

Identification of hepatocellular carcinoma risk using a novel prognostic model 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, Prognosis model, TCGA, Nomogram

Posted Date: January 8th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-141169/v1>

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 in human hepatocarcinogenesis. There is a possible association of apoptosis dysregulation with poor prognosis in many malignant tumors, such as hepatocellular carcinoma (HCC). However, the prognostic effect of Apoptosis-related genes (ARGs) on HCC is still unclear.

Methods

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

Results

Compared to normal specimens, 43 highly upregulated and 8 downregulated ARGs respectively and their normal counterparts in HCC specimens were screened. KEGG analysis demonstrated pathways correlated with these 51 genes which included MAPK, P53, TNF, PI3K-Akt signaling pathways. With Cox regression, 5 prognostic correlated with ARGs (PPP2R5B, SQSTM1, TOP2A, BMF, and LGALS3) were obtained to develop the prognosis model. According to the median of risk scores, patients were categorized into high-risk and low-risk groups. Patients in low-risk groups had a significantly higher two-year or five-year survival probability ($p < 0.0001$). The risk model had better potency than other clinical characteristics, with the area under the ROC curve (AUC = 0.741). Prognosis of HCC patients was established from a plotted nomogram.

Conclusion

This present study established a novel prognostic risk model for predicting HCC according to the expression of ARGs. The present advancement can potentially contribute to prediction prognosis and individualized treatment of HCC patients.

Introduction

Liver cancer has been ranked sixth among the most common tumors worldwide and the fourth leading cause of cancer-related mortality [1]. Among the frequent primary liver cancer, hepatocellular carcinoma (HCC) accounts for approximately 75% cases [1]. Despite the advancements in diagnosing, techniques and treatment of HCC, the survival of the patients worsens following the high rate of recurrence and metastasis [2–4]. TNM staging system is a traditional method for prognostic prediction, lacks performance accuracy in its clinical application due to substantive variations in HCC clinical outcomes [5]. Over the past decades, serum alpha-fetoprotein (AFP) has been the only biomarker available for detecting and predicting prognosis, through with low sensitivity limits its clinical utility [6]. Therefore, identification of a novel prognostic biomarker and an advanced prognostic model for HCC patients is of paramount importance.

Bioinformatics analysis had been extensively adopted over the past few years as an effective tool in exploring functions of numerous differentially expressed genes (DEGs) and determining the complexity of HCC occurrence and development [7, 8]. Hub genes and pathways associated with pathogenesis and prognosis of HCC via a series of bioinformatics analyses were established by Meng et al [9]. However, these studies usually ignore clinical information such as sex, age, grade, and stage of tumors. It may be very innovative to construct a prognostic model combining patient's gene expression level and clinical information. Hepatocarcinogenesis develops following the imbalance between proliferation and apoptosis [10]. Results of a recent study had revealed that the spindle and kinetochore-related complex subunit 3 (SKAT3) overexpressed in HCC potentially inhibits p53 activation by binding to cyclin-dependent kinase 2 (CDK2), before impeding cell apoptosis and promoting cancer cell proliferation [11]. Besides, several biomolecules may influence HCC prognosis by regulating apoptosis-related genes or apoptosis-related pathways [12–14]. Association of the haplotype of the apoptosis-related gene CDKN1B with the prognosis of HCC patients who underwent surgical resection were reported by Yu et al [15]. CCT/ACT haplotype patients had lower overall survival rates than those with common ACT/CCT haplotype. Thus, ARGs can potentially assess HCC prognosis.

Herein, based on the gene expression and clinical characteristics data obtained from the Cancer Genome Atlas (TCGA) database, we determined the ARG related to the prognosis and develop a prognostic prediction model. The model potentially calculates the risk score to predict and evaluate the prognosis of HCC patients.

Materials And Methods

Data collection

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

Gene Set Enrichment Analysis and Differentially Expressed ARGs

Gene set with prominent differences between in HCC and normal samples were explored by GSEA. Subsequently, the limma package and the Wilcoxon signed-rank test in R software 3.6.2 ($|\logFC| > 1$, $FDR < 0.05$) were used to show the significantly differentiated expressed ARGs in mRNA expression profiles of 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 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were performed to explore the potential biological functions of ARG participation before visualized through R software with packages such as ggplot2, DOSE, clusterProfiler, enrichplot, GOplot, digest, etc. Interactions among the selected ARGs were determined through protein-protein interaction (PPI) network from the STRING database (<http://www.string-db.org/>) and visualized by Cytoscape software.

Establishment of Prognostic risk model based on ARGs

Univariate and multivariate Cox regression analyses were conducted to identify the prognosis-associated ARGs in HCC. A prognosis-associated prediction formula acquired from multivariate Cox regression was used to construct a prognostic model using package “glmnet” in R. According to the prognostic model, the Kaplan-Meier (K-M) method was performed to analyze the survival rate of the high- and the low-risk groups. Subsequently, we use the area under the ROC curve (AUC), and KEGG enrichment analysis 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, the statistical analyses and plots were performed and constructed using R software 3.6.2 (<https://www.r-project.org/>) and Perl language packages. Respectively, p -value < 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 was identified from the gene set “HALLMARK_APOPTOSIS”. Thereafter, the 161 ARGs in this gene set was further assessed using GSE analysis to ascertain the presence of biological significance in HCC. Results in (Fig. 1a) showed this ARG set was significantly differentially expressed between HCC and normal specimens.

Identification of differentiated expressed ARGs

The mRNA sequence data of 374 cases of HCC and 50 cases of normal tissue were obtained from TCGA. The limma package and the Wilcoxon signed-rank test in R ($|\log\text{FC}| > 1$, $\text{FDR} < 0.05$) were used to screen the differentiated expressed ARGs in HCC and non-tumor samples. In this study, 43 and 8 ARGs were significantly upregulated and downregulated, respectively. In addition, their results were plotted by volcano plot, heatmap, and box plot (Fig. 1(b-d)).

Functional enrichment and PPI network analysis of differentiated expressed ARGs

To further study the main biological functions and significant pathways of these 51 identified genes, the GO enrichment and KEGG pathway enrichment analysis were performed. The plotted results are shown in (Fig. 2). The 51 identified ARGs were involved in the pathways related to cellular apoptosis (Fig. 2a and 2b). According to KEGG enrichment analysis results, the 51 ARGs possible contribute to platinum drug resistance and transcriptional misregulation in cancer as well as associated with some oncogenic pathways such as MAPK, P53, TNF, PI3K-Akt signaling pathways (Fig. 2c). The STRING online tool was used to establish a PPI network to explore the interactions among the proteins coded by ARGs. Results observed using the Cytoscape software (Fig. 2d).

Association of ARGs with HCC patient's survival and prognosis

Univariate Cox regression analyses were performed on both mRNA expressional and corresponding clinical data of 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 identified genes were subsequently analyzed by multivariate Cox regression ($p < 0.05$) to determine ARGs association with the prognosis of patients and acquire corresponding regression coefficients. Following the multivariate Cox regression, five prognosis-related ARGs were screened which contained PPP2R5B, SQSTM1, TOP2A, BMF, and LGALS3. Expression levels in normal and HCC specimens were further compared to establish the prognostic value of these 5 genes. The 5 identified ARGs in HCC specimens had significantly higher expression levels than normal specimens as indicated by heatmap and boxplot (Fig. 3b and 3c). Besides, the K-M curve was constructed from the comparisons of the survival rates differences in high- and low-expressed groups in the identified ARGs. Interestingly, the high expression levels in the 5 identified ARGs indicated a low survival rate (Fig. 3d).

On the other hand, mutations of these 5 genes in HCC were analyzed by the cBioPortal database (<http://cbioportal.org>). Among the data of 353 HCC patients in this database, 22 patients (6.3%) had mutations. Notably, among the 22 patients with the mutation, 0.85% had missense mutations, 0.56% had amplifications, 0.56% had deep deletions in PPP2R5B. However, in the SQSTM1, 0.28% had missense mutations, 0.85% had amplifications and 0.28% had deep deletions whereas in TOP2A, 0.85% had

missense mutations, 0.28% had truncating mutations, 0.56% had amplifications and 0.28% had deep deletions. Moreover, in BMF, 0.28% had missense mutations, and 0.28% had deep deletions whereas 0.28% had amplifications in LGALS3 (Fig. 3e).

Simultaneously, we revealed that these 5 genes were significantly upregulated in HCC from the immunohistochemical results of the Human Protein Atlas (HPA) database (Fig. 4).

Establishment of prognostic risk signature based on ARGs

We combined the expression level of ARGs and the regression coefficients of multiple Cox regression analysis to summarize the risk scoring formula. The 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 indicating that the high expressions of identified genes were negatively associated with prognosis. The risk score in HCC patients was calculated by this formula, and these patients were divided into high- and low-risk groups based on the median risk score as a cutoff (Fig. 5). The plotted heatmap exhibit the 5 ARGs expression levels which showed that patients in the same group had distinct expression levels (Fig. 5a). Patients scores were ranked in ascending order (Fig. 5b) and their survival time displayed in the scatterplot (Fig. 5c) revealed that the low-risk patients had a longer survival time and higher survival rate than those with high-risk.

Survival analysis of high- and low-risk groups were performed to further illustrate the correlation between the risk score and the prognosis of patients. First, we demonstrated the prognostic significance of the risk score using K-M curves and the log-rank method as shown in (Fig. 6). Patients in low-risk groups had a significantly high two-year or five-year survival probability ($p < 0.0001$; Fig. 6) compared with high-risk groups, revealing a worse prognosis.

Independent prediction of HCC prognosis by the risk score

Both univariate and multivariate Cox regression analyses were used to compare the predictive value of the prognostic model and other clinical variables on prognosis. The risk score was established as an independent prognosis predictor in patients with HCC (Fig. 7a and 7b). On the contrary, clinical variables were not independent prognosis predictor as shown in the result of multivariate cox regression analysis ($p > 0.05$; Fig. 7b). Moreover, the ROC curve was plotted to validate the predictive value of the risk score (Fig. 7c). The curves demonstrated that the AUC of the risk score was 0.741 as the highest value than that of the clinical variables. The results suggested that the risk score had a higher predictive value than the clinical variables.

KEGG enrichment analysis between different risk patients

The KEGG analysis was utilized to explore the potential biological pathways which were associated with 5 ARGs. The results showed that the major 5 correlated pathways were apoptosis, cell cycle, mTOR

signaling pathway, WNT signaling pathway, and phosphatidylinositol signaling system (Fig. 7d).

Construction of a nomogram to predict the prognosis of HCC patients

To better apply the risk score to predict 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. Given 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 among the HCC patients (Fig. 8).

The correlations between the risk score/ 5 ARGs and clinical variables

Based on the gene expression and corresponding clinical data acquired from the TCGA database, we analyzed the correlations between these clinical variables and the risk score formed by 5 ARGs. The results indicated that the risk score was correlated with the stage; BMF was associated with age, gender, grade, and stage; PPP2R5B and LGALS3 were related to the stage; SQSTM1 was associated with age and gender; TOP2A was connected with grade and stage (Fig. 9).

Discussion

HCC is the most common type of malignancy with a high recurrence rate and worse prognosis [16, 17]. Studies have shown limitations in achieving early diagnosis and treatment of HCC [18]. Therefore, the identification of a prognosis-related biomarker and the development of a prognostic model is of critical importance for HCC patients. Currently, bioinformatics analysis plays an increasingly significant role in identifying therapeutic targets for diagnosis, treatment, and prognosis of numerous tumors [19]. The process of apoptosis plays a crucial role in liver tumor development and regeneration [12, 20]. Insufficient apoptosis potentially leads to occurrence and development of liver tumors [21].

In this work, we developed a novel prognosis-predictive model for HCC based on the expression of ARGs. 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 specimens compared with normal specimens were screened out. These 51 genes mainly participated in the pathways associated with cellular apoptosis according to GO enrichment analysis. KEGG analyzes revealed pathways correlated with these 51 genes which included MAPK, P53, TNF, PI3K-Akt signaling pathways. Besides, five prognosis-related ARGs were screened after the multivariate Cox regression which contained PPP2R5B, SQSTM1, TOP2A, BMF, and LGALS3. The high expressions having these five ARGs were negatively associated with prognosis. Based on this, we designed a risk model that considered an independent prognostic model of HCC with multivariate regression analysis. The predictive value of the model was confirmed by performing the K-M curve and ROC curve that showed the highest prognostic predictive power of the model with other clinical features of HCC patients. Therefore, by combining the risk score

and clinical features as the basis of the nomogram, facilitate efficiency in predicting the prognosis of HCC patients.

The potential impact of these five hub ARGs on the progression of HCC were earlier observed. For example, Topoisomerase II alpha (TOP2A), was identified as a potential biomarker for cancer therapy in ovarian cancer, colon cancer, pancreatic adenocarcinoma and HCC [9, 19, 22, 23]. However, TOP2A was up-regulated in HCC [24]. Studies have proved that overexpressed TOP2A in HCC leads to a worse prognosis and its reactive agents has a potential therapeutic effect on HCC patients [25]. SQSTM1, an autophagy-related protein, degrades during autophagy. Moreover, autophagy was demonstrated to suppress spontaneous tumorigenesis in the liver [26]. Besides, BMF, one of the Bcl-2 family members, promotes apoptosis by inactivating pro-survival Bcl-2-like proteins via BH3 domain following its activation by stress signals [27]. Laura and Xinping et al [28, 29] revealed a close relationship between BMF and activated caspase-3as well as apoptosis inhibition using miR-221inhibits by targeting BMF in HCC and ovarian cancer cells. Besides, BMF inhibition promotes survival of multiple myeloma [30]. LGALS3 plays a significant role in the progression and metastasis of colon cancer, acute myeloid leukemia, melanoma and pituitary tumors [31–34]. Recent research had indicated that LGALS3 enhanced the tumorigenesis and metastasis of HCC cells via β -catenin signaling [35]. PPP2R5B (B56 β) as the regulatory subunit of PP2A can influence cell growth, survival, and metabolism [36]. PPP2R5B mutation is hypothesized to cause human overgrowth [37], though its role in HCC had remained elusive.

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

Declarations

Availability of data and materials

All the data and materials are available.

Ethics approval and consent participate

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

Consent for publication

Not applicable.

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' contributions

RL and GW designed the study; RL collected data and wrote the manuscript; GW performed the data analysis and drew figures and tables; RL, GW, and DB reviewed and revised the manuscript. Besides, the organization, revision and submission of this manuscript were completed by DB. Final draft were 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).

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. 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.
17. 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.
18. 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.
19. 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.
20. Guicciardi ME, Gores GJ: Apoptosis: a mechanism of acute and chronic liver injury. *Gut* 2005, 54(7):1024-1033.
21. Fabregat I, Roncero C, Fernández M: Survival and apoptosis: a dysregulated balance in liver cancer. *Liver Int* 2007, 27(2):155-162.
22. 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.
23. 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.

24. Panvichian R, Tantiwetrueangdet A, Angkathunyakul N, Leelaudomlapi S: TOP2A amplification and overexpression in hepatocellular carcinoma tissues. *Biomed Res Int* 2015, 2015:381602.
25. 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.
26. Madrigal-Matute J, Cuervo AM: Regulation of Liver Metabolism by Autophagy. *Gastroenterology* 2016, 150(2):328-339.
27. Willis SN, Adams JM: Life in the balance: how BH3-only proteins induce apoptosis. *Current opinion in cell biology* 2005, 17(6):617-625.
28. 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.
29. 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.
30. 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.
31. 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.
32. 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.
33. 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.
34. 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.
35. 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.
36. 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.
37. 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

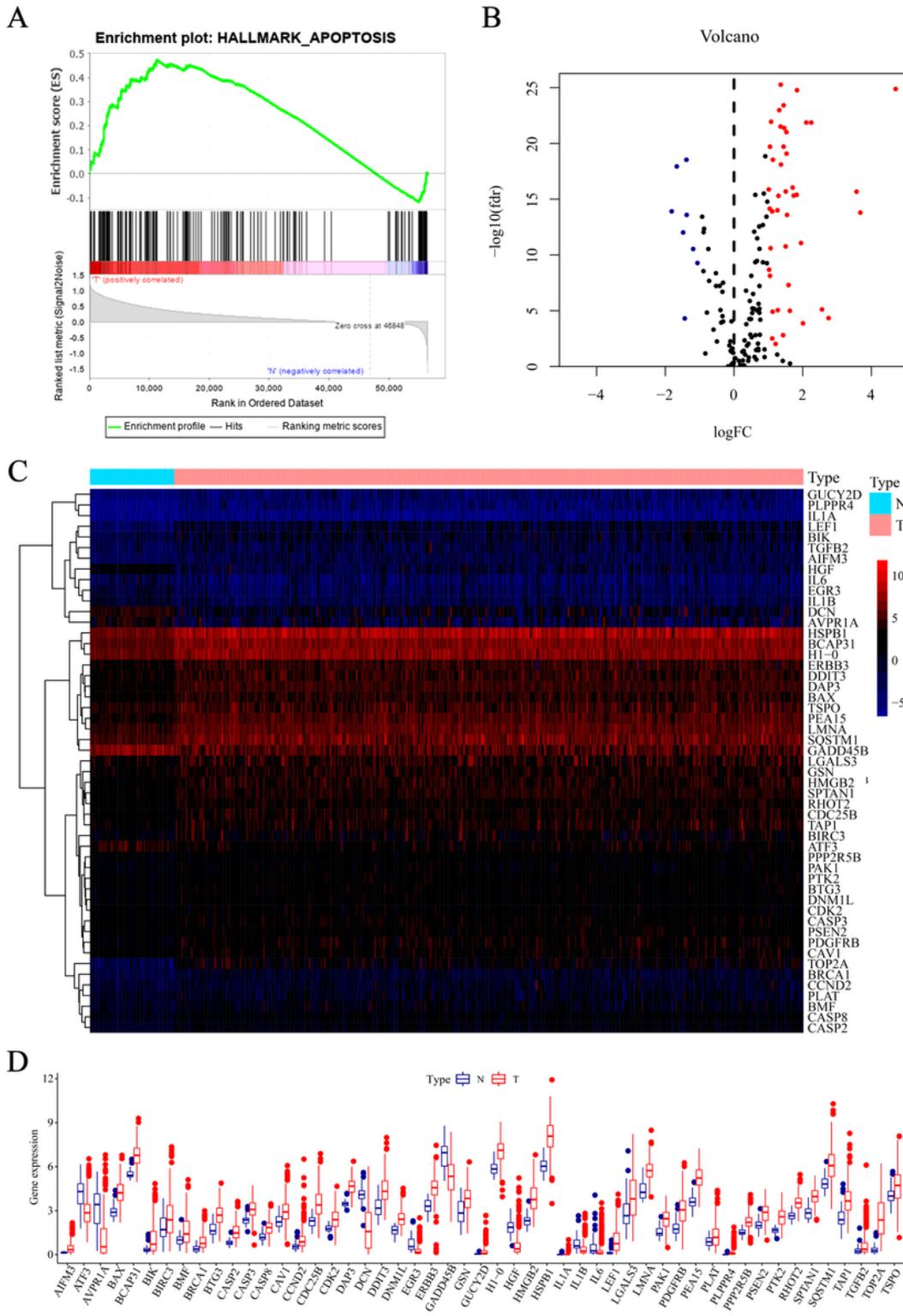


Figure 1

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

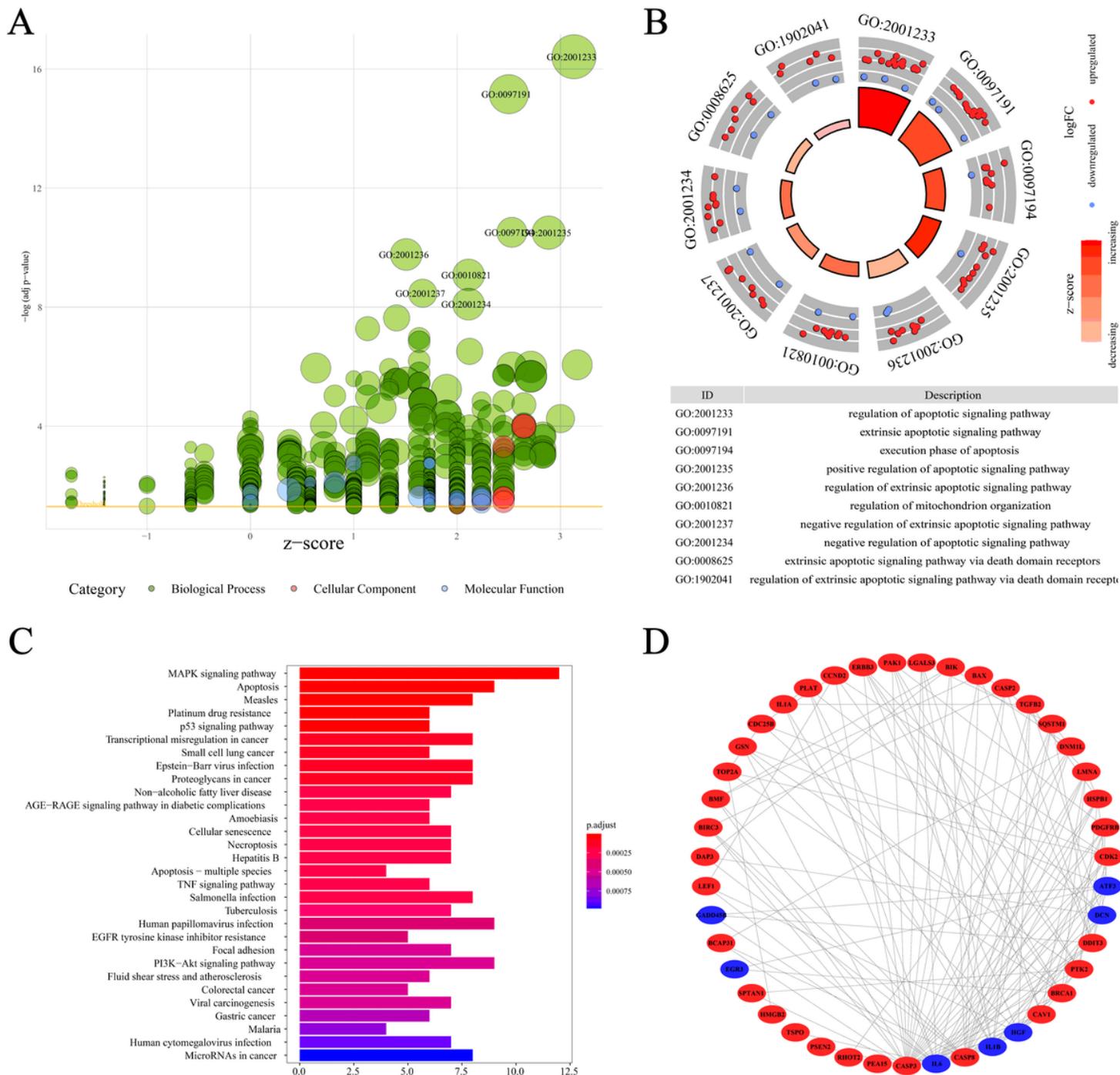


Figure 2

GO enrichment, KEGG pathway and PPI analysis of 51 identified ARGs. The results of the GO analysis were displayed in the bubble chart (a) and the circle plot (b). c The results of KEGG analysis are shown in the bar plot. d The PPI network of 51 identified ARGs.

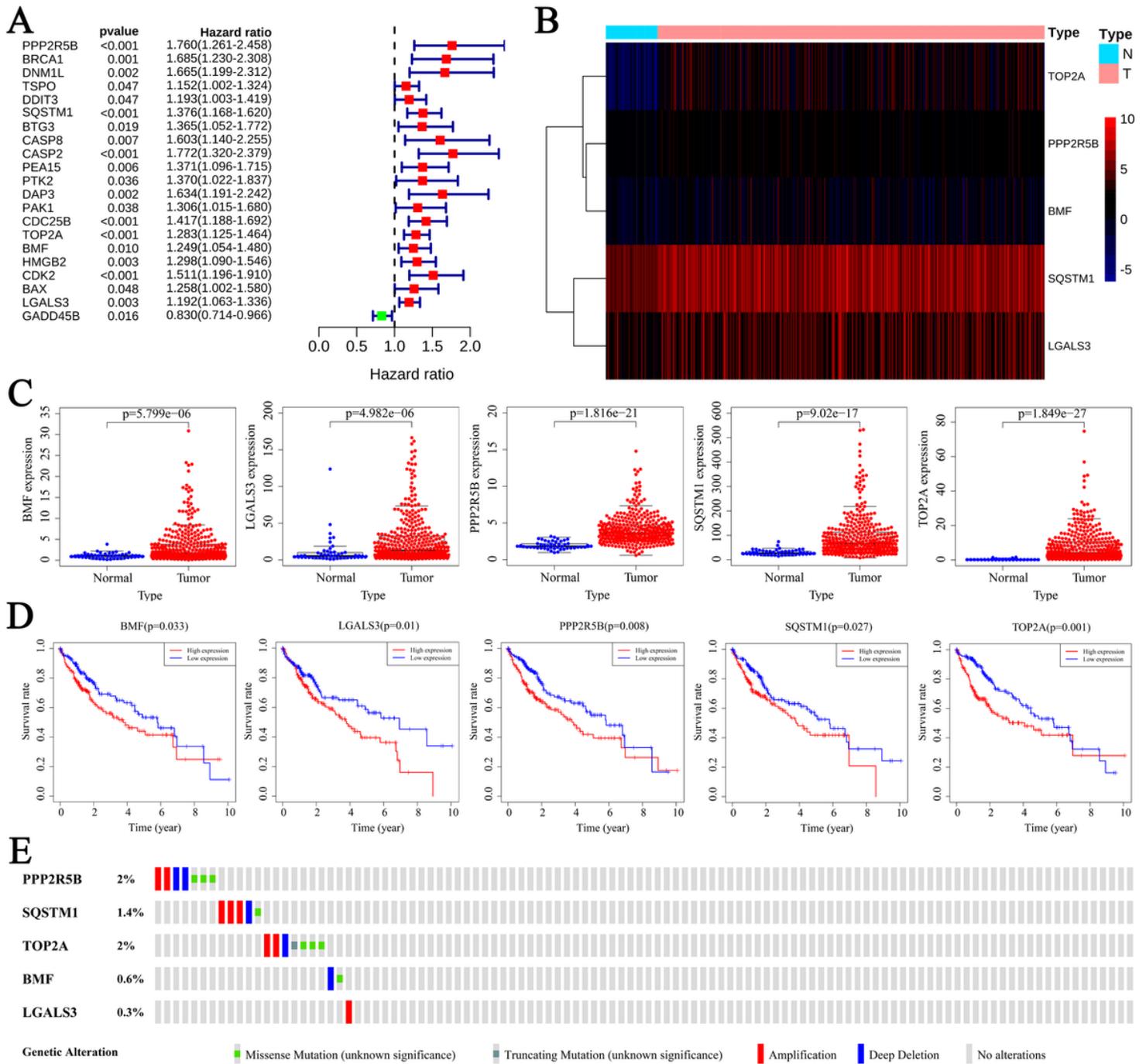


Figure 3

Identification of ARGs related to 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 between HCC and normal specimens. d The relationship between OS in HCC patients and expression levels of 5 screened ARGs in K-M curve. e Mutation data of 5 screened ARGs among 353 HCC specimens according to the cBioPortal database. OS: overall survival rates.

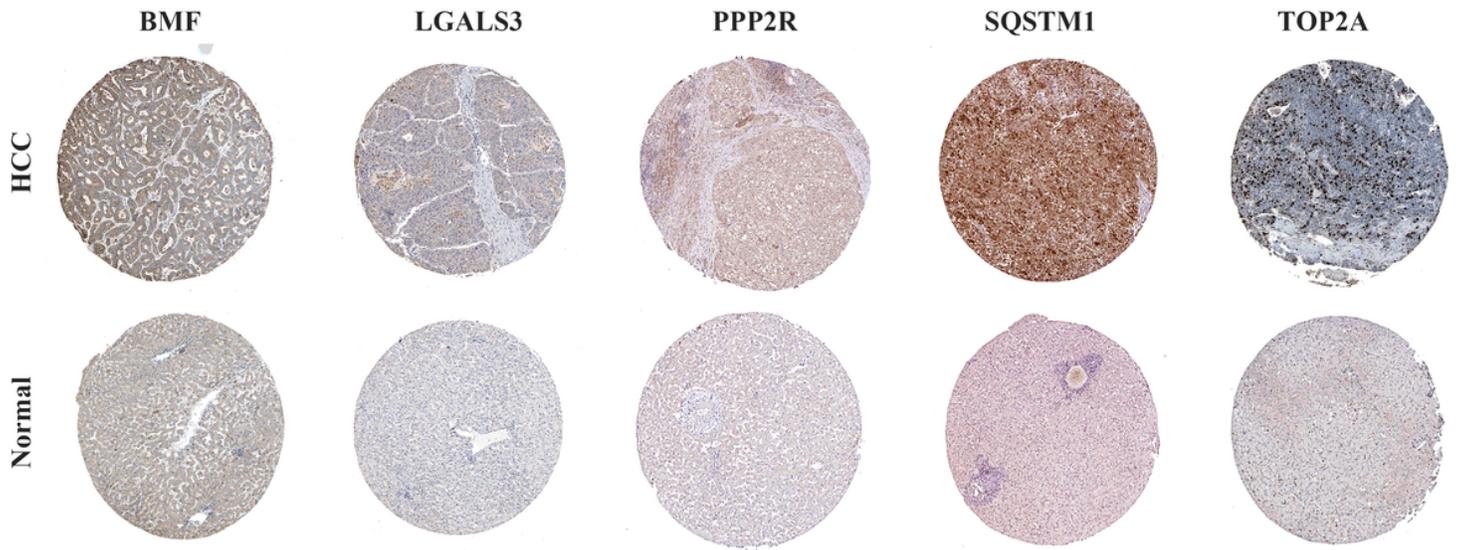


Figure 4

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.

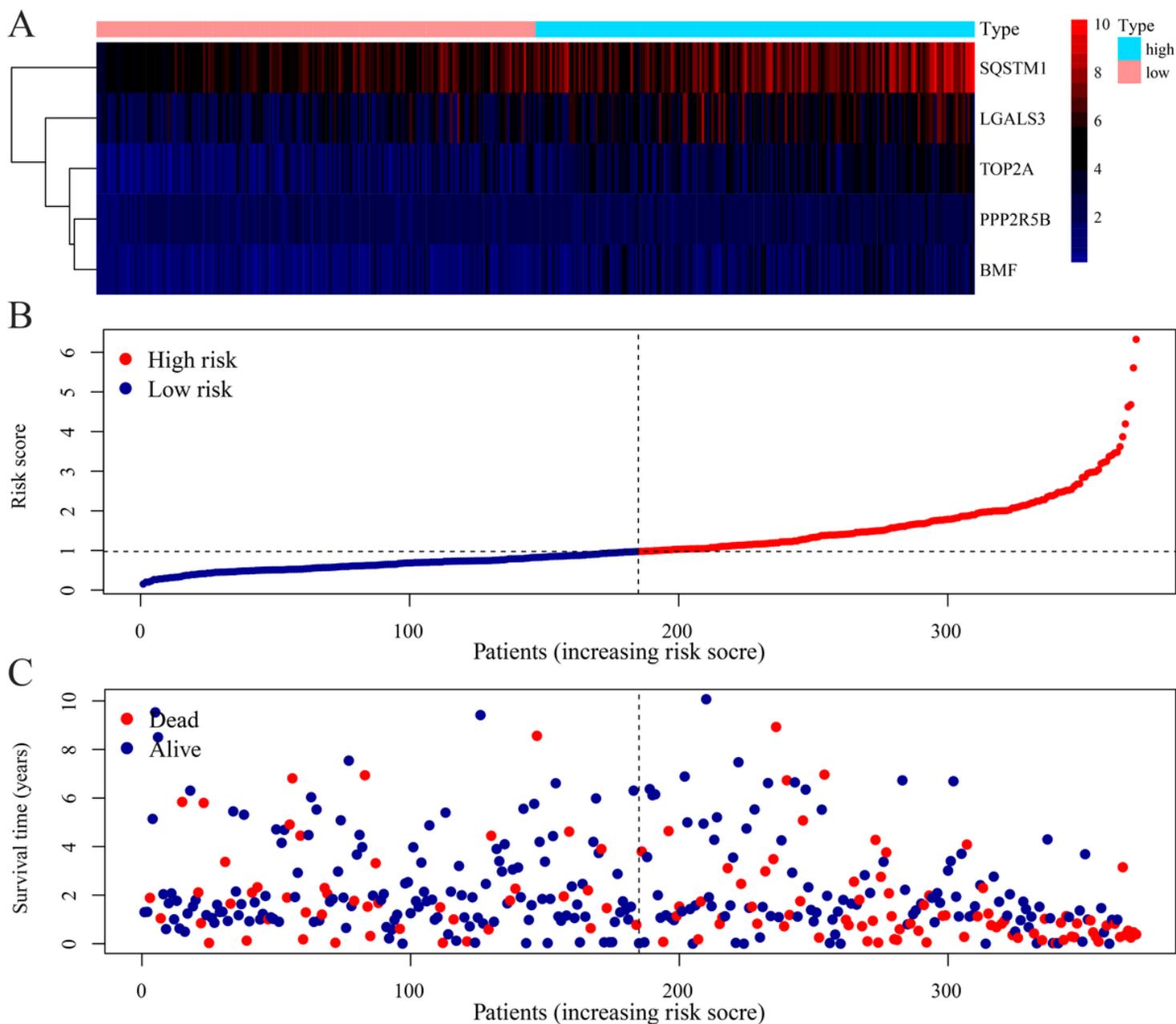


Figure 5

Prediction of OS 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 The distribution of risk score in HCC patients. c Overall survival and survival status of HCC patients in the TCGA database.

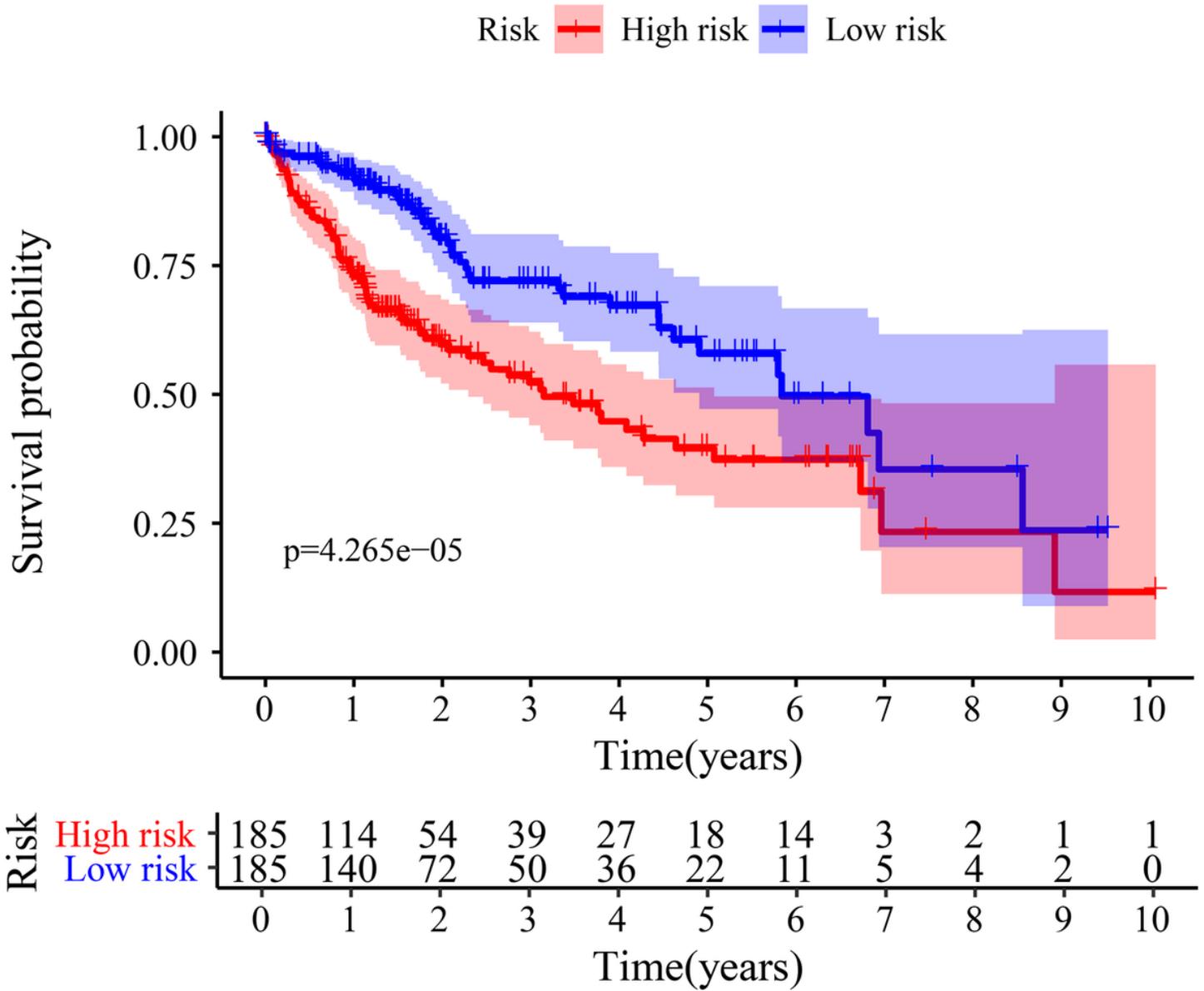


Figure 6

Survival analysis represents that compared with high-risk groups, patients in low-risk groups have significantly longer OS.

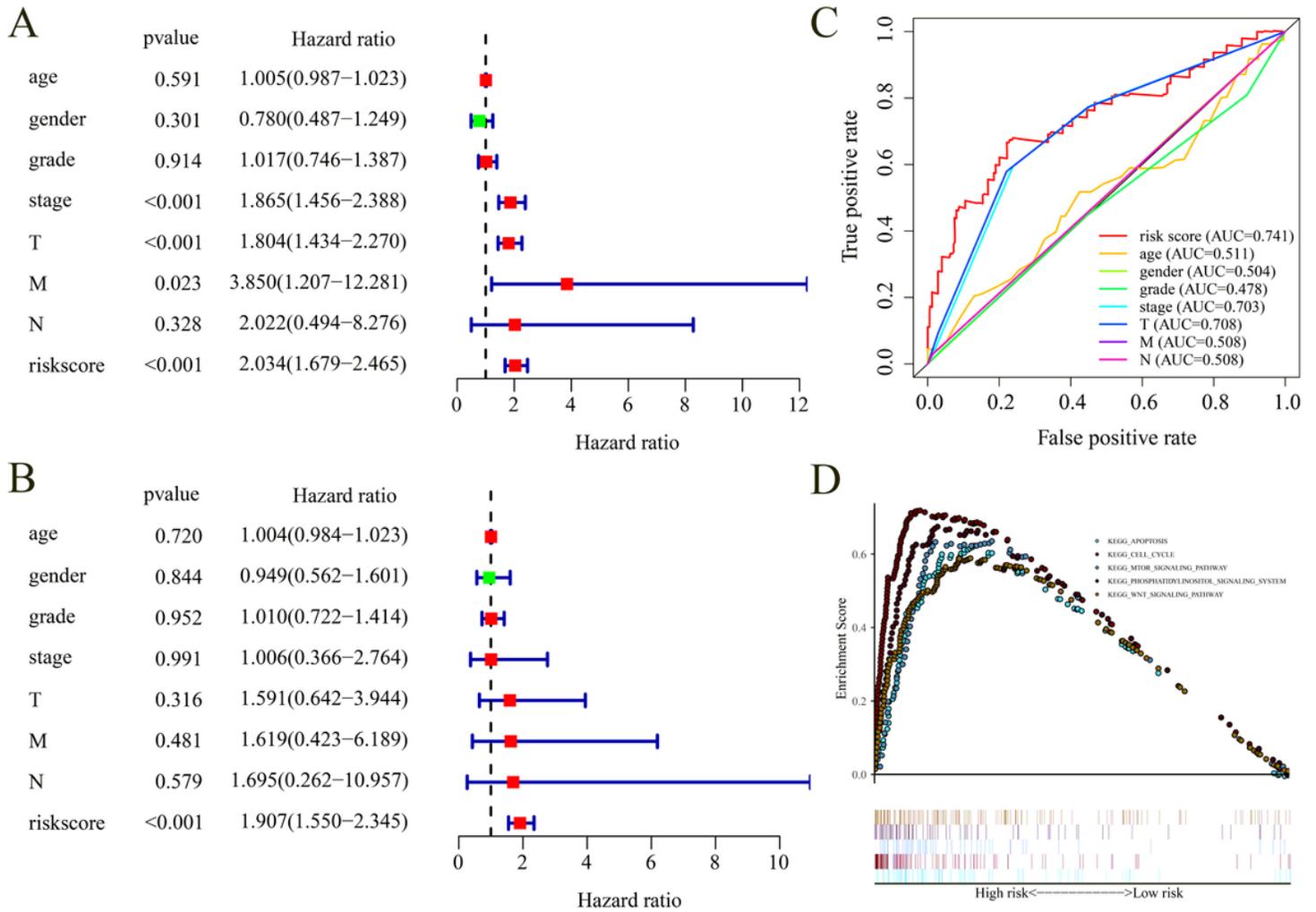


Figure 7

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 involved in according to KEGG analysis.

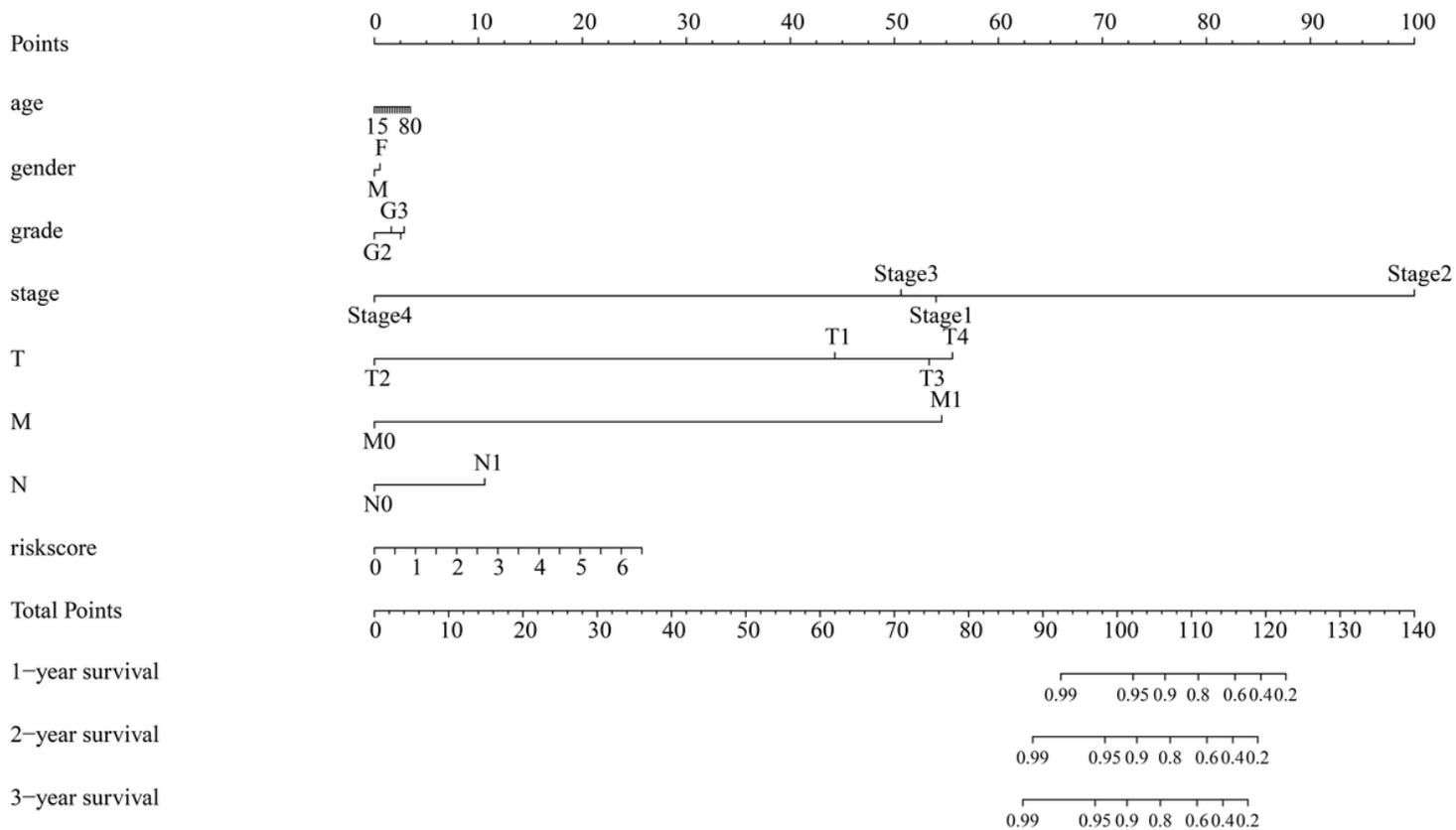


Figure 8

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

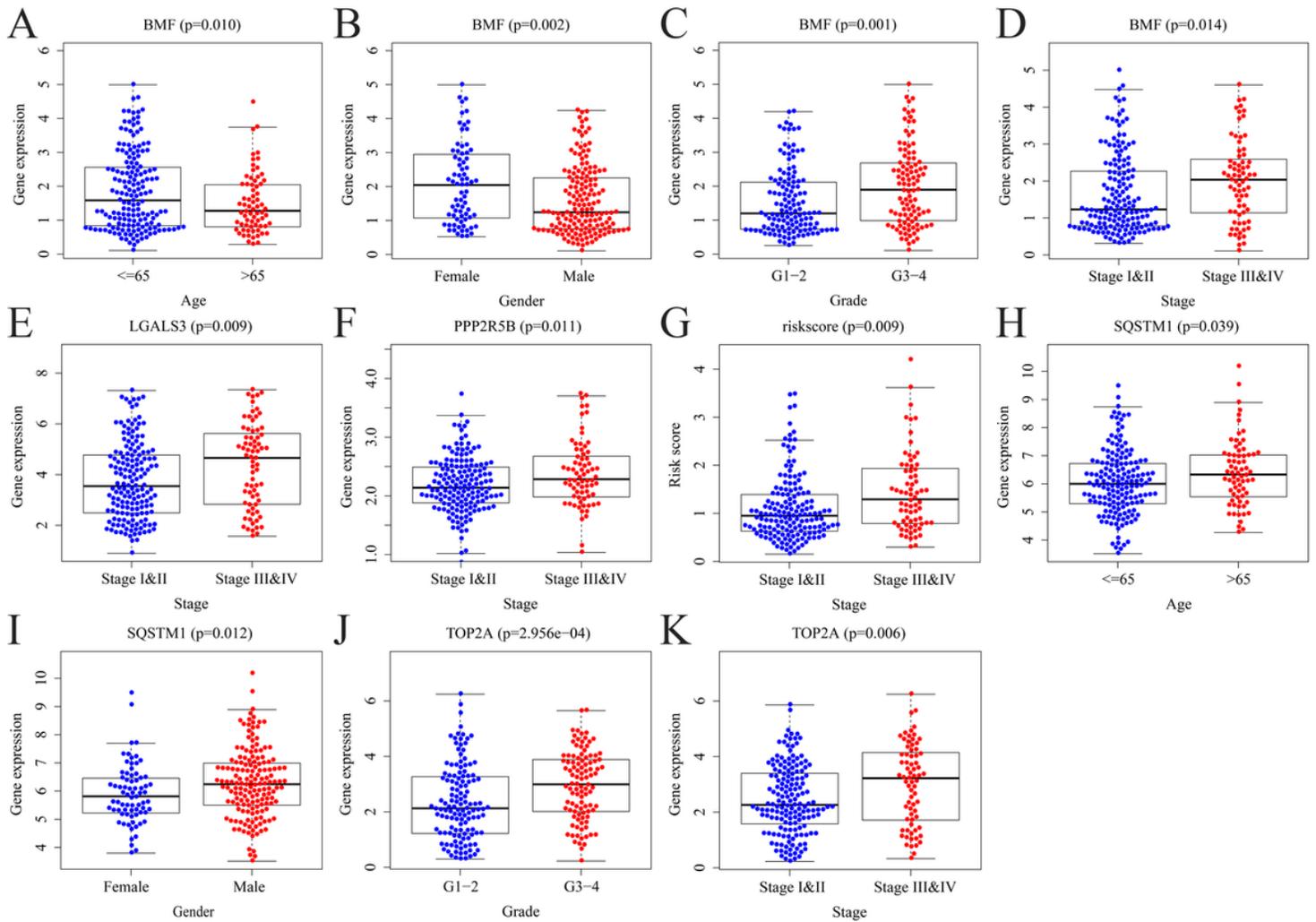


Figure 9

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

Supplementary Files

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

- [SupplementaryTable1.docx](#)