Development and validation of a survival model based on autophagy-associated genes for predicting prognosis of hepatocellular carcinoma

Wanli Yang
Fourth Military Medical University
Liaoran Niu
Fourth Military Medical University
Xinhui Zhao
Northwest University
Lili Duan
Fourth Military Medical University
Yiding Li
Fourth Military Medical University
Xiaoqian Wang
Fourth Military Medical University
Yujie Zhang
Fourth Military Medical University
Wei Zhou
Fourth Military Medical University
Jinqiang Liu
Fourth Military Medical University
Qingchuan Zhao
Fourth Military Medical University
Yu Han
Fourth Military Medical University
Daiming Fan
Fourth Military Medical University
Liu Hong (✉ Hongliufmmu@163.com)
Fourth Military Medical University https://orcid.org/0000-0002-8276-2345

Primary research

Keywords: Hepatocellular carcinoma, Autophagy-associated genes, The Cancer Genome Atlas, Prognosis, Nomogram

Posted Date: May 28th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-31172/v1

Abstract

Background: Hepatocellular carcinoma (HCC) is one of the devastating tumors with increasing incidence. Autophagy-associated genes (ARGs) are widely participated in the cellular processes of HCC. This study proposed to identify the novel prognostic gene signature based on ARGs in HCC.

Methods: We downloaded the RNA sequencing data and clinical information of HCC and normal tissues from The Cancer Genome Atlas (TCGA) database. The differentially expressed ARGs were screened by the Wilcoxon signed-rank test. Functional enrichment analyses were conducted to explore the biological implications and mechanisms of ARGs in HCC. Cox regression analysis and Lasso regression analysis were performed to screen the ARGs which related to overall survival (OS). The OS-related ARGs were then used to establish a prognostic prediction model. Kaplan-Meier curves and receiver operating characteristic (ROC) curves were both applied to evaluate the accuracy of the model. GSE14520 dataset was downloaded as the testing cohort to validate the prognostic risk model in TCGA. A nomogram based on the clinical features and risk signature was established to predict the 3-year and 5-year survival rate of HCC patients.

Results: Totally 27 differentially expressed ARGs were screened in this study. Then, 3 OS-related ARGs (SQSTM1, HSPB8, and BIRC5) were identified via the Cox regression and Lasso regression analyses. Based on these 3 ARGs, a prognostic prediction model was constructed. HCC patients in high-risk group presented poorer prognosis than those with low risk score in TCGA cohort (3-year OS, 53.7% vs 70.2%; 5-year OS, 42.0 % vs 55.2%; P=4.478e-04) and in the testing group (3-year OS, 57.7% vs 73.5%; 5-year OS, 43.2% vs 63.0%; P=1.274e-03). The risk score curve showed a well feasibility in predicting the patients' survival both in TCGA and GEO cohort with the area under the ROC curve (AUC) of 0.756 and 0.672, respectively. Besides, the calibration curves and C-index indicated that the clinical nomogram performs well to predict the 3-year and 5-year survival rate in HCC patients.

Conclusions: The survival model based on the ARGs may be a promising tool to predict the prognosis in HCC patients.

1. Background

Hepatocellular carcinoma (HCC), which is the most frequent liver malignancy, ranks as the third leading cause for cancer deaths in the world [1]. According to the studies, hepatitis B virus, hepatitis C virus, and alcoholism are the most common risk factors for HCC [2, 3]. Despite the advances made in the diagnosis and treatments for HCC patients, this disease still is a challengeable threat to the health of individuals. Moreover, due to the tumor recurrence, metastasis, and drug resistance, the 5-years survival rate of HCC patients is unsatisfied [4]. Therefore, it is pressing to identify the novel and specific biomarkers and targets for diagnosis, prognostic analysis, and targeted therapy in HCC.

Autophagy is an important process which allows lysosomes to degrade the damaged and non-functional proteins or organelles [5]. Recently, increasing number of investigations demonstrated that the abnormal

autophagy is involved in many types of tumors, including esophageal cancer, gastric cancer, and breast cancer [5–7]. What's more, some autophagy genes have shown the potential to serve as the ideal biomarkers or therapeutic targets for cancer management [5, 6]. Recent studies also revealed the relationships between pathophysiological processes of HCC and autophagy. For example, Fang et al. found that suppression of autophagy can inhibit the hepatitis C virus replication in human hematoma cells [8]. Pan et al reported that up-regulation of p62/SQSTM1 can decrease the sensitivity of HCC cells towards sorafenib [9]. LC3, a vital marker for autophagy, was also demonstrated to be a promising indicator for predicting the prognosis of HCC patients [10]. In addition, some studies have given evidence that small molecules that regulate the autophagy regulatory mechanisms may provide new clues for targeted therapy in advanced HCC [11]. For instance, ATG5 siRNA could suppress autophagy and enhance norcantharidin-induced apoptosis in HCC [12]. Xue et al. found that ULK1 may act as a novel target for HCC treatment [13]. However, little literature is available regarding the prediction of HCC prognosis using the prognostic models based on autophagy associated genes (ARGs). Our study proposed to explore the impact of ARGs on the prognosis of HCC patients, thus helping to improve the prognostic judgement and targeted therapy.

In this study, we screened 3 ARGs which closely related to the overall survival (OS) of HCC patients from the TCGA database. A prognosis prediction model was constructed based on these 3 ARGs and demonstrated to perform well for HCC patients in both training and testing cohort. Moreover, we established a clinical nomogram combining the OS-related ARGs and clinicopathological factors (age, gender, grade, stage, T (primary tumor), N (lymph nodes), M (metastasis), and risk score) and showed its good performance for predicting the 3-year and 5-year survival rate of HCC patients.

2. Materials And Methods

2.1 Data collection and procession

The Human Autophagy Database (http://www.autophagy.lu/index.html) was used to download the entire 232 ARGs [14]. RNA sequencing data and the corresponding clinical data for 374 HCC and 50 non-tumor tissues were acquired from the TCGA database (https://tcga-data.nci.nih.gov/tcga/) [15]. The Data procession was conducted as the previous study [16]. The cBio Cancer Genomics Portal (http://cbioportal.org) was utilized to explore the genetic alterations and clinical information of the selected ARGs in HCC [17].

2.2 Differentially expressed ARGs in HCC

The Wilcoxon signed-rank test in package "limma" in R software (version 3.6.3) was applied to screen the differentially expressed ARGs between HCC and normal tissues, with the criteria of $|\log_2$ fold change (FC)$|>1.5$ and adjusted P-value < 0.05 [16, 18]. Then, we combined the expression data of the differentially expressed ARGs with the corresponding clinical data. Univariate Cox regression and multivariate Cox regression analyses were applied to pick out the ARGs that were closely related to OS in HCC[18]. We also performed the Lasso regression to remove ARGs that might be closely correlated with others [14, 16].

2.3 Functional enrichment analyses

To explore the biological implications and potential mechanisms of ARGs in HCC, GO annotation and KEGG pathway analyses were conducted for ARGs via R software with “GO plot”, “ggplot2”, “Cluster Profiler” and “DOSE” packages, etc [16]. Top enriched terms with p-value < 0.05 and q-value < 0.05 were regarded noteworthy.

2.4 Establishment of prognostic model for HCC

A linear combination of the expression of ARGs and regression coefficient which based on the multivariate Cox regression was used to construct the risk signature[14]. The OS-related predictive formula was calculated as following: risk score = (Expression $_{gene1}$ × Coefficient $_{gene1}$) + (Expression $_{gene2}$ × Coefficient $_{gene2}$) +…+ (Expression $_{genen}$ × Coefficient $_{genen}$) [14]. Then, each patient with HCC was assigned a risk score according to the formula.

2.5 Assessment of the prognostic model

Kaplan-Meier survival curves and time-dependent ROC curves were performed to evaluate the efficiency of the prognostic model[19].

2.6 External validation of the prognostic gene signature

GSE14520 dataset with 221 HCC patients was downloaded from the GEO database (http://www.ncbi.nlm.nih.gov/geo/) [20, 21]. Each patient received a risk score which was calculated using the same prognostic gene-signature based risk model in TCGA dataset. Then, the Kaplan-Meier curves and ROC curve were performed to evaluate the predictive performance of the prognostic gene signature.

2.7 Construction of the clinical nomogram

We downloaded the clinical data of HCC patients from the TCGA database. Then, a clinical nomogram based on several factors (age, gender, grade, stage, T (primary tumor), N (lymph nodes), M (metastasis), and risk score) was built for predicting the 3-year and 5-year survival rate of HCC patients by using the “survival” and “rms” packages in R software[19]. Moreover, the concordance index (C-index) and calibration curves were both applied to evaluate the accuracy of the nomogram.

2.8 Validation of the expression of ARGs at protein level and mRNA level

The Human Protein Atlas database (https://www.proteinatlas.org/) contains more than 11,200 unique proteins [22], we therefore utilized it to evaluate the proteins levels of the prognostic ARGs in HCC tissues and normal tissues. TIMER database (https://cistrome.shiny apps.io/timer/) was used to validate the mRNA levels of ARGs in HCC tissues and normal samples[23].

2.9 Construction of transcription factors-genes networks and miRNA-genes networks

The NetworkAnalysis (http://www.networkanalyst.ca) database was conducted to predict the transcription factors (TFs) and miRNAs of SQSTM1, HSPB8, and BIRC5 [24]. The TFs prediction was based on the ENCODE database with ChIP-seq data [24]. Only the results with a peak intensity signal value < 500 and a potential score value < 1 were identified for further study[24]. The miRNA-genes network was predicted and constructed by the TarBase and miRTarBase in the NetworkAnalysis database.

2.10 Statistical analysis

The Perl language and R software 3.6.3 were applied to conduct all the statistical tests and graphics. $P < 0.05$ was regarded as statistical significance.

3. Results

3.1 Differentially expressed ARGs between HCC and adjacent nontumor tissues

In this study, a total of 374 HCC and 50 non-tumor tissues with RNA sequencing data were analyzed. Eventually, 27 differentially expressed ARGs were identified, including 25 up-regulated genes (DDIT3, BAX, TSC1, HGS, BAK1, HSP90AB1, RAB24, CLN3, SQSTM1, PEA15, IKBKE, TP63, HSPB8, ITGA6, ITGA3, DAPK2, TMEM74, ITGB4, IRGM, NKX2-3, SPHK1, NRG2, TP73, CDKN2A, and BIRC5) and 2 down-regulated genes (FOS and DIRAS3) with $|\log_2 FC| > 1.5$ (Fig. 1a-1b). Figure 1c illustrated these 27 differentially expressed ARGs in HCC and normal tissues.

3.2 Function enrichment analyses for ARGs

To investigate the biological functions and molecular mechanisms of the above 27 ARGs in HCC, GO enrichment analysis and KEGG pathway analysis were conducted (Fig. 2, Fig. 3). The GO enrichment analysis can be divided into three parts: biological processes (BP), cellular components (CC) and molecular function (MF). As showed in Fig. 2, the top enriched terms for BP included autophagy, process utilizing autophagic mechanism, and regulation of apoptotic signaling pathway. As for CC, the most significant terms involved in autophagosome, chaperone complex, and integrin complex. In MF group, these differentially expressed ARGs were mainly associated with BH domain binding, chaperone binding, and heat shock protein binding. Besides, KEGG enrichment analysis showed that the selected differentially expressed ARGs were mainly participated in pathways in cancer, human papillomavirus infection, apoptosis, measles, and platinum drug resistance (Fig. 3a-3b).

3.3 Prognostic gene signature for HCC cohorts

Using a univariate Cox regression analysis, totally 13 differentially expressed ARGs (BAX, SQSTM1, PEA15, CDKN2A, HSPB8, HGS, IKBKE, HSP90AB1, RAB24, BIRC5, DDIT3, BAK1, and TMEM74) were found to be markedly related to OS in HCC patients (Fig. 4a). All these 13 survival-related ARGs were regarded as risk factors (HRs, 1.154–1.715; $P < 0.05$) and their overexpression may cause worse prognosis. Then, we validated the prognostic roles of these 13 survival-related ARGs in Kaplan Meier-plotter website (http://kmplot.com/analysis/index.php). The results are consistent with the findings in TCGA dataset (**Fig. S1**; The OS curve of BAX presented similar trend to these of other genes, but not statistically significant). Finally, the 13 differentially expressed ARGs were entered into a Lasso regression analysis. Figure 4b presented the regression coefficient of these 13 ARGs in HCC. While 7 ARGs (SQSTM1, PEA15, CDKN2A, HSPB8, RAB24, BIRC5, and TMEM74) were included, the model reached the best performance (Fig. 4c). The biological functions and risk coefficients of 7 ARGs were listed in Table 1, which mainly associated with the formation of autophagosome, regulation of autophagy, as well as regulation of apoptosis.

Table 1
Biological functions and coefficient of 7 ARGs.

No	Gene symbol	Full name	Function	Risk coefficient
1	SQSTM1	Sequestosome 1	Functions as a bridge between polyubiquitinated cargo and autophagosomes	0.76707221
2	PEA15	Proliferation And Apoptosis Adaptor Protein 15	Functions as a regulator of apoptosis and autophagy	0.24377181
3	CDKN2A	Cyclin Dependent Kinase Inhibitor 2A	Involved in regulation of autophagy and caspase-independent cell death	0.03399741
4	HSPB8	Heat Shock Protein Family B (Small) Member 8	Functions as a regulator of macroautophagy	0.19802769
5	RAB24	Ras-Related Protein Rab-24	Involved in autophagy-related processes	0.07147873
6	BIRC5	Baculoviral IAP Repeat Containing 5	Functions as a regulator of apoptosis and autophagy	0.50048513
7	TMEM74	Transmembrane Protein 74	Plays an essential role in autophagy	0.21098334

To explore the contributions of the 7 ARGs to hepatocellular carcinogenesis, the genetic alteration information of these genes were explored in the cBio Cancer Genomics database. The PanCancer Atlas dataset (353 samples) and Firehose Legacy dataset (366 samples) of HCC were both included. ARGs of interest are altered in 149 (41%) of 366 sequenced patients (TCGA, Firehose Legacy dataset) (**Fig. S2**), compared with that altered queried ARGs were detected in 110 (31%) of 353 sequenced patients (TCGA,

PanCancer Atlas dataset) (**Fig. S3a**). Moreover, HCC patients in genes-altered group presented poorer PFS (progression-free survival) (**Fig.S3b**) and DFS (disease-free survival) (**Fig.S3c**) than these in genes-unaltered group. The OS curves presented similar trend to that of PFS, but not statistically significant (**Fig.S3d**). These results suggested that the 7 ARGs play a crucial role in hepatocellular carcinogenesis.

The 7 ARGs were finally analyzed by a multivariate Cox regression analysis, and 3 candidate genes (SQSTM1, HSPB8, and BIRC5) were selected as the prognostic markers for HCC patients (Table 2). Each patient received a risk score calculated as follows: risk score = (0.2578 × expression value of SQSTM1) + (0.1190 × expression value of HSPB8) + (0.3049 × expression value of BIRC5).

Table 2
Univariate and multivariate Cox regression analyses of OS in HCC patients.

Genes	HR(95% CI)	P	HR(95% CI)	P	Coef
SQSTM1	1.3759(1.1682–1.6204)	0.000132	1.2940(1.1007–1.5213)	0.001792	0.257771
PEA15	1.3709(1.0956–1.7153)	0.005811	—	—	—
CDKN2A	1.2434(1.0706–1.4441)	0.004323	—	—	—
HSPB8	1.1542(1.0371–1.2846)	0.008595	1.1264(1.0079–1.2588)	0.035779	0.119042
RAB24	1.7155(1.2278–2.3970)	0.001566	—	—	—
BIRC5	1.3523(1.1790–1.5510)	1.60E-05	1.3565(1.1810–1.5581)	1.61E-05	0.304895
TMEM74	1.5356(1.0999–2.1439)	0.011769	—	—	—

3.4 Identification of independent risk factors of OS for HCC patients

In order to screen the independent risk factors of OS for HCC patients, the univariate cox and multivariate cox regression analyses were carried out to explore the independent risk factors of OS. According to Fig. 5a, tumor stage, T (primary tumor), and risk score were closely related to OS (HR = 1.669(95% CI: 1.357–2.053), P < 0.001; HR = 1.649(95% CI: 1.354–2.009), P < 0.001, and HR = 1.755(95% CI: 1.511–2.039), P < 0.001, respectively). As shown in Fig. 5b, the results of multivariate cox regression indicated that M (metastasis) (HR = 1.394(95% CI: 1.065–1.824), P = 0.016) and risk score (HR = 1.769(95% CI: 1.478–2.116), P < 0.001) should be considered as the independent risk factors of OS.

3.5 Validation of the risk model

According to the median values of risk score, we divided 374 HCC cases into high-risk and low-risk groups. To validate the performance of risk model, we plotted Kaplan-Meier curves to compare the HCC survival in two groups. The results showed that HCC patients in low-risk group have better prognosis than those in high-risk group (3-year OS, 70.2% vs 53.7%; 5-year OS, 55.2% vs 42.0%; P = 4.478e-04) (Fig. 6a). The ROC curves were also plotted using the risk factors related to OS (age, gender, grade, stage, T (primary tumor), M (metastasis), N (lymph nodes), and risk score). We then evaluated the area under curve (AUC) values for each risk factor, and the results showed that risk score curve has a better feasibility in predicting the individuals' survival with AUC of 0.756 (Fig. 6b). In addition, **Figure. 6c-6e** showed that the survival time of HCC patients decreases along with the rising of risk score.

The prognostic role of the three-gene risk model was then validated in the GEO validation cohort (GSE14520; n = 221). According to the median value of risk score, the patients were also divided into two groups (high risk group and low risk group). The Kaplan-Meier survival result showed that HCC patients in high risk group have a significantly worse OS compared to cases with low risk in the validation dataset (3-year OS, 57.7% vs 73.5%; 5-year OS, 43.2% vs 63.0%; P = 1.274e-03) (Fig. 7a). Moreover, Fig. 7b demonstrated that the risk score presents a well predictive ability with AUC of 0.672. Figure 7c-7e showed that patients in high risk group have shorter survival time than patients with low risk score in validation cohort. Taking together, the three-gene signature has good feasibility for predicting the prognosis in HCC.

3.6 Construction of a nomogram for predicting survival rate of HCC

The clinical nomogram was constructed to quantitatively assess the patients' survival rate through combining several risk factors (age, gender, grade, stage, T (primary tumor), N (lymph nodes), M (metastasis), and risk score). As shown in Fig. 8a, the total points of risk factors were utilized to evaluate the individuals' 3-year and 5-year survival rates. The concordance index (C-index) was 0.68 (95% CI: 0.63–0.73). In addition, calibration curves presented good concordance between actual survival and nomogram-predicted survival (Fig. 8b-8c), especially for the 3-year survival.

3.7 The relationships between ARGs and clinical factors

The Student's t test was applied to investigate the relationships between the expression levels of 3 ARGs (SQSTM1, HSPB8, and BIRC5) and clinical factors. The results showed that the risk scores increased along with the T (primary tumor) and tumor grade (Fig. 9a-9b). In addition, the expression of SQSTM1 was higher in the groups of patients aged > 65, male, and patients with higher tumor grade (Fig. 9c-9e). We also found the level of BIRC5 was related to the grade, tumor stage and T (primary tumor) in HCC patients (Fig. 9f-9 h).

3.8 Validation of the expression of ARGs at protein level and mRNA level

The Human Protein Atlas database (https://www.proteinatlas.org/) was used to evaluate the protein levels of 3 prognostic ARGs (SQSTM1, HSPB8, and BIRC5) in HCC tissues with their expression in normal tissues (Fig. 10a-10c). As expected, the protein levels of 3 prognosis-related ARGs (SQSTM1, HSPB8, and BIRC5) were markedly higher in HCC tissues as compared with normal samples. Then, the TIMER database (https://cistrome.shinyapps.io/timer/) was applied to validate the mRNA expression levels of SQSTM1, HSPB8, and BIRC5. The results showed that the mRNA levels of these 3 ARGs were obviously higher in HCC compared to normal controls (**Fig. S4**).

3.9 Construction of TFs-genes networks and miRNA-genes networks for SQSTM1, HSPB8, and BIRC5

To better understand the contributions of SQSTM1, HSPB8, and BIRC5 to the development and progression of HCC. The TFs-genes networks and miRNA-genes networks for SQSTM1, HSPB8, and BIRC5 were established (Fig. 11a-11b). The numbers of TFs and miRNAs in the networks were 100 and 117, respectively. In the TFs-genes networks (Fig. 11a), ZNF394 was identified as the hub TF for these 3 target genes. Moreover, SQSTM1 and BIRC5 shared 10 TFs (ZNF394, ZBTB7A, SCRT1, ZNF644, SSRP1, MLX, PPARG, CTCF, MDX3, and ZNF501). In the miRNA-genes networks (Fig. 11b), miR-218-5p, miR-646, miR-93-5p, miR-16-5p, miR-484, miR-335-5p, and miR-1252-3p could regulate both SQSTM1 and BIRC5. In addition, SQSTM1 and HSPB8 were predicted as the targeted genes of miR-1226-3p. Taken together, the TFs-genes networks and miRNA-genes may provide new clues for the further studies on molecular mechanisms of HCC.

4. Discussion

HCC is a common and frequently occurring malignancy worldwide. Due to the lack of ideal prognostic biomarkers, HCC patients usually can’t receive reasonable treatment immediately [25]. Traditional prognostic risk factors (such as tumor size, histological type, stage, and grade) could only be adopted and evaluated post-surgery. Some scholars even believe that the present TNM stage system should be perfected for its insufficient to accurately predict the prognosis of cancer patients [26–28]. Moreover, different patients can present different treatment responses. Thus, more specific and effective markers are required to be identified for evaluating prognosis and screening potentially HCC patients with high risk. To date, increasing number of biomarkers have been screened for the prediction of prognosis in HCC [29, 30]. For example, Dai et al. confirmed that high expression of HIF-1α is an independent prognostic factors for OS and DFS of HCC patients[31]. Similarly, our lab previously found that high levels of SOX12 is an independent and important risk factor for HCC patients [32]. However, the translation of these biomarkers into clinical application still leaves much to be desired. Firstly, researchers need to investigate the molecular mechanisms behind the dysregulation of marker. What’s more, these biomarkers are expected to be evaluated in large samples. In addition, the single gene study may provide inaccurate conclusion since the expression levels of single gene can be affected by many factors. The prognostic

gene signature may solve these problems and eventually provide more convincing results since it is based on the statistical model and comprise of multiple related genes.

Numerous researches have reported that autophagic dysfunctions were associated with several pathophysiological processes, including inflammation, metabolic disorder, neurodegeneration and cancer [33, 34]. Autophagy can sever as both tumor suppresser role and oncogenetic role in tumorigenesis, which may depend on the tumor microenvironments and tumor heterogeneity[6, 35]. For example, Fan et al. reported that autophagy can promote the metastasis and glycolysis of HCC cells through Wnt/β-catenin pathway[36]. Conversely, exenatide induced autophagy can inhibit the cell proliferation of HCC cells [37]. Considering the emerging role of autophagy in cancers, autophagy-related studies may provide us a better understanding on the pathogenesis and prognosis of HCC. Importantly, the risk gene signature obtained from the entire set of ARGs could be superior to single gene for predicting the survival.

So far, the rapid development of gene chip assays and second-generation gene sequencing have greatly facilitated the development of personalized medicine and precision medicine. Increasing numbers of biomarkers have been identified by integratedly analyzing the genomic data form individual specimens. It is indisputable that these methods can finally consolidate and improve the current clinical setting for cancer management. To our knowledge, this study firstly explored the prognostic roles of entire reported ARGs in HCC. Here, 27 differentially expressed ARGs was identified from 374 HCC tissues and 50 normal tissues. Then, the functional enrichment analyses were conducted to explore the roles and mechanisms of the differentially expressed ARGs in HCC. Though the Cox regression and Lasso regression analyses, we established a risk model based on 3 OS-related ARGs (SQSTM1, HSPB8, and BIRC5). The HCC patients were then divided into two groups (high risk group and low risk group) according to the risk score derived from this model. Kaplan-Meier curves and ROC curves suggested that the risk model performed well in both training cohort and testing cohort. What's more, the clinical nomogram combining the clinicopathological features and risk score was applied to predict the 3-year and 5-year survival rate of HCC patients. And the calibration plots and *C*-index illustrated a good performance of the nomogram to predict patient' survival.

However, there are still some deficiencies in the study. First, our study mainly focused on the prognostic role of selected ARGs and we didn't deeply study the other ARGs. Second, we failed to validate the expression levels of prognosis-related ARGs by *in vitro* or *in vivo* assays. Third, the prognostic model should be evaluated in large clinical samples. Finally, in-depth studies on the 3 prognosis-related ARGs may improve the targeted therapy of HCC patients.

5. Conclusions

In conclusion, 3 OS-related ARGs (SQSTM1, HSPB8, and BIRC5) were screened. Based on these genes, we established a prognosis prediction model, which presented a good efficacy in guiding the personalized medicine for HCC patients. The above results suggested that ARGs signature can act as the effective and

promising prognostic indicator for HCC patients. Moreover, further studies for these ARGs may also improve the cancer management of HCC.

Abbreviations

HCC: Hepatocellular carcinoma; FC: Fold change; FDR: False discovery rate; OS: Overall survival; TCGA: The Cancer Genome Atlas; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; ROC: Receiver operating characteristic; AUC: The area under the ROC curve.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data used in the study were downloaded from The Cancer Genome Atlas (TCGA) database and GEO database.

Competing interests

The authors declare that they have no conflicts of interest.

Funding

This study was supported in part by grant from the Scientific Foundation of Shaanxi Province (S2019ZDCXL01-02-01; 2018SF-240) and grant from the National Clinical Research Center for Digestive Diseases (2015BAI13B07).

Authors' contributions

YH, DF, and LH conceived and designed the study; WY, LN, XZ and LD wrote the manuscript; XW, YL, and YZ collected the literature; QZ, YH, DF, and LH reviewed and revised the manuscript; WZ and JL conducted the statistical analysis. All authors approved the final version of manuscript.

Acknowledgements

Not applicable.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.
2. Llovet JM, Montal R, Sia D, Finn RS. Molecular therapies and precision medicine for hepatocellular carcinoma. Nature reviews Clinical oncology. 2018;15(10):599–616.
3. Fujiwara N, Friedman SL, Goossens N, Hoshida Y. Risk factors and prevention of hepatocellular carcinoma in the era of precision medicine. J Hepatol. 2018;68(3):526–49.
4. Chang Lee R, Tebbutt N. Systemic treatment of advanced hepatocellular cancer: new hope on the horizon. Expert Rev Anticancer Ther. 2019;19(4):343–53.
5. Onorati AV, Dyczynski M, Ojha R, Amaravadi RK. **Targeting autophagy in cancer**. 2018, 124(16):3307–3318.
6. Huang T, Song X, Yang Y, Wan X, Alvarez AA, Sastry N, Feng H, Hu B, Cheng SY. Autophagy and Hallmarks of Cancer. Crit Rev Oncog. 2018;23(5–6):247–67.
7. Amaravadi R, Kimmelman AC, White E. Recent insights into the function of autophagy in cancer. Genes Dev. 2016;30(17):1913–30.
8. Fang S, Su J, Liang B, Li X, Li Y, Jiang J, Huang J, Zhou B, Ning C, Li J, et al. Suppression of autophagy by mycophenolic acid contributes to inhibition of HCV replication in human hepatoma cells. Scientific reports. 2017;7:44039.
9. Pan J, Lu C, Jun W, Wu Y, Shi X, Ding Y. The up-regulation of P62 levels is associated with resistance of sorafenib in hepatocarcinoma cells. Int J Clin Exp Pathol. 2019;12(7):2622–30.
10. Meng YC, Lou XL, Yang LY, Li D, Hou YQ: **Role of the autophagy-related marker LC3 expression in hepatocellular carcinoma**: **a meta-analysis**. 2020, **146**(5):1103–1113.
11. Lee YG, Jeon TI. **Modulation of the Autophagy-lysosomal Pathway in Hepatocellular Carcinoma Using Small Molecules**. *Molecules (Basel, Switzerland)* 2020, 25(7).
12. Xiong X, Wu M, Zhang H, Li J, Lu B, Guo Y, Zhou T, Guo H, Peng R, Li X, et al. Atg5 siRNA inhibits autophagy and enhances norcantharidin-induced apoptosis in hepatocellular carcinoma. Int J Oncol. 2015;47(4):1321–8.
13. Xue ST, Li K, Gao Y, Zhao LY, Gao Y, Yi H, Jiang JD, Li ZR. **The role of the key autophagy kinase ULK1 in hepatocellular carcinoma and its validation as a treatment target**. *Autophagy* 2020:1–15.
14. Wang Y, Zhao W, Xiao Z, Guan G, Liu X, Zhuang M. **A risk signature with four autophagy-related genes for predicting survival of glioblastoma multiforme**. 2020, 24(7):3807–3821.
15. Wang Z, Jensen MA, Zenklusen JC. A Practical Guide to The Cancer Genome Atlas (TCGA). Methods Mol Biol. 2016;1418:111–41.

16. Wang X, Yao S, Xiao Z, Gong J, Liu Z, Han B, Zhang Z. Development and validation of a survival model for lung adenocarcinoma based on autophagy-associated genes. J Transl Med. 2020;18(1):149.
17. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401–4.
18. Hu D, Jiang L, Luo S, Zhao X, Hu H, Zhao G, Tang W. Development of an autophagy-related gene expression signature for prognosis prediction in prostate cancer patients. J Transl Med. 2020;18(1):160.
19. Liu Y, Wu L, Ao H, Zhao M, Leng X, Liu M, Ma J, Zhu J. Prognostic implications of autophagy-associated gene signatures in non-small cell lung cancer. Aging. 2019;11(23):11440–62.
20. Roessler S, Jia HL, Budhu A, Forgues M, Ye QH, Lee JS, Thorgeirsson SS, Sun Z, Tang ZY, Qin LX, et al. A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. Cancer research. 2010;70(24):10202–12.
21. Roessler S, Long EL, Budhu A, Chen Y, Zhao X, Ji J, Walker R, Jia HL, Ye QH, Qin LX, et al. Integrative genomic identification of genes on 8p associated with hepatocellular carcinoma progression and patient survival. Gastroenterology. 2012;142(4):957–66.e912.
22. Ponten F, Schwenk JM, Asplund A, Edqvist PH. The Human Protein Atlas as a proteomic resource for biomarker discovery. J Intern Med. 2011;270(5):428–46.
23. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B, Liu XS. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. Cancer research. 2017;77(21):e108–10.
24. Xia J, Gill EE, Hancock RE. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. Nature protocols. 2015;10(6):823–44.
25. Hartke J, Johnson M, Ghabril M. The diagnosis and treatment of hepatocellular carcinoma. Semin Diagn Pathol. 2017;34(2):153–9.
26. Tellapuri S, Sutphin PD, Beg MS, Singal AG, Kalva SP. Staging systems of hepatocellular carcinoma: A review. Indian journal of gastroenterology: official journal of the Indian Society of Gastroenterology. 2018;37(6):481–91.
27. Cattoni M, Vallières E, Brown LM, Sarkeshik AA, Margaritora S, Siciliani A, Filosso PL, Guerrera F, Imperatori A, Rotolo N, et al. Improvement in TNM staging of pulmonary neuroendocrine tumors requires histology and regrouping of tumor size. J Thorac Cardiovasc Surg. 2018;155(1):405–13.
28. Ilhan E, Ureyen O, Meral UM. Ongoing problems concerning 7(th) TNM Staging System and Proposals for 8(th) TNM Staging System of Gastric Cancer. Przeglad gastroenterologiczny. 2016;11(4):223–5.
29. De Stefano F, Chacon E, Turcios L, Marti F, Gedaly R. Novel biomarkers in hepatocellular carcinoma. Dig Liver Dis. 2018;50(11):1115–23.
30. Yim SY, Hae NJ, Shin JH, Jeong YS, Kang SH, Park YN, Um SH, Lee JS. Identification of prognostic biomarker in predicting hepatocarcinogenesis from cirrhotic liver using protein and gene signatures.

Exp Mol Pathol. 2019;111:104319.

31. Dai CX, Gao Q, Qiu SJ, Ju MJ, Cai MY, Xu YF, Zhou J, Zhang BH, Fan J. Hypoxia-inducible factor-1 alpha, in association with inflammation, angiogenesis and MYC, is a critical prognostic factor in patients with HCC after surgery. BMC Cancer. 2009;9:418.
32. Huang W, Chen Z, Shang X, Tian D, Wang D, Wu K, Fan D, Xia L: **Sox12, a direct target of FoxQ1, promotes hepatocellular carcinoma metastasis through up-regulating Twist1 and FGFBP1**. *Hepatology (Baltimore, Md)* 2015, **61**(6):1920–1933.
33. Saha S, Panigrahi DP, Patil S, Bhutia SK. Autophagy in health and disease: A comprehensive review. Biomedicine pharmacotherapy = Biomedecine pharmacotherapie. 2018;104:485–95.
34. Mizushima N. A brief history of autophagy from cell biology to physiology and disease. Nat Cell Biol. 2018;20(5):521–7.
35. Mowers EE, Sharifi MN, Macleod KF. Functions of autophagy in the tumor microenvironment and cancer metastasis. Febs j. 2018;285(10):1751–66.
36. Fan Q, Yang L, Zhang X, Ma Y, Li Y, Dong L, Zong Z, Hua X, Su D, Li H, et al. Autophagy promotes metastasis and glycolysis by upregulating MCT1 expression and Wnt/β-catenin signaling pathway activation in hepatocellular carcinoma cells. J Exp Clin Cancer Res. 2018;37(1):9.
37. Krause GC, Lima KG, Levorse V, Haute GV, Gassen RB, Garcia MC, Pedrazza L, Donadio MVF, Luft C, de Oliveira JR. Exenatide induces autophagy and prevents the cell regrowth in HepG2 cells. Excli j. 2019;18:540–8.

Figures

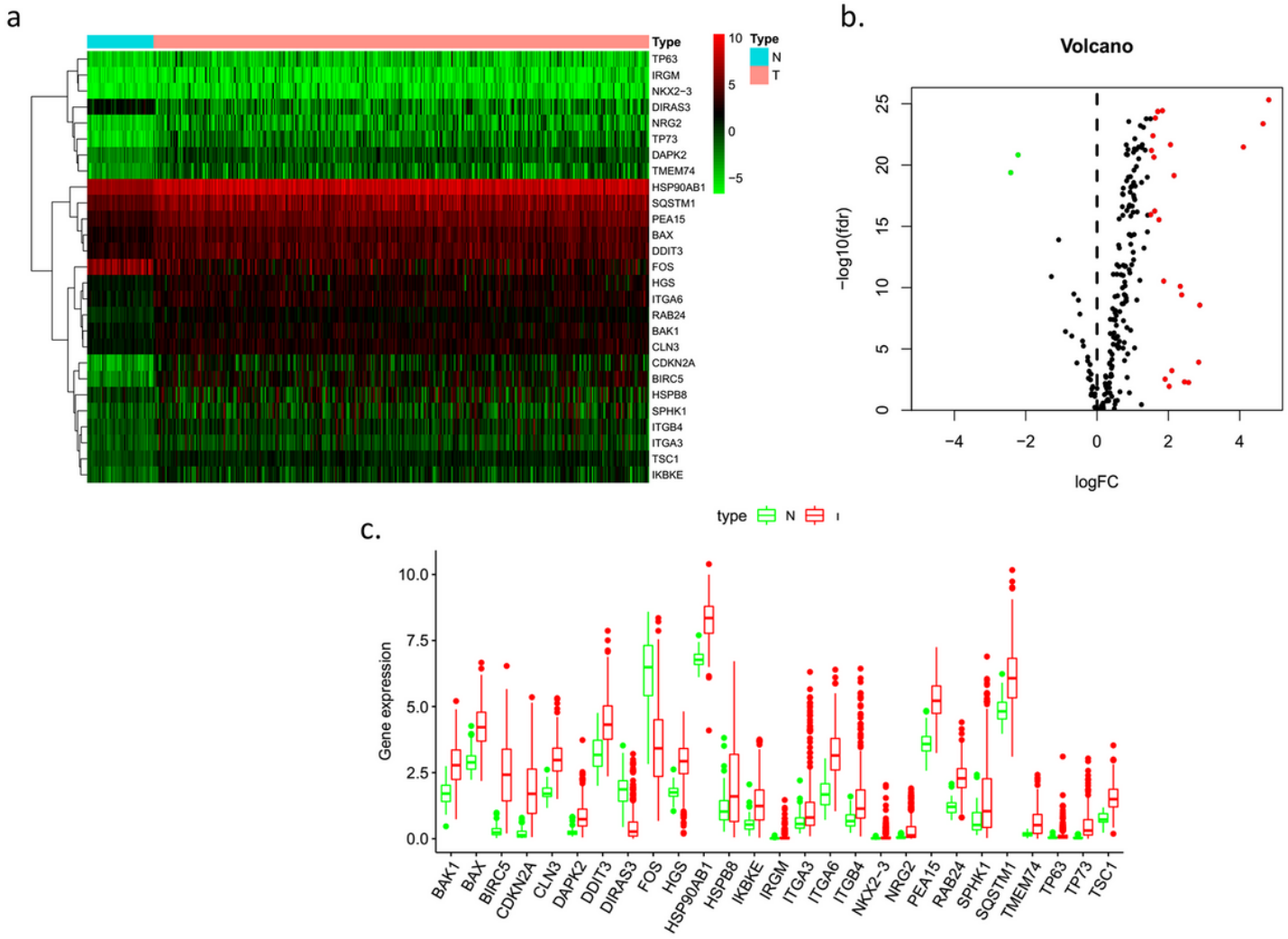
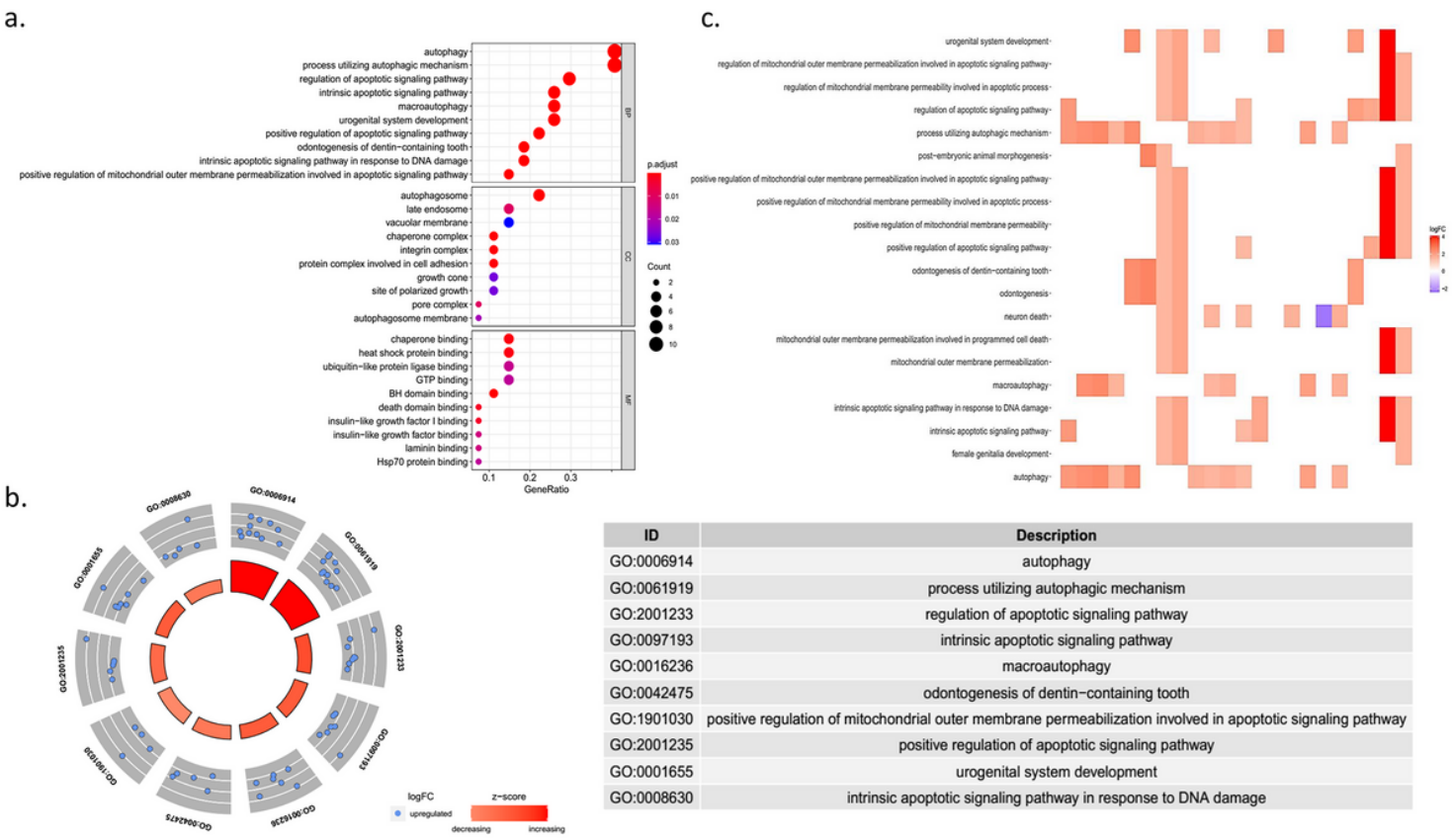


Figure 1

Differentially expressed ARGs in HCC and non-tumor tissues. a. The heatmap of differentially expressed ARGs between 374 HCC tissues and 50 non-tumor tissues; b. The volcano map of differentially expressed ARGs. The green dots mean down-regulated genes and the red dots represent up-regulated genes. c. The expression patterns of 27 differentially expressed ARGs in HCC and normal samples. The green box and red plots represent the normal samples and HCC, respectively.

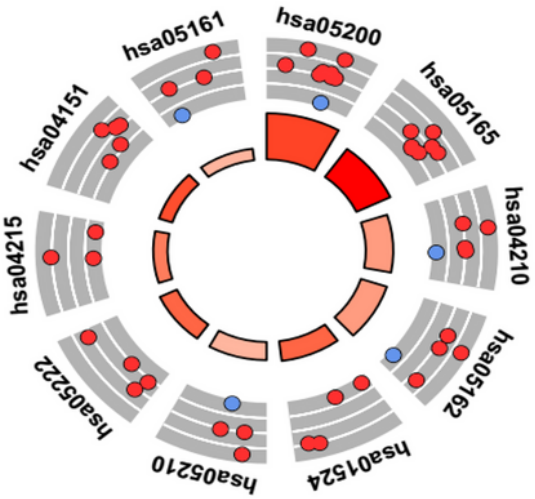


ID	Description
GO:0006914	autophagy
GO:0061919	process utilizing autophagic mechanism
GO:2001233	regulation of apoptotic signaling pathway
GO:0097193	intrinsic apoptotic signaling pathway
GO:0016236	macroautophagy
GO:0042475	odontogenesis of dentin−containing tooth
GO:1901030	positive regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway
GO:2001235	positive regulation of apoptotic signaling pathway
GO:0001655	urogenital system development
GO:0008630	intrinsic apoptotic signaling pathway in response to DNA damage

Figure 2

GO enrichment analysis of differentially expressed ARGs. a. Bubble plot of significant GO terms. b. GOCircle plot of significant GO terms. c. The relationship between the differentially expressed ARGs and GO terms is presented by a heatmap.

a.



ID	Description
hsa05200	Pathways in cancer
hsa05165	Human papillomavirus infection
hsa04210	Apoptosis
hsa05162	Measles
hsa01524	Platinum drug resistance
hsa05210	Colorectal cancer
hsa05222	Small cell lung cancer
hsa04215	Apoptosis – multiple species
hsa04151	PI3K–Akt signaling pathway
hsa05161	Hepatitis B

b.

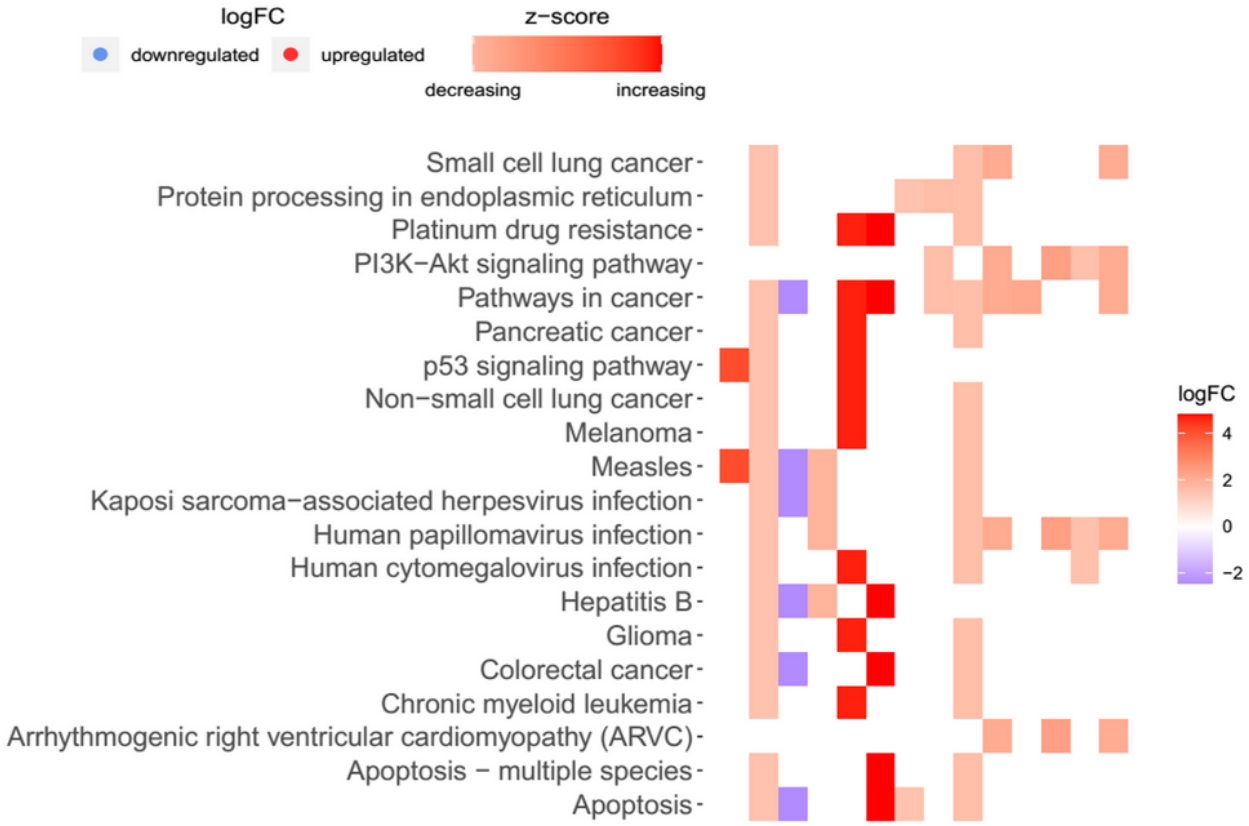


Figure 3

KEGG pathway analysis of differentially expressed ARGs. a. The significant KEGG terms of differentially expressed ARGs. b. The heatmap illustrates the relationships between the differentially expressed ARGs and KEGG pathways.

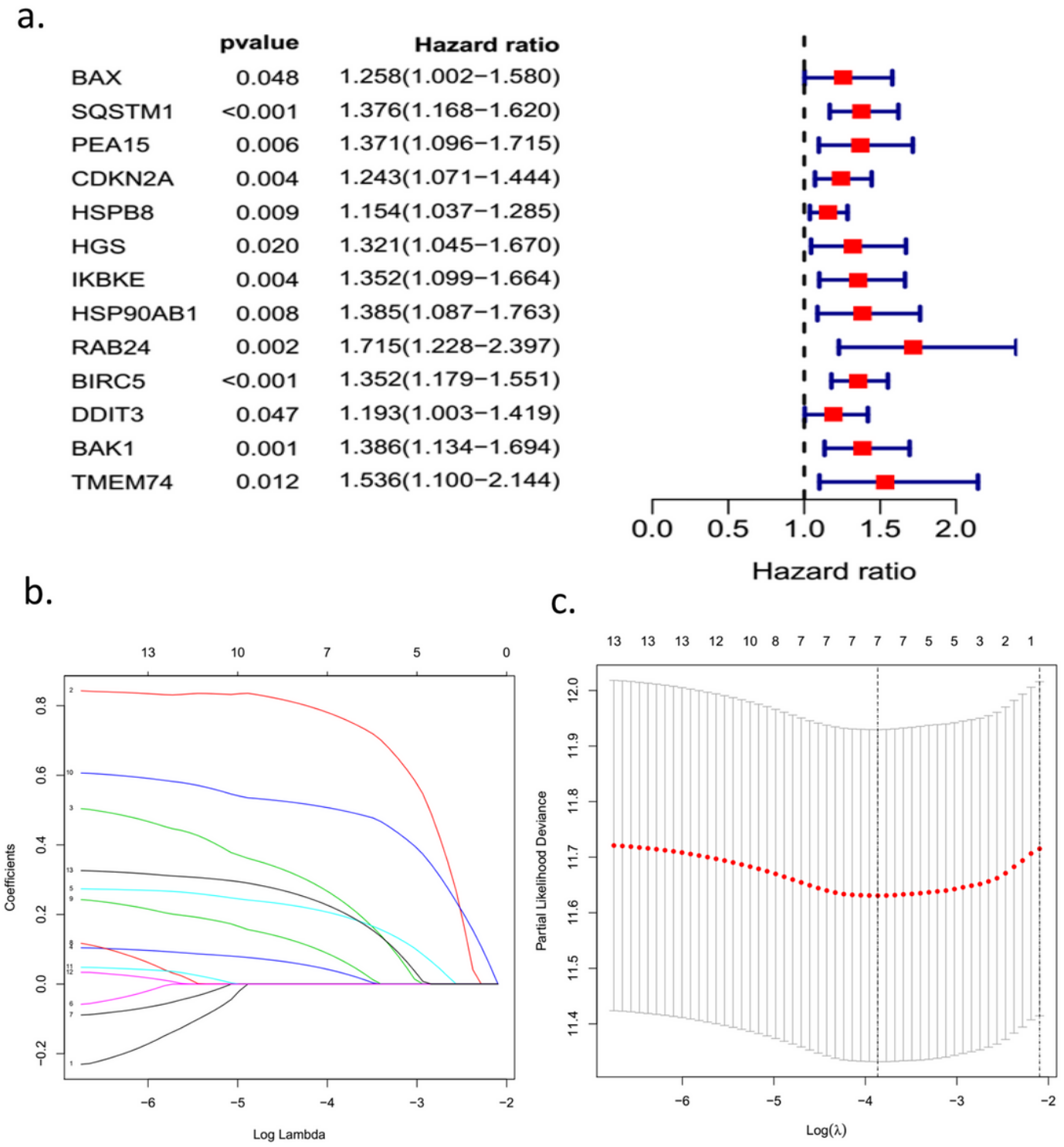


Figure 4

Predictive gene signature constructed by Lasso regression. a. Forest plot of ARGs associated with OS in HCC. b. Lasso coefficient profiles of the 13 ARGs in HCC. c. The optimal lambda value in Lasso model for HCC.

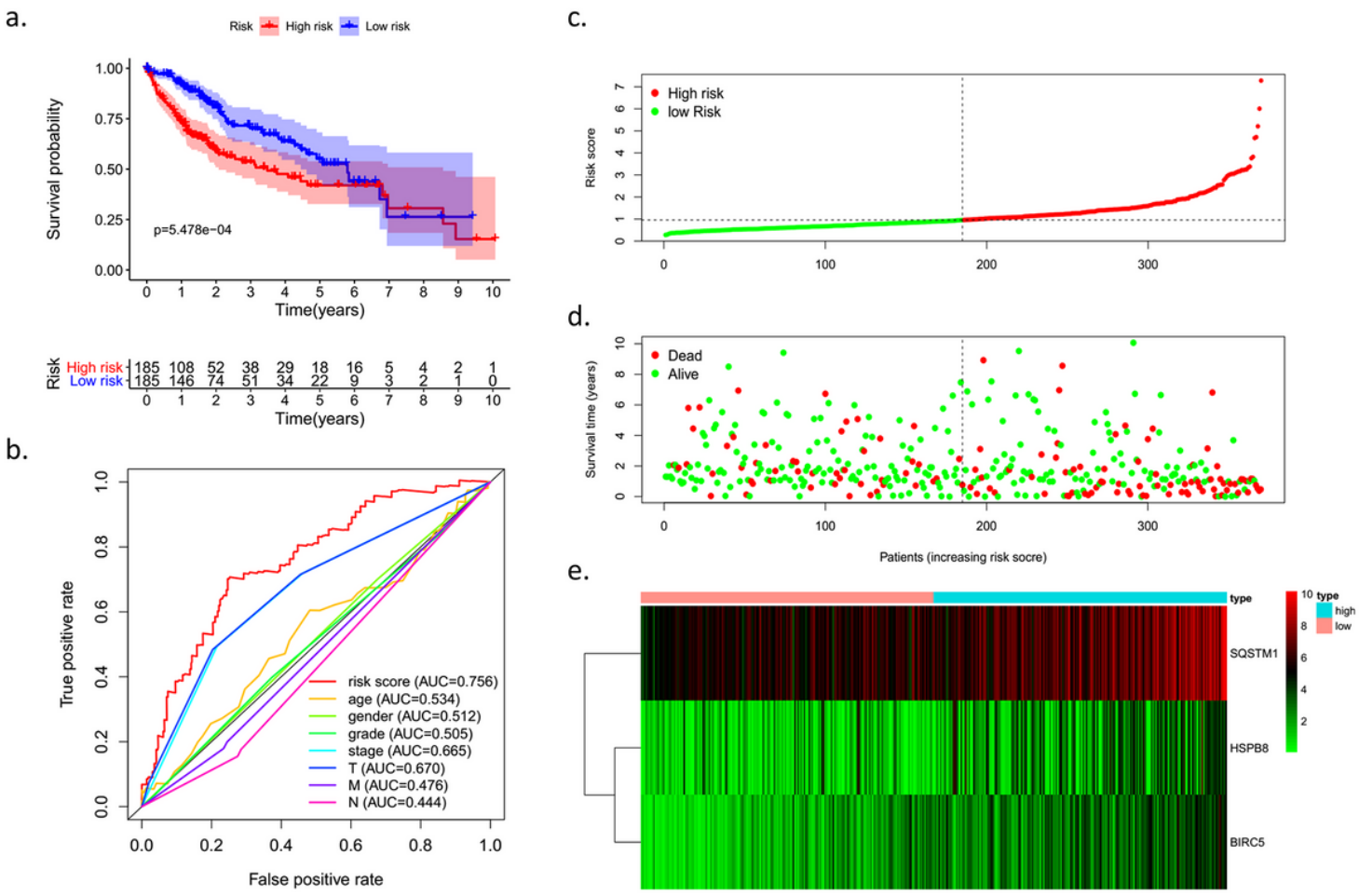


Figure 5

Prognostic prediction model of HCC patients. a. Kaplan-Meier curves show that HCC patients in low-risk group have better OS than those in high-risk group. b. ROC curves of different variable in the prognostic model. c. The distribution of risk scores of HCC patients in prognostic model. d. The distribution of HCC patients with different survival status. e. The heatmap of 3 risk genes (SQSTM1, HSPB8, and BIRC5) in HCC patients with different risk score.

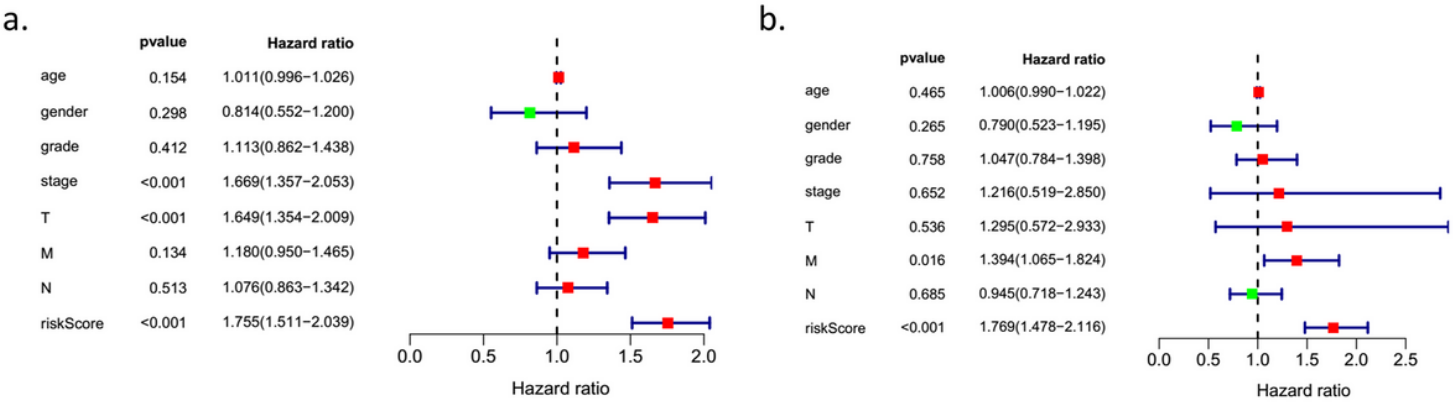


Figure 6

Independent risk factors of OS for HCC patients a. Assessment of the contribution of each factor to HCC survival by univariate Cox regression analysis. b. Assessment of the contribution of each factors to HCC survival by multivariate Cox regression analysis.

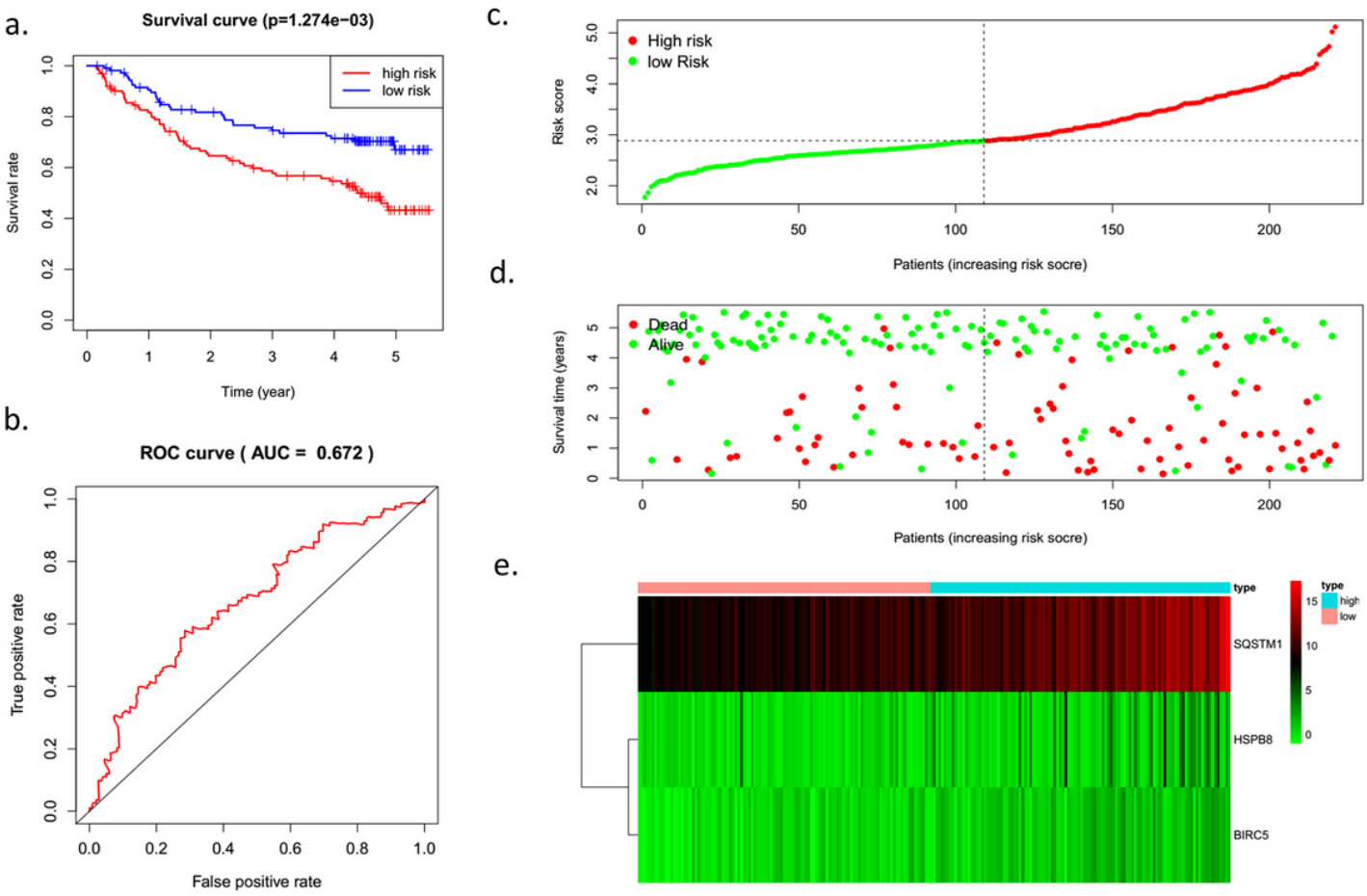


Figure 7

Validation of the risk signature in the testing cohort. a. Kaplan-Meier analysis shows that HCC patients with high risk score have poorer OS than those with low risk score; b. The ROC curve for assessing the prediction performance of the risk gene signature; c. The distribution of risk scores of HCC patients in prognostic model. d. The distribution of HCC patients with different survival status and survival time; e. The distributions gene expression profiles of SQSTM1, HSPB8, and BIRC5.

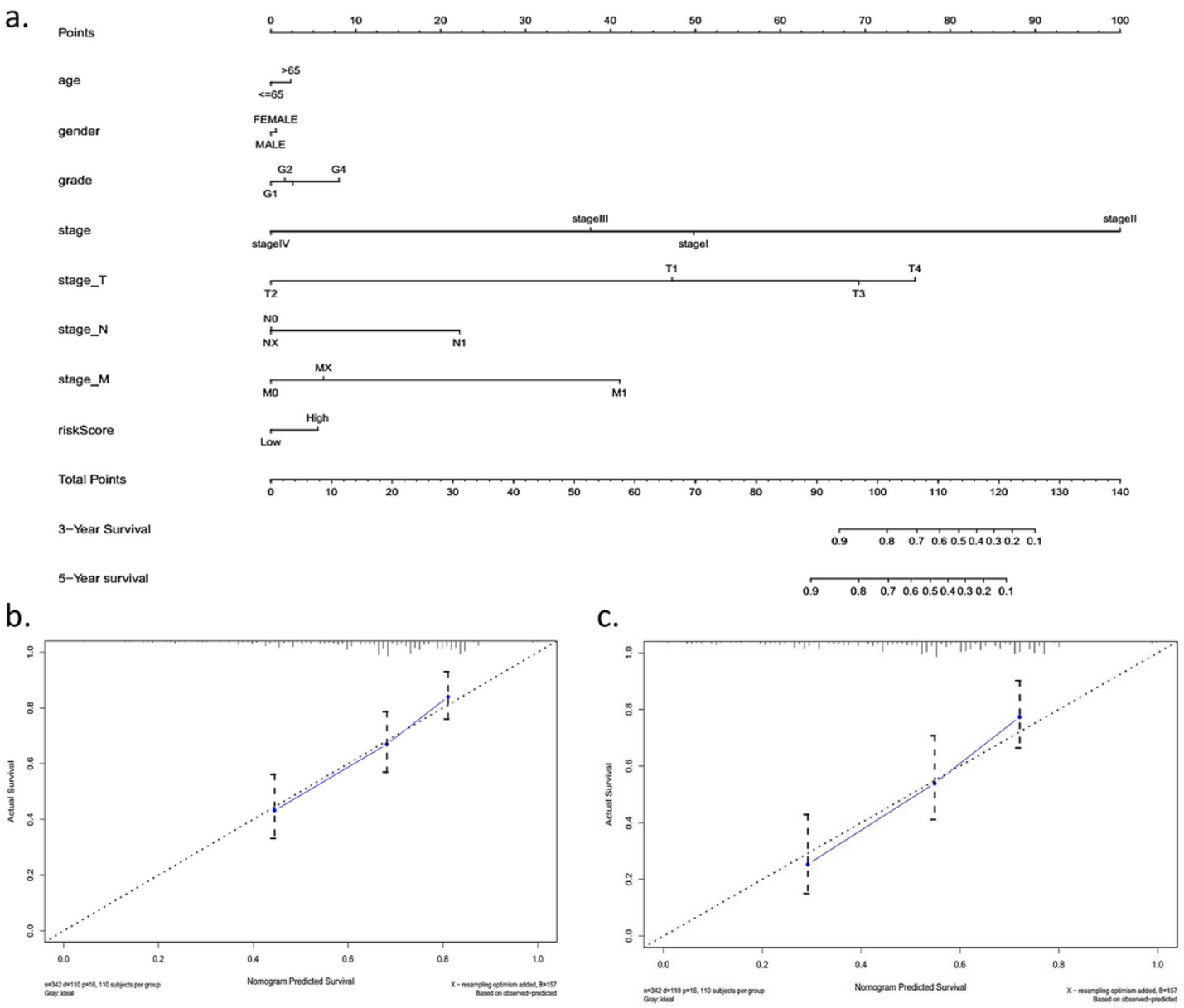


Figure 8

The nomogram for predicting survival rate of HCC. a. The clinical nomogram for predicting the 3-yerar and 5-year survival rate in HCC patients. b. The calibration curves present the concordance of 3-year survival between the observation and the prediction. c. The calibration curves present the concordance of 5-year survival between the observation and the prediction.

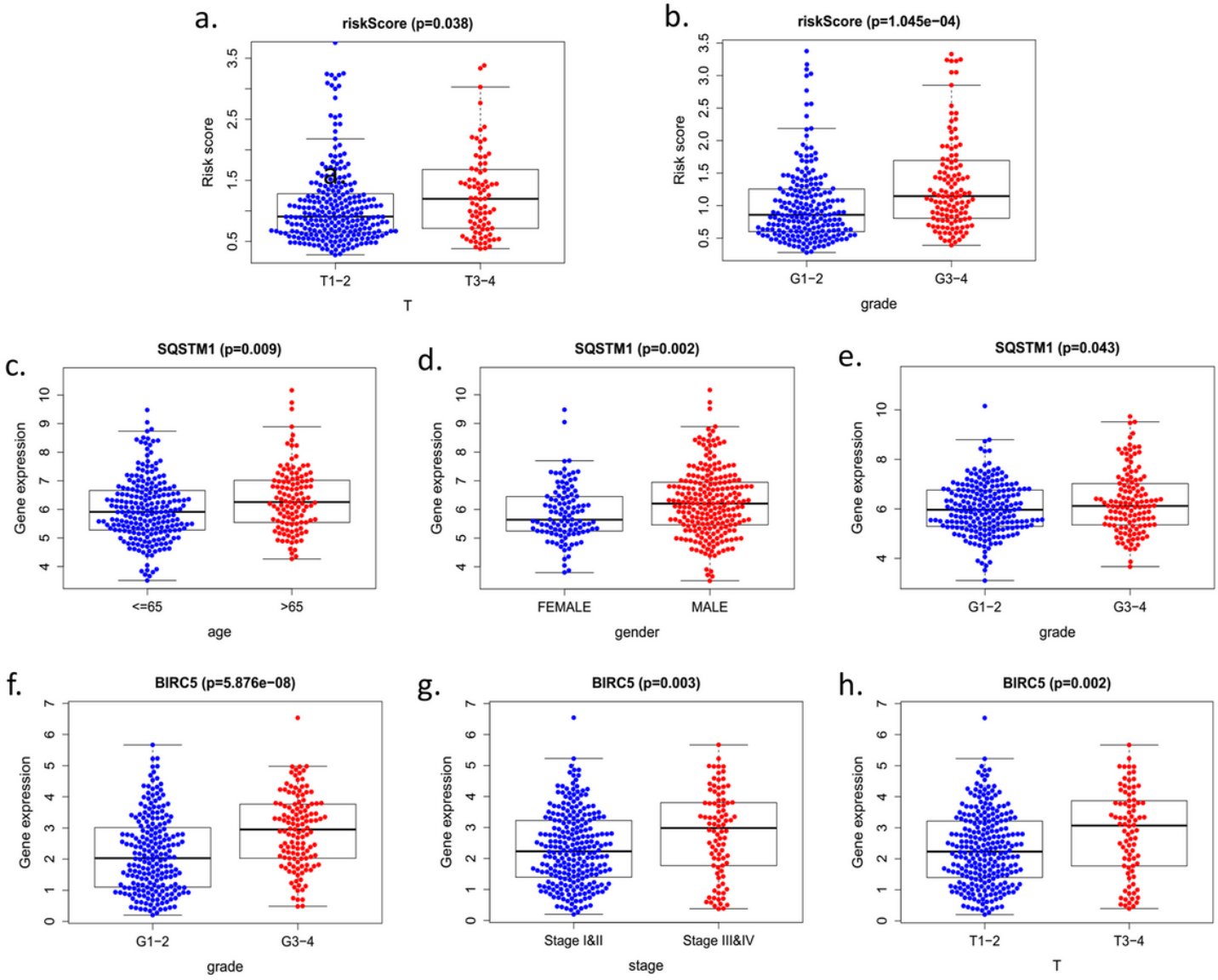


Figure 9

The relationships between ARGs and clinical factors. a. The risk scores increase along with the T (primary tumor). b. The risk scores increase along with tumor grade. c. The expression of SQSTM1 is higher in the groups of patients aged >65. d. The expression of SQSTM1 is higher in the groups of male. e. The expression of SQSTM1 is higher in the groups of patients with higher tumor grade. f-g. The level of BIRC5 is associated with the grade, tumor stage and T (primary tumor) in HCC patients.

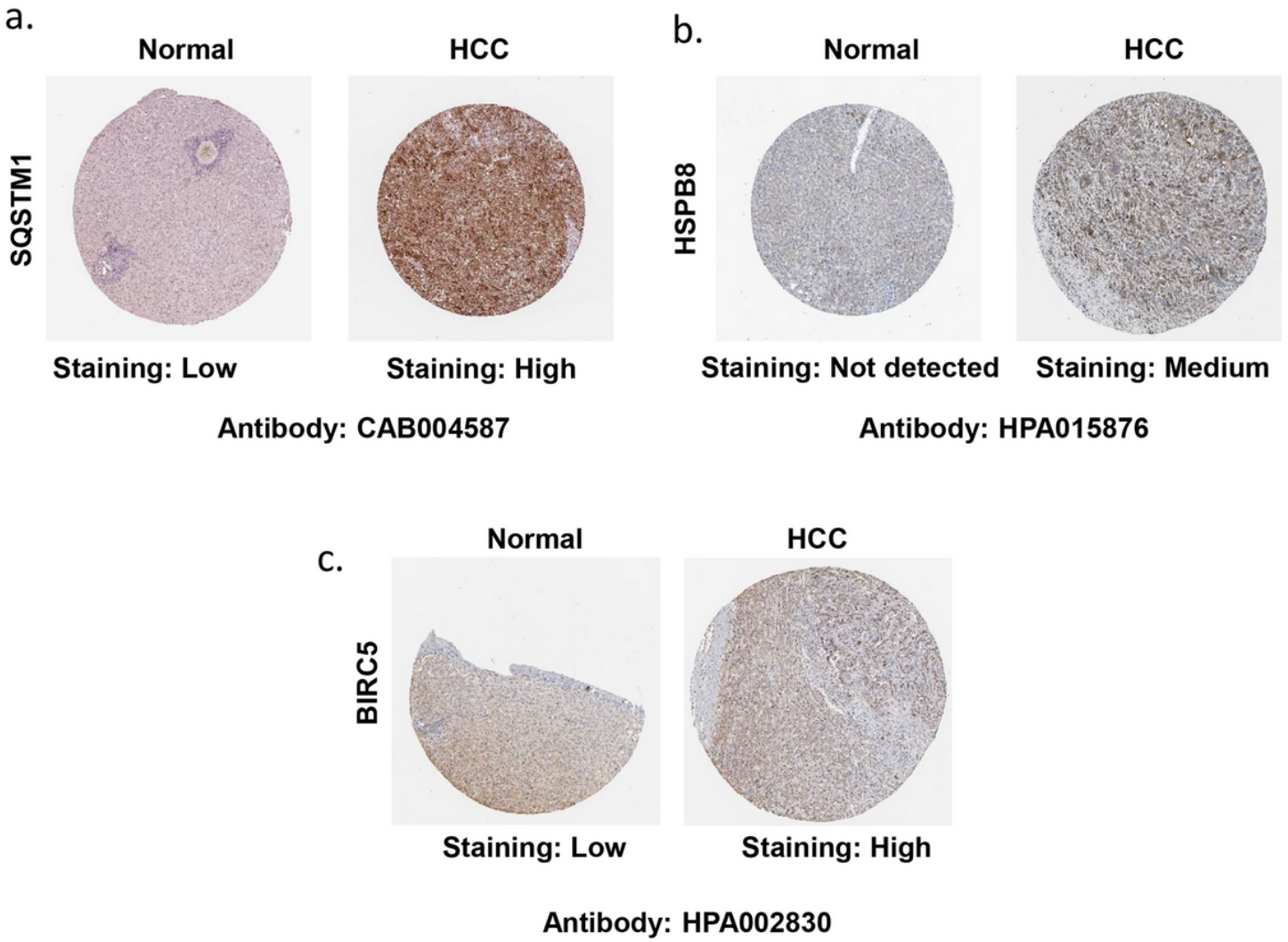


Figure 10

Protein levels of SQSTM1, HSPB8, and BIRC5 in HCC and normal tissues a. Immunohistochemistry results of SQSTM1 in normal tissue (staining: low; intensity: moderate; quantity: <25%; location: nuclear) and in HCC (staining: high; intensity: strong; quantity: >75%; location: cytoplasmic/membranous/nuclear). b. Immunohistochemistry results of HSPB8 in normal tissue (staining: not detected; intensity: negative; quantity: none; location: none) and in HCC (staining: medium; intensity: moderate; quantity: >75%; location: cytoplasmic/membranous). c. Immunohistochemistry results of BIRC5 in normal tissue (staining: low; intensity: weak; quantity: >75%; location: cytoplasmic/membranous) and in HCC (staining: high; intensity: strong; quantity: 75-25%; location: cytoplasmic/membranous/nuclear)

Figure 11

TFs-genes networks and miRNA-genes networks of SQSTM1, HSPB8, and BIRC5 a. The predicted networks of TFs and 3 ARGs (SQSTM1, HSPB8, and BIRC5); The red circles represent SQSTM1, HSPB8, and BIRC5; The blue squares represent the predicted TFs; b. The predicted networks of miRNAs and 3 ARGs (SQSTM1, HSPB8, and BIRC5). The red circles represent SQSTM1, HSPB8, and BIRC5; The blue squares represent the predicted miRNA.

Supplementary Files

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

- Supplementary.docx