified prognostic ARGs screened out by multivariate Cox regression. c Expression levels of 5 screened ARGs in HCC and normal samples. d K-M curve of the relationship between OS in HCC patients and expression levels of 5 screened ARGs. e Mutation data of 5 screened ARGs among 353 HCC specimens according to the cBioPortal database. OS: overall survival rates.

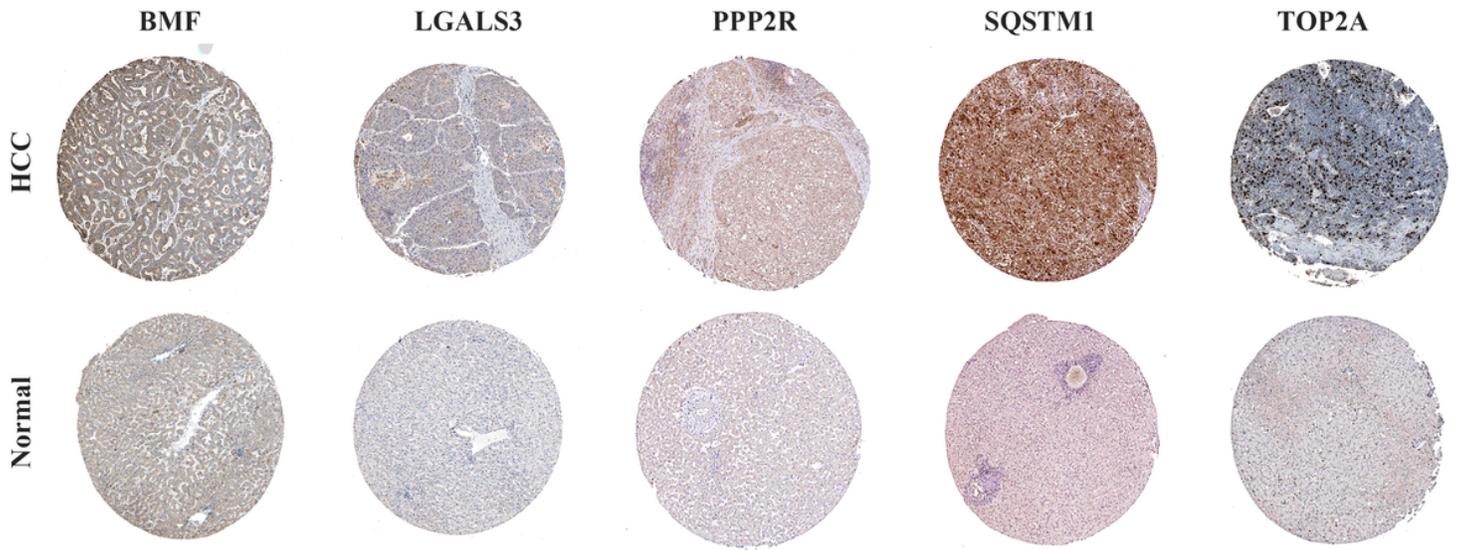


Figure 4

Boxplot and immunohistochemical results of the expression levels of 5 screened ARGs between HCC and para-carcinoma tissues according to the Human Protein Atlas (HPA) database.

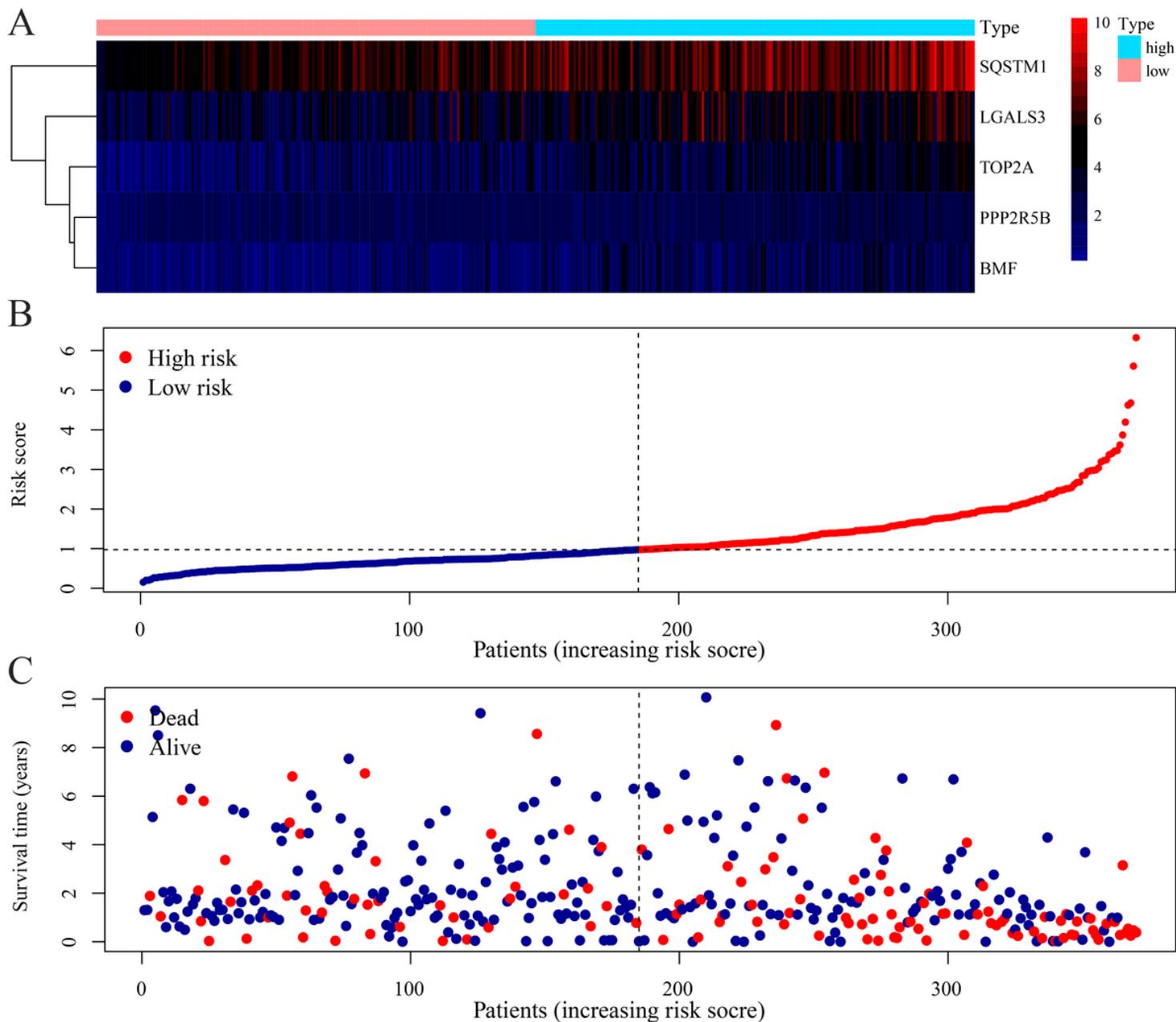


Figure 5

OS prediction in HCC patients by using the risk score which was established by 5 screened ARGs. a Expression heatmap of 5 prognostic ARGs in high and low-risk groups. b Risk score distribution in HCC patients. c Overall survival and survival status of HCC patients in the TCGA database.

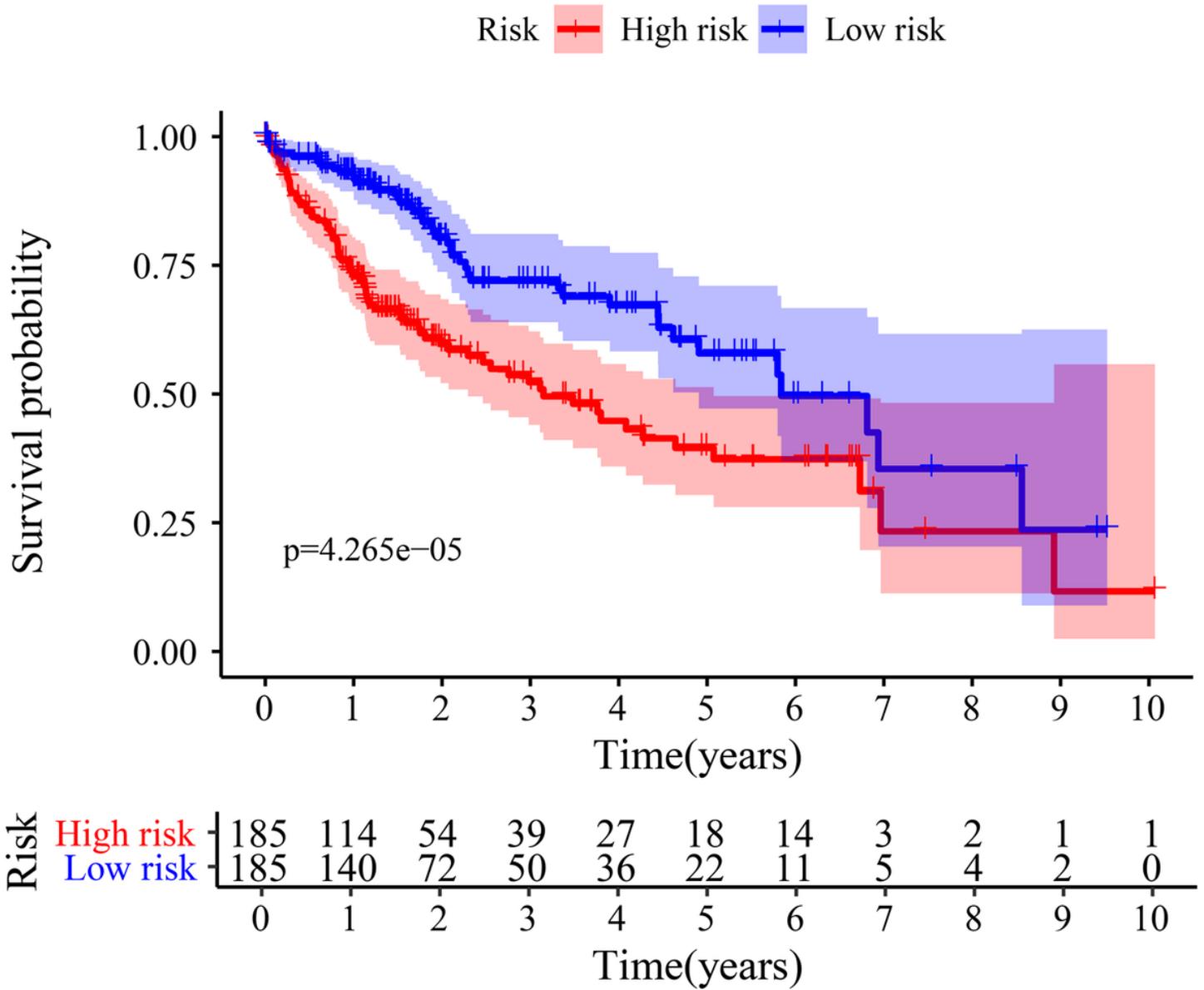


Figure 6

Compared to the high-risk groups, patients in low-risk groups have significantly longer OS outcomes.

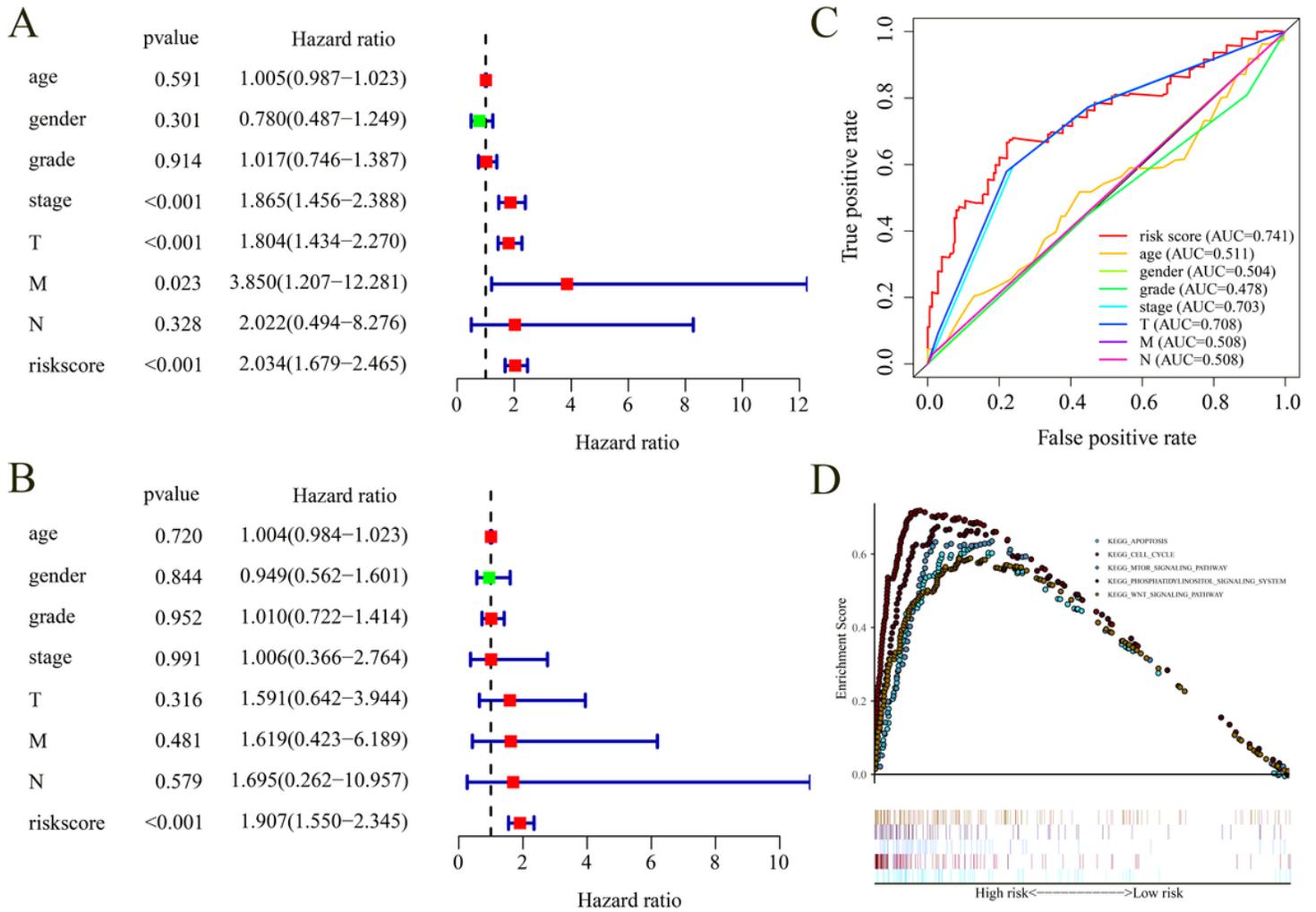


Figure 7

a Prognostic effect analysis of risk score and clinical features in HCC with univariate Cox regression analysis. b Independent prognostic effect analysis of risk score and traditional prognostic clinical features in HCC with multivariate Cox regression analysis. c The ROC analysis of the risk score and other prognostic clinical features in HCC. d Pathways that the 5 screened ARGs are enriched in according to KEGG analysis.

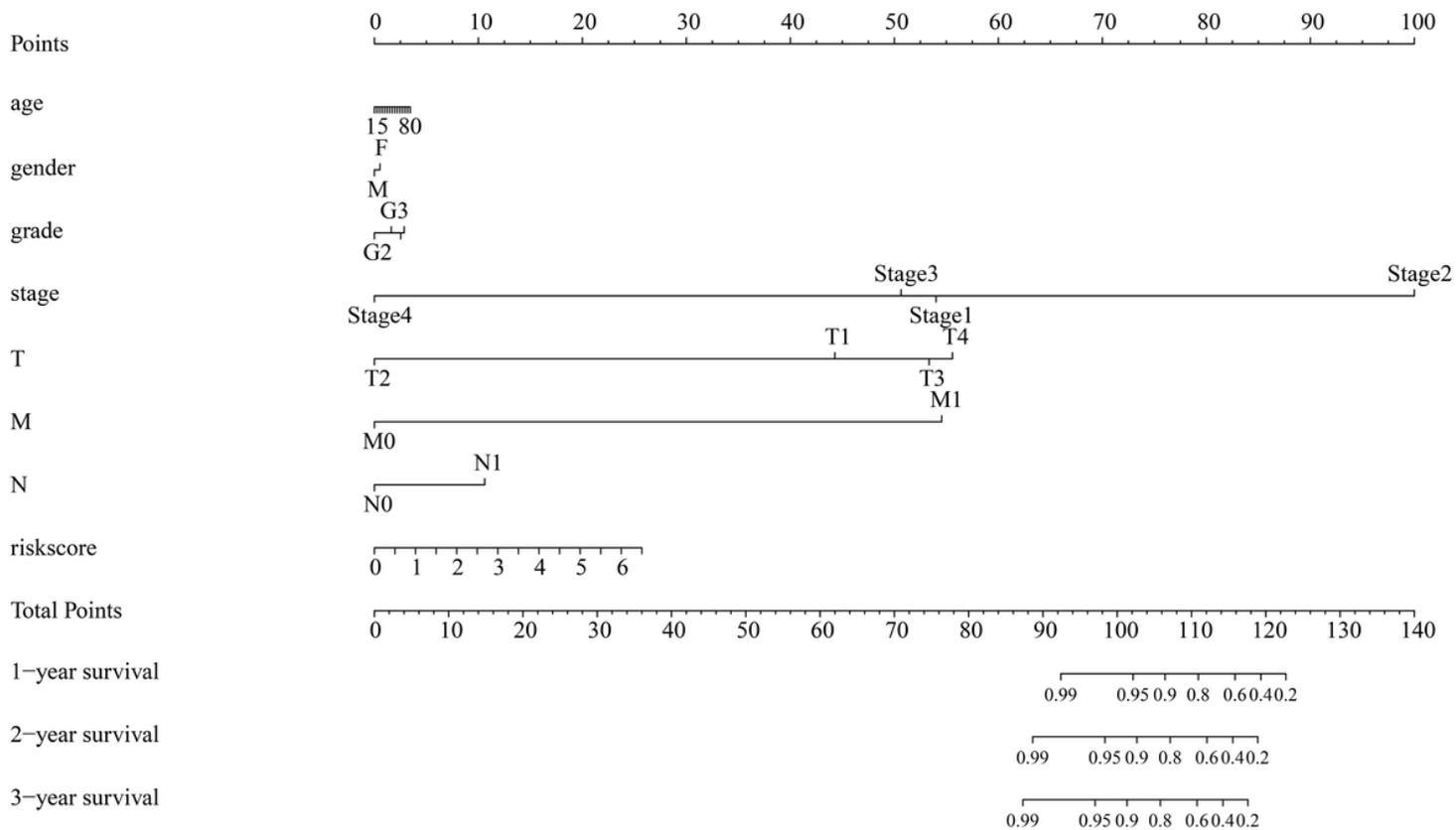


Figure 8

Prediction of the 1-, 2-, and 3- year OS of HCC patients with the nomogram.

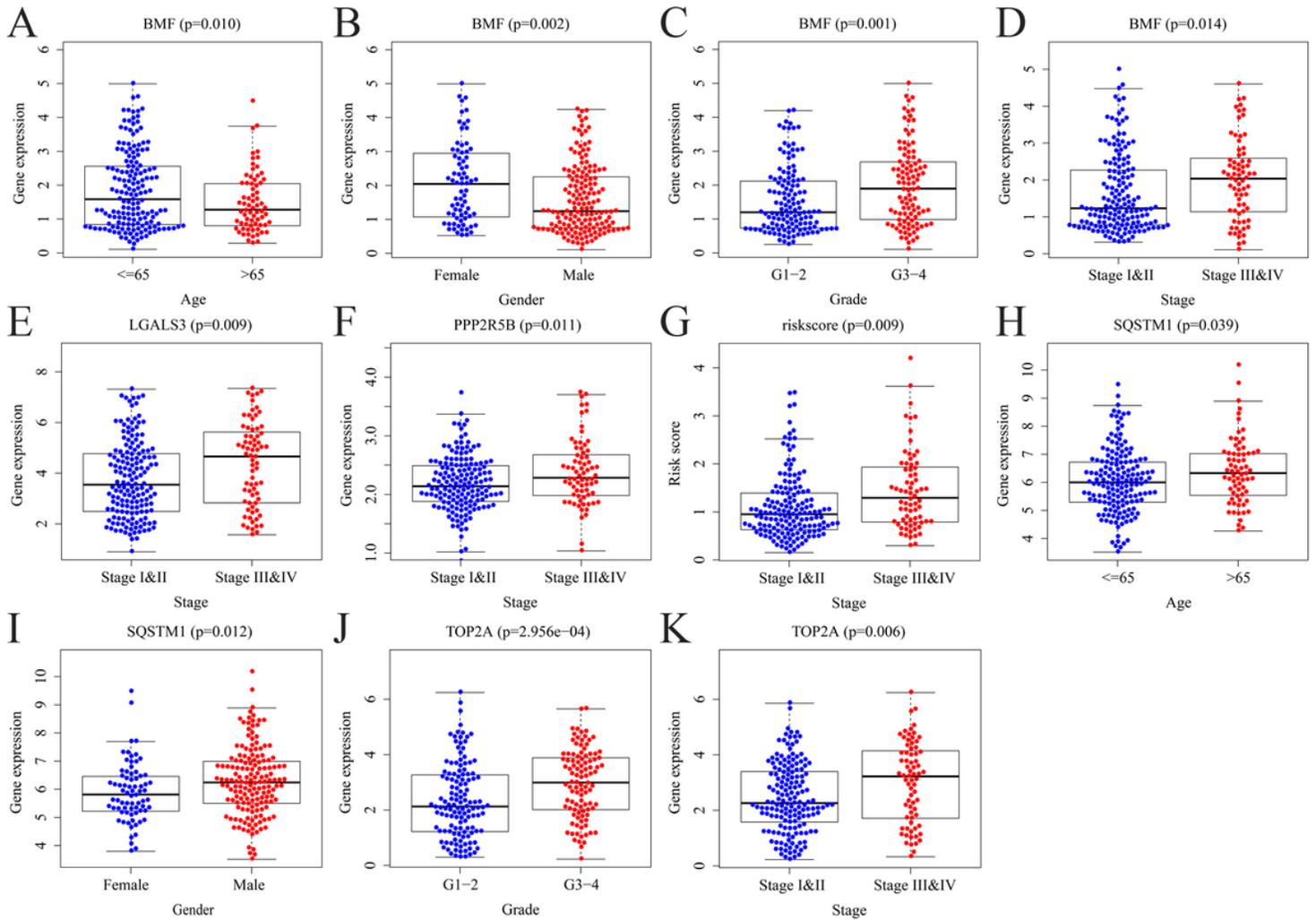


Figure 9

Associations between the 5 significant ARGs and clinical features. a-d Association between BMF expression level and age, gender, grade, and stage. e Association between LGALS3 expression level and pathology stage. f Association between PPP2R5B expression level and pathology stage. g Risk score and age. h-i Association between SQSTM1 expression level, age and gender. j-k Association between TOP2A expression level, grade and pathology stage.

Supplementary Files

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

- [SupplementaryTable1.docx](#)