A novel prognostic index model constructed with five autophagy-related genes may be a potential prognostic biomarker and therapeutic target in liver hepatocellular carcinoma

Qingqing Zhu

1 School of Medicine and Life Sciences, University of Jinan-Shandong Academy of Medical Sciences. 2 Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences

Jia Wang

1 Department of Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences. 2 Department of Oncology, Zibo Maternal and Child Health Hospital.

Menghan Liu

Basic Medicine College, Shandong First Medical University, Taian 271016, Shandong Province, China

Qiujing Zhang

School of Medicine and Life Sciences, University of Jinan-Shandong Academy of Medical Sciences, Jinan 250022, Shandong Province, China. Department of Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy o

Miaomiao Shen

Hematology of tumor, Chengwu Hospital Affiliated to Shandong First Medical University, Chengwu Hospital, Heze 274200, Shandong Province, China

Chao Xie

Department of Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan 250117, Shandong Province, China

Jianjun Zhang

Department of Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan 250117, Shandong Province, China

Shuai Fu

Department of Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan 250117, Shandong Province, China

Congcong Han

School of Medicine and Life Sciences, University of Jinan-Shandong Academy of Medical Sciences, Jinan 250022, Shandong Province, China. Department of Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy o

Bao Song
Basic Laboratory, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan 250117, Shandong Province, China
Fuxia Wang
Department of Oncology, Yuncheng County People's Hospital, Heze 274700, Shandong Province, China
Zining Liu
Department of Radiotherape, The Third Affiliated Hospital of Shandong First Medical University, Jinan 250031, Shandong Province, China
Jie Liu (wzzljsb123@126.com)
Department of Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan 250117, Shandong Province, China
https://orcid.org/0000-0002-4966-1753

Research

Keywords: liver hepatocellular carcinoma, autophagy-related gene, prognosis, biomarker, The Cancer Genome Atlas

Posted Date: May 11th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-27304/v1

Abstract

Background Early diagnosis and effective treatment of liver hepatocellular carcinoma (LIHC) are keys to improving the prognosis of patients. Increasing evidences clarify that autophagy-related genes (ARGs) make great differences to the generation and progression of LIHC, and may serve as prognostic biomarkers for LIHC.

Methods We randomly divided the LIHC patients in The Cancer Genome Atlas (TCGA) into the training and testing group. Next, use the training group to perform univariate Cox, LASSO and multivariate Cox analysis to construct our prognostic index (PI) model for LIHC; use the testing group and the whole TCGA set to make internal validations; use the International Cancer Genome Consortium to make external validations; and use the whole TCGA, GSE14520 and Oncomine to exam the expression patterns of the five ARGs. Then, we performed the ROC curve as well as univariate and multivariate analysis to evaluate the independent prognostic prediction power of the PI model, and made nomograms to estimate 1,3,5-year survival rate of LIHC patients. Besides, we conducted functional enrichment analyses of differentially expressed ARGs with GO, KEGG and GSEA, and made drug sensitivity analysis for the PI model via the GDSC database.

Results A novel PI model which was composed of five key ARGs (ATG9A , EIF2S1 , GRID1 , SAR1A and SQSTM1) succeeded to be constructed. All the internal and external validations testified that the PI model could well distinguish high-risk patients from low-risk ones, with AUC values > 0.60. Further comparison analysis showed that the PI model was no less than some common prognostic factors. People can estimate the 1,3,5-year survival rate of individual LIHC patient with the nomograms. Additionally, we obtained 62 differentially expressed ARGs and studied the potential mechanisms or pathways. Furthermore, we also found some potential targeted drugs associated to the five ARGs for LIHC patients.

Conclusions The novel five-ARGs PI model has great potential to serve as a diagnostic or prognostic biomarker and therapeutic target in LIHC, which may guide future clinical applications to some extent and improve the outcome of LIHC patients.

Background

In GLOBOCAN 2018, liver cancer occupied the 6th incidence (about 4.7%) and the 4th mortality (about 8.2%) all over the word, and more seriously, in males, its mortality was second only to lung cancer (1). Liver hepatocellular carcinoma (LIHC) is the primary pathological type, accounting for about 85% – 90% of liver cancer (2). Hepatectomy is the most important treatment for LIHC. However, because the early symptoms of LIHC are not obvious, mostly patients have turned into the late stage at the diagnosis, losing the opportunity of surgical resection (3). In addition, LIHC has low sensitivity to chemotherapy and radiotherapy, recurrence and metastasis being the main causes of patients' death (4). It's well acknowledged that early detection, diagnosis and effective treatment of LIHC are key to improving the

prognosis of patients. The most effective way to achieve this goal is to screen and monitor LIHC people at high risk. At present, imaging and three biomarkers, α-fetoprotein (AFP), des-γ-carboxy prothrombin and lens culinaris agglutinin-reactive fraction of AFP, play the main roles in the surveillance and diagnosis of LIHC (5, 6). But their sensitivity and specificity remain insufficient for clinicians. Therefore, to search for better biomarkers for the diagnosis and prognosis of LIHC is particularly important.

Recently, some researchers have found autophagy-related genes (ARGs) may better guide patients' prognosis in some cancers, such as clear-cell renal cell carcinoma, bladder cancer and serous ovarian cancer (7–9). Autophagy generally refers to the cellular metabolic process in which damaged or redundant large molecular components or organelles in cytoplasm are orderly transported to lysosomes for degradation (10). It is an essential biological process for cell survival and is beneficial to maintain cell homeostasis, closely linked to various diseases, including cancer (11, 12). In recent years, a large number of studies have shown that autophagy could play dual roles (inhibition or promotion) in the generation and progression of LIHC by regulating immunity, oxidative stress, cell homeostasis and so on (13, 14). It have been testified that some ARGs, proteins and pathways had a certain relationship with LIHC, and guided the diagnosis and prognosis of LIHC patients (15–17). Specially, Shanshan Wang et al. found that AFP may be able to inhibit autophagy to promote the proliferation, migration and invasion of LIHC via activating PI3K/AKT/mTOR signaling (18). However, those previous studies only examined a single ARG in LIHC, which merely reflected part of the autophagy function. Thus, we wondered whether some ARGs could combine to better guide the diagnosis and prognosis of LIHC patients. We aimed to carry on systematic and comprehensive analyses on the overall known ARGs so as to identify better prognostic biomarkers in LIHC, which has not been studied before.

We found the most detailed and latest list of ARGs from the Human Autophagy Database (HADb), including a total of 234 ARGs which have ever been directly or indirectly mentioned in literatures. Not long ago, Wang et al. identified three ARGs related to the prognosis of patients with bladder cancer among these 234 ARGs(7). Hence, in this study, we used similar approaches to study the prognostic value of these 234 ARGs in LIHC with the data of The Cancer Genome Atlas (TCGA) database, and conducted repeated filtration to identify the key ARGs in LIHC, which were further utilized to construct a novel prognostic index (PI) model to serve as a prognostic biomarker for individual LIHC patient. Meanwhile, we made repeated verification for the predictive performance of the PI model and pictured the nomograms to estimate the 1,3,5-year survival rate of individual LIHC patient. In addition, by functional enrichment analysis and drug sensitivity analysis, we made primary explorations on the potential mechanism and pathways as well as potential targeted drugs associated to the ARGs in LIHC. Our study may be helpful to early diagnose the prognosis and improve the survival outcomes in LIHC patients.

Methods

Data preparation

The main data were downloaded from the cohort labeled as LIHC in the TCGA dataset (http://cancergenome.nih.gov/), including the transcriptome data and patients' clinical information (19). It covered a total of 374 tumor samples and 50 normal samples. Then, we randomly divided the overall LIHC patients of the TCGA database into two parts with creatDataPartition R package: the training group and the testing group. Besides, we also obtained another set of the transcriptome data and clinical information of 243 LIHC patients from the International Cancer Genome Consortium (ICGC) database (https://icgc.org/), which was used as external validations. The list of the 234 ARGs was downloaded from the HADb website (http://www.autophagy.lu/clustering/index.html/). Then, we extracted the gene expression levels of the 234 ARGs in each sample, and arranged the expression levels of ARGs and the patients' OS into matrixes.

Prognostic index model construction based on ARGs

We used the training group to construct a PI model as a novel prognostic marker in LIHC. We firstly carried on univariate Cox regression analysis to primarily screened out candidate prognostic ARGs from the whole ARGs, which were statistically associated with the OS of LIHC patients ($p < 0.05$). Subsequently, the Least Absolute Shrinkage and Selector Operation (LASSO) algorithm was used to further filtrate prognosis-related ARGs. LASSO regression can make variable selection and adjust complexity, so as to better avoid overfitting of high-dimensional predictors (20). Ten-fold cross validation method was set to make penalty parameter turning and to determine the optimal parameter for the smallest partial likelihood deviance. Lastly, in order to ensure the accuracy and validity of the model, a multivariate Cox regression analysis was conducted to remove the ARGs that didn't belong to independently prognostic biomarkers, the rest key ARGs entering the construction of the PI model. The PI model was showed with a sum formula in which expression level of each prognostic ARG was multiplied by its regression coefficient.

Verification of the prognostic index model

First of all, we verified the expression pattern of the key ARGs in the PI model in LIHC compared to non-cancer liver tissues. It was finished in the GSE14520 database and the whole TCGA database. The samples of GSE14520 database came from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/), covering 220 non-tumor liver tissues and 225 LIHC tissues. Meanwhile, we also observed the differential expression of the key ARGs in many other cancers with the help of Oncomine database (http://www.oncomine.org) (21). The thresholds of p-value, fold change (FC) and gene rank were set as 1E-4, 2 and top 10%, respectively.

Next, we performed the internal validations (in the training group, the testing group and the whole TCGA database, respectively) and the external validations (in the ICGC database) for the PI model, all according to the following four steps. Firstly, according to the PI model formula, every LIHC patient got a risk score and divided into high-risk group or low-risk group, the median risk score set as the cutoff value. Specially, among the 243 LIHC patients from ICGC database, 64 were judged as low-risk and 179 as high-risk. Secondly, receiver operating characteristic (ROC) curve for one-year overall survival time (OS) with

survivalROC R package was plotted to assess the sensitivity and specificity of the PI model. The closer the ROC curve was to the upper left corner, the larger the AUC value, indicating that the PI model had a better prediction power on OS. Thirdly, Kaplan-Meier survival curve was drawn to judge the OS change trend between high-risk group and low-risk group. Fourthly, we also directly observed the association between patients' survival status, the expression patterns of prognostic ARGs and the risk score of LIHC patients.

The evaluation of the independent prognostic prediction power of the prognostic index model

For better assessing our PI model, we investigated the association between the risk score and some common clinicopathologic features in LIHC patients, including primary tumor staging (T), regional lymph node staging (N), distant metastasis staging (M), grade, stage, gender, age, survival state and survival time. The statistical method to identify the correlation between PI model and these clinicopathologic features was chisq.test. Then, we pictured multivariate ROC curves to quantificationally compare the prediction power of these prognostic markers. The higher the area under the ROC curve (AUC) value, the stronger the power. Finally, we also conducted univariate and multivariate independent prognostic analyses to judge whether they belonged to independent prognostic marker correlated with OS in LIHC. In addition, nomograms can be developed to quantize the prognosis of individual patient by integrating multiple risk factors (22). Here we added age, gender, grade, stage and risk score into the nomograms to individually estimate 1,3,5-year survival rate of LIHC patients.

Functional enrichment analyses for ARGs

First of all, we conducted differential expression analysis to screen out the differentially expressed ARGs between LIHC and normal tissue. It was realized by R programming language, the statistical method being Wilcoxon signed-rank test. The threshold of false discovery rate (FDR)for filer was set as 0.05, and that of FC was 2 (8). Next, functional enrichment analyses of differentially expressed ARGs were made with two methods: gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis. Here clusterProfiler R package was utilized (23). The screening conditions were p-value < 0.05 and q-value < 0.05. In GO analysis, all the GO terms were divided into 3 groups: biological processes (BP), cellular components (CC) and molecular function (MF). We also pictured the circle graph of the top ten GO terms with GOplot package (24). The procedure and screening conditions of KEGG enrichment analysis were basically the same as those of GO analysis. Both of the two methods played important roles in helping us to understand the dominating biological attributes of ARGs.

Furthermore, we carried onthe Gene Set Enrichment Analysis (GSEA, http://software.broadinstitute.org/gsea/) to explore the enrichment of KEGG pathways in which high and low risk groups. It mainly contains several key steps: calculation of Enrichment Score (ES), assessment of the significance of ES and corrective multiple hypothesis testing (25).

Drug sensitivity and resistance analysis

Aiming to uncover the potential targeted drugs in LIHC, we used the Drug Sensitivity in Cancer (GDSC) database (https://www.cancerrxgene.org/) to make the drug sensitivity and resistance analysis for the prognostic index model (26). Correlation coefficient (Cor) < 0 represented the resistance of LIHC cell lines to the drug increased, and Cor> 0 represented the sensitivity of LIHC cell lines to the drug increased. The cutoff of p value was 0.05.

Results

Construction of the five-ARGs prognostic index model

On the base of the training group, 222 ARGs were statistically associated with OS of LIHC patients in the univariate Cox regression analysis ($p < 0.05$), becoming candidate prognostic ARGs (**Additional file 1**). Subsequently, by the further filtration of a LASSO algorithm, only 16 ARGs may be related to prognosis (**Additional file 2** and **3**). Lastly, multivariate Cox regression analysis ultimately identified five independently prognostic key ARGs —— genes *ATG9A* (autophagy related 9A), *EIF2S1* (eukaryotic translation initiation factor 2 subunit alpha), *GRID1* (glutamate ionotropic receptor delta type subunit 1), *SAR1A* (secretion associated Ras related GTPase 1A) and *SQSTM1* (sequestosome 1) —— which made up the PI model (**Additional file 4**).

The formula of our PI model was: . The five key ARGs were all characterized by hazard ratio (HR)> 1 and the positive coefficient, indicating that they all belonged to the risk factors for LIHC patients.

Verification of the key ARGs and prognostic index model

There was no data on gene *GRID1* in the GSE14520 database. **Fig 1a-d** showed that in the whole TCGA database, the expression levels of genes *ATG9A*, *EIF2S1*, *SAR1A* and *SQSTM1* were all increased in LIHC compared to non-cancer liver tissues ($p < 0.05$). The verification results in GSE14520 database were generally consistent, except gene *ATG9A* being slightly decreased however without statistical significance ($p = 0.307$) (**Fig 1e-h**). The validation in Oncomine database showed that the expressions of the five key ARGs were also generally increased in many other cancers (**Additional file 5**).

In the ROC curve analysis, the AUC values of the PI model were 0.841, 0.724, 0.787 and 0.678 in the training group, the testing group, the whole TCGA database and the ICGC database, respectively (**Fig 1i-l**). It suggested that the PI model should possess a good sensitivity and specificity. Subsequently, in the Kaplan-Meier survival analysis, we all got statiscial difference no matter in the internal validations of the three TCGA sets or in the independently external validations of the ICGC database, revealing that low-risk group patients always had significantly longer OS than high-risk group ones ($p < 0.05$, **Fig 1m-p**). In addition, as expected, the higher the risk scores of LIHC patient, the lower the survival rate and the shorter the OS. And the expressions of the five prognostic ARGs were obviously increased in high-risk LIHC patients (**Fig 1q-t**). In conclusion, both the internal and external validations supported that our PI model manifested excellent predictive performance in predicting the survival of LIHC patients.

Prognostic index model being an excellent independent prognostic marker in LIHC

Fig 2a displayed the association between the PI model and some common clinicopathologic features in LIHC patients, including T, N, M, grade, stage, gender, age, survival state and survival time. It was showed that the PI model was significantly correlated to T (p = 0.016), stage (p = 0.011) and survival state (p = 0.004). It's worth noting that, in multivariate ROC curves analysis (**Fig 2b**), the AUC value of risk score (AUC = 0.792) was distinctly higher than that of age (AUC = 0.534), gender (AUC = 0.512), grade (AUC = 0.505), stage (AUC = 0.665), T (AUC = 0.670), N (AUC = 0.448) and M (AUC = 0.479). It revealed that using our PI model may improve the power to predict survival of LIHC patients when compared to using other clinicopathologic features. Finally, univariate independent prognostic analysis showed that the PI model, stage, T and M were significantly correlated to the patients' OS in LIHC (**Fig 2c**). Multivariate independent prognostic analysis showed that the PI model and M were significantly correlated to the patients' OS in LIHC (**Fig 2d**). Therefore, we determined that the PI model might be an excellent independent prognostic marker of OS in LIHC. In addition, **Fig 3a-b** were the nomograms of the training group and the testing group, which could be used to estimate the 1,3,5-year survival rate of individual LIHC patient with the data of age, gender, grade, stage and risk score.

Exploration on enriched function and pathways

Via differential analysis, we obtained 62 differentially expressed ARGs between LIHC and normal tissue in all (**Additional file 6**). Looking at the heat map (**Fig 4a**), the volcano figure (**Fig 4b**) and the box plot (**Fig 4c**), we found there were 4 ARGs down-regulated and 58 ARGs up-regulated. The 4 ARGs down-regulated were genes *DIRAS3*, *FOS*, *FOXO1* and *NRG1*.In GO analysis, the top 5 BP terms were all related to autophagy, the top 5 CC terms all related to membrane (autophagosome is a kind of vacuolar membrane) and the top 2 MF terms all related to kinase regulator activity (**Additional file 7**). As known in the circle graph of the top ten GO terms (**Fig 5a**), the top ten GO terms all had z-scores greater than zero, which meant the dominating biological attributes of the 62 differentially expressed ARGs were increased in LIHC. In brief, autophagy was upregulated in LIHC. In addition, KEGG analysis revealed that these differentially expressed ARGs focused on the pathways associated to the cell cycle, including autophagy and apoptosis (**Additional file 8**). Similarly, the z-scores of the top 10 KEGG terms were also greater than zero, corresponding functional pathways being upregulated in LIHC (**Fig 5b**).

Furthermore, GSEA analysis suggested that in TCGA1 database, there were 155 KEGG pathways upregulated in high risk group (**Additional file 9**) and 23 KEGG pathways upregulated in low risk group (**Additional file 10**). With FDR< 0.05 as criterion, 77 KEGG pathways in high risk group and one KEGG pathway in low risk group were significantly enriched. **Fig 5c** displayed the five leading-edge subsets (including thyroid cancer, oocyte meiosis, RNA degradation, ubiquitin mediated proteolysis and aminoacyl tRNA biosynthesis) in high risk group and one leading-edge subset (complement and coagulation cascades) in low risk group.

Exploration on potential sensitive or resistant drugs of LIHC cell lines

With the help of GDSC database and R package, we observed the association between the expression of the five key ARGs in the PI model and the sensitivities of LIHC cell lines to some common drugs. As the five key ARGs all belonged to risk factors, high expression of them would increase the resistance of LIHC cell lines to some drugs which were shown as red points in **Fig 6** or carried with Cor > 0 in **Table 1**, and increase the sensitivity to some drugs which were shown as green points in **Fig 6** or carried with Cor < 0 in **Table 1**. For example, high expression of gene *ATG9A* increased the resistance of LIHC cell lines to HG-5-113-01 and HG-5-88-01, and increased the sensitivity to OSI-930, SL 0101-1, selumetinib, EHT 1864, CCT018159, trametinib, selumetinib and RDEA119 ($p < 0.05$).

Discussion

LIHC is always a chief lethal malignant tumor worldwide (1).Early detection, diagnosis and effective treatment of LIHC are key to improving the prognosis of patients. To look for better biomarker for the diagnosis and prognosis of LIHC has been a research hotspot. More and more evidences clarify that autophagy makes a great difference to the generation and progression of carcinoma. As previously mentioned, by bioinformatics analyses for the high-throughput sequencing data of the whole ARGs, several essential ARGs could be screened out and make up a novel prognostic model, which might better guide the prognosis of patients (7–9). Previous studies have also testified that some individual ARGs were related to the prognosis of LIHC patients (15–17). Thus, we have reasons to believe that, with the similar bioinformatics analysis for the global ARGs, we could construct an autophagy-related prognostic index model which may better help clinicians to monitor the prognosis of LIHC patients. This method was first applied in LIHC but theoretically feasible.

Under the convenience of the big data era, we acquired the transcriptome data and clinical information a total of 424 LIHC patients from TCGA database, which was randomly divided into the training group and the testing group. With a series of filtration (Univariate Cox regression, LASSO regression and multivariate COX regression analysis) for the training group, we ultimately identified 5 independently prognostic ARGs (genes *ATG9A*, *EIF2S1*, *GRID1*, *SAR1A* and *SQSTM1*) to made up our PI model. Subsequently, we conducted repeated validation on the expression patterns of the five key ARGs in the PI model via the testing group, the TCGA database, GEO database and the Oncomine database. Subsequently, we observed the expression patterns of the five key ARGs in the PI model via the testing group, the whole TCGA database, GEO database and the Oncomine database, and conducted the internal and external validations for performance of the PI model. In the internal validations with the training group, the testing group and the whole TCGA database, the AUC value of the PI model was 0.724–0.841; and in the independently external validations of the ICGC database, the AUC value of the PI model was 0.678: Therefore, we thought there was a good sensitivity and specificity of the PI model for predicting prognosis of LIHC patients. In the Kaplan-Meier survival analysis, we all got statistical difference in both the internal and external validations, showing that low-risk LIHC patients had significantly longer OS than high-risk ones. Meanwhile, the higher the risk scores of LIHC patient, the lower the survival rate and the shorter the OS. Further studies showed that the PI model may be an excellent independent prognostic marker in LIHC, with a significantly higher AUC value (AUC = 0.792) than other clinicopathologic features.

Therefore, we determined that the PI model might be a potential prognostic marker for LIHC. In addition, in the model formula, the coefficients of the 5 prognostic ARGs were all positive, indicating that they all belonged to the risk factors for LIHC patients. In other words, when the expression levels of the 5 prognostic ARGs were increased, the LIHC patient would be likely to have a poor prognosis. The preceding validation also showed that the five prognostic ARGs were generally high expressed in high-risk LIHC patients. Thus, we regarded genes *ATG9A*, *EIF2S1*, *GRID1*, *SAR1A* and *SQSTM1* as oncogenes of LIHC.

Besides, we made primary exploration on the potential mechanism and pathways of ARGs in LIHC. Via differential analysis between tumor and normal samples, we obtained 62 differentially expressed ARGs, four genes being down-regulated and 58 ARGs up-regulated. The results of functional enrichment analysis for the 62 differentially expressed ARGs showed that most of the dominating GO and KEGG terms were directly or indirectly involved in autophagic mechanisms or pathways. In addition, the circle plots indicated that all the dominating biological attributes of the 62 differentially expressed ARGs were increased in LIHC. Therefore, we speculated that autophagy was likely to be increased in LIHC. It's worth noting that, except autophagy, these differentially expressed ARGs were yet in touch with apoptosis, longevity regulating pathway and other pathways. Under normal circumstances, autophagy is a mechanism to resist the occurrence of cancer. The mechanism is that autophagy can remove damaged and unnecessary macromolecules or organelles, reduce intracellular pressure and stabilize the genome, so as to reduce the cancerization. Moreover, autophagy can also eliminate cancerous cells through mediating apoptosis and immune response (11, 27). Yet, once the carcinogenesis occurs, the role of autophagy would turn into promote the development of tumors. The main mechanism is that the proliferation of cancer cells is accompanied by the increase of nutritional requirements, so that autophagy will be activated to provide tumor cells with nutrients and energy (28). Meanwhile, autophagy can enhance the chemoradiotherapy resistance of tumor cells (29). With our results, we guessed that autophagy may play a role in promoting the developing of LIHC, or the existence of LIHC may enhance the autophagy. Furthermore, we also made exploration on potential sensitive or resistant drugs of LIHC. However, the roles of autophagy in LIHC remain unclear and to be explored further.

Autophagy is mediated by autophagic vacuole, the formation of which relies on a series of autophagy-related proteins (30). Among them, autophagy-related protein 9A (ATG9A protein) is the unique transmembrane protein (31). Under normal circumstances, ATG9A protein chiefly exists in trans-Golgi network and endosomal system, whist it instantly partially localizes to the autophagy membranes after the induction of autophagy (32, 33). ATG9A protein plays crucial roles in mediating membrane trafficking from the recycling endosomes to the autophagosome (34). It has been found that the expression of gene *ATG9A* was increased in LIHC cells compared to normal human liver epithelial cell (35). Kunanopparat et al. attributed it to the autophagy activation of tumor cells under innutrition (35). It is necessary to conduct further mechanism research about the actions of gene *ATG9A* in LIHC autophagy.

Gene *EIF2S1* (also called as *EIF2A*) is a kind of protein coding gene, coding EIF2S1 protein. EIF2S1 protein is the alpha subunit of EIF2, phosphorylation of EIF2S1 subunit making the action of EIF2 inhibited (36). It was reported that EIF2A may be involved in the regulation of autophagy in yeast and

mammals (37). Upon various stress stimulations, such as oxygen deficiency and endoplasmic reticulum stress, EIF2AK4-EIF2A-ATF4 pathway would be activated, then EIF2A be phosphorylated, transcription factor ATF4 bind to the promoter of some ARGs, thereby activating their transcription (38). In addition, studies havefound that endoplasmic reticulum stress may activate the formation of autophagosome with LC3 conversion via PERK-EIF2A pathway (39). Above all, gene *EIF2S1* indeed plays an important role in autophagy, however the roles in autophagy of LIHC remain unclear. Combined with our results, we speculated that the phenomenon that the increasing expression of gene *EIF2S1* was associated with a poor prognosis of LIHC patients may be put down to the launch of autophagy by cancer cells.

Gene *GRID1* is located on chromosome 10q23 and encodes glutamate receptor δ1, a subunit of glutamate receptor channels. glutamate receptor δ1 is mainly distributed in plasma membrane, acting as an excitatory neurotransmitter of many synaptic transmission in the central nervous system(40).It was reported that *GRID1* was associated with schizophrenia, bipolar I disorder and Rett Syndrome (41, 42). *GRID1* was involved in MECP2 pathway, peptide ligand-binding receptors pathway and associated Rett Syndrome pathway. MF annotations of GO related to gene GRID1 mainly included glutamate receptor activity, ionotropic glutamate receptor activity and transmitter-gated ion channel activity involved in regulation of postsynaptic membrane potential. However, there is a lack of researches on the association between gene *GRID1*, autophagy and LIHC.

Gene *SAR1A* is also a protein coding gene, located on chromosome 10q22. GO annotations related to gene *SAR1A* mainly included GTP binding and obsolete signal transducer activity. It was reported that gene *SAR1A* was involved in the transportation between the endoplasmic reticulum and the golgi apparatus (43). Petrosyan et al. thought that downregulation of gene *SAR1A* is the key event leading to ethanol-induced Golgi fragmentation in hepatocytes (44).Likewise, there is a lack of researches on the association between gene *SAR1A*, autophagy and LIHC.

Gene *SQSTM1* has been widely studied, with a number of alternative names, such as *P62, A170, DMRV, OSIL, PDB3, ZIP3, p62B, NADGP, FTDALS3* and so on. This gene encodes a multifunctional protein (SQSTM1 protein),which is an adaptor protein to bind ubiquitin and LC3, and selectively targets polyubiquitinated cargo to the autophagosome(45).It was studied that, in autophagy-defective tumor cells, P62/SQSTM1 was preferentially accumulated, and promoted tumor growth by cooperating with autophagy-deficiency(46). Combined with our results, we speculated that the increasing expression of gene *SQSTM1*may lead to tumor deterioration by selective autophagy pathways, and be linked to a poor prognosis of LIHC patients. The specific mechanisms remain to be further studied.

There also exist some limitations in this study. First, our results were calculated on basis of the public TCGA database, and needed to be verified in further prospective clinical trials, that was the main limitation of our study. Second, it required further investigation on the mechanisms of the PI model in the initiation and progression of LIHC, which was limited by limited financial support but was our later research direction.

Conclusions

All in all, the novel five-ARGs PI model has great potential to serve as a diagnostic or prognostic biomarker and therapeutic target in LIHC, improving the outcome of LIHC patients. It may guide clinical applications to some extent in the future, and remain to need further investigated and substantiated.

Abbreviations

AFP: α-fetoprotein; ARGs: autophagy-related genes; ATG9A: autophagy related 9A; AUC: the area under the ROC curve; BP: biological processes; CC: cellular components; Cor: correlation coefficient; EIF2S1: eukaryotic translation initiation factor 2 subunit alpha; ES: Enrichment Score; FC: fold change; FDR: false discovery rate; GDSC: Drug Sensitivity in Cancer; GEO: Gene Expression Omnibus; GO: gene ontology; GRID1: glutamate ionotropic receptor delta type subunit 1; GSEA: Gene Set Enrichment Analysis; HADb: Human Autophagy Database;HR: hazard ratio; ICGC: International Cancer Genome Consortium; KEGG: Kyoto Encyclopedia of Genes and Genomes; LASSO: Least Absolute Shrinkage and Selector Operation; LIHC: liver hepatocellular carcinoma; M: distant metastasis staging; MF: molecular function; N: regional lymph node staging; OS: overall survival time; PI: prognostic index; ROC: receiver operating characteristic; SAR1A: secretion associated Ras related GTPase 1A; SQSTM1: sequestosome 1; T: primary tumor staging; TCGA: The Cancer Genome Atlas.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

The data generated and/or analyzed during the current study are included in this published article and its additional files. The initial data are available in the public databases (including TCGA, GEO, ICGC databases and so on), or available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by grants from the Key Research and Development Program of Shandong Province, China (GG201710060039).

Authors' contributions:

QQZ and JL contributed the idea of the design of the study; JW and ZNL performed data analysis and curation; MHL, QJZ, MMS prepared the figures and tables; CX, JJZ and SF took part in the analysis for the results; QQZ and JW drafted the manuscript; CCH, BS and FXW reviewed and revised the manuscript. All authors have read and approved the final manuscript to be submitted.

Acknowledgements

We are thankful for TCGA because the results published here are mostly based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

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: A Cancer Journal for Clinicians. 2018 Nov;68(6):394–424.
2. McGlynn KA, Petrick JL, London WT. Global epidemiology of hepatocellular carcinoma: an emphasis on demographic and regional variability. Clin Liver Dis. 2015 May;19(2):223–38.
3. European Association for the Study of the Liver. Electronic address: easloffice@easloffice.eu, European Association for the Study of the Liver. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. J Hepatol. 2018;69(1):182–236.
4. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. CA Cancer J Clin. 2016 Apr;66(2):115–32.
5. Daniele B, Bencivenga A, Megna AS, Tinessa V. Alpha-fetoprotein and ultrasonography screening for hepatocellular carcinoma. Gastroenterology. 2004 Nov;127(5 Suppl 1):108–12.
6. Marrero JA, Feng Z, Wang Y, Nguyen MH, Befeler AS, Roberts LR, et al. Alpha-fetoprotein, des-gamma carboxyprothrombin, and lectin-bound alpha-fetoprotein in early hepatocellular carcinoma. Gastroenterology. 2009 Jul;137(1):110–8.
7. Wang S-S, Chen G, Li S-H, Pang J-S, Cai K-T, Yan H-B, et al. Identification and validation of an individualized autophagy-clinical prognostic index in bladder cancer patients. Onco Targets Ther. 2019;12:3695–712.
8. Wan B, Liu B, Yu G, Huang Y, Lv C. Differentially expressed autophagy-related genes are potential prognostic and diagnostic biomarkers in clear-cell renal cell carcinoma. Aging (Albany NY). 2019 Oct 17;11.
9. An Y, Bi F, You Y, Liu X, Yang Q. Development of a Novel Autophagy-related Prognostic Signature for Serous Ovarian Cancer. J Cancer. 2018;9(21):4058–71.
10. Levine B, Klionsky DJ. Development by self-digestion: molecular mechanisms and biological functions of autophagy. Dev Cell. 2004 Apr;6(4):463–77.

11. Levine B, Kroemer G. Autophagy in the pathogenesis of disease. Cell. 2008 Jan 11;132(1):27–42.
12. Ueno T, Komatsu M. Autophagy in the liver: functions in health and disease. Nat Rev Gastroenterol Hepatol. 2017;14(3):170–84.
13. Huang F, Wang B-R, Wang Y-G. Role of autophagy in tumorigenesis, metastasis, targeted therapy and drug resistance of hepatocellular carcinoma. World J Gastroenterol. 2018 Nov 7;24(41):4643–51.
14. Tian Y, Kuo C-F, Sir D, Wang L, Govindarajan S, Petrovic LM, et al. Autophagy inhibits oxidative stress and tumor suppressors to exert its dual effect on hepatocarcinogenesis. Cell Death Differ. 2015 Jun;22(6):1025–34.
15. Shen M, Lin L. Functional variants of autophagy-related genes are associated with the development of hepatocellular carcinoma. Life Sci. 2019 Oct 15;235:116675.
16. Lin C-W, Chen Y-S, Lin C-C, Lee P-H, Lo G-H, Hsu C-C, et al. Autophagy-related gene LC3 expression in tumor and liver microenvironments significantly predicts recurrence of hepatocellular carcinoma after surgical resection. Clin Transl Gastroenterol. 2018 02;9(6):166.
17. Wu D-H, Wang T-T, Ruan D-Y, Li X, Chen Z-H, Wen J-Y, et al. Combination of ULK1 and LC3B improve prognosis assessment of hepatocellular carcinoma. Biomed Pharmacother. 2018 Jan;97:195–202.
18. Wang S, Zhu M, Wang Q, Hou Y, Li L, Weng H, et al. Alpha-fetoprotein inhibits autophagy to promote malignant behaviour in hepatocellular carcinoma cells by activating PI3K/AKT/mTOR signalling. Cell Death Dis. 2018;09(10):1027. 9(.
19. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn). 2015;19(1A):A68–77.
20. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med. 1997 Feb;28(4):385–95. 16(.
21. Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Varambally R, Yu J, Briggs BB, et al. Oncomine 3.0: Genes, Pathways, and Networks in a Collection of 18,000 Cancer Gene Expression Profiles. Neoplasia. 2007 Feb;9(2):166–80.
22. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. Lancet Oncol. 2015 Apr;16(4):e173–80.
23. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012 May;16(5):284–7.
24. Walter W, Sánchez-Cabo F, Ricote M. GOplot: an R package for visually combining expression data with functional analysis. Bioinformatics. 2015 Sep 1;31(17):2912–4.
25. Hung J-H, Yang T-H, Hu Z, Weng Z, DeLisi C. Gene set enrichment analysis: performance evaluation and usage guidelines. Brief Bioinformatics. 2012 May;13(3):281–91.
26. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013 Jan;41(Database issue):D955–61.

27. Takamura A, Komatsu M, Hara T, Sakamoto A, Kishi C, Waguri S, et al. Autophagy-deficient mice develop multiple liver tumors. Genes Dev. 2011 Apr 15;25(8):795–800.
28. Galluzzi L, Pietrocola F, Bravo-San Pedro JM, Amaravadi RK, Baehrecke EH, Cecconi F, et al. Autophagy in malignant transformation and cancer progression. EMBO J. 2015 Apr;1(7):856–80. 34(.
29. Wang X, Deng Q, Feng K, Chen S, Jiang J, Xia F, et al. Insufficient radiofrequency ablation promotes hepatocellular carcinoma cell progression via autophagy and the CD133 feedback loop. Oncol Rep. 2018 Jul;40(1):241–51.
30. Shibutani ST, Yoshimori T. A current perspective of autophagosome biogenesis. Cell Res. 2014 Jan;24(1):58–68.
31. Young ARJ, Chan EYW, Hu XW, Köchl R, Crawshaw SG, High S, et al. Starvation and ULK1-dependent cycling of mammalian Atg9 between the TGN and endosomes. J Cell Sci. 2006 Sep 15;119(Pt 18):3888–900.
32. Orsi A, Razi M, Dooley HC, Robinson D, Weston AE, Collinson LM, et al. Dynamic and transient interactions of Atg9 with autophagosomes, but not membrane integration, are required for autophagy. Mol Biol Cell. 2012 May 15;23(10):1860–73.
33. He C, Song H, Yorimitsu T, Monastyrska I, Yen W-L, Legakis JE, et al. Recruitment of Atg9 to the preautophagosomal structure by Atg11 is essential for selective autophagy in budding yeast. J Cell Biol. 2006 Dec 18;175(6):925–35.
34. Imai K, Hao F, Fujita N, Tsuji Y, Oe Y, Araki Y, et al. Atg9A trafficking through the recycling endosomes is required for autophagosome formation. J Cell Sci. 2016 Oct 15;129(20):3781–91.
35. Kunanopparat A, Kimkong I, Palaga T, Tangkijvanich P, Sirichindakul B, Hirankarn N. Increased ATG5-ATG12 in hepatitis B virus-associated hepatocellular carcinoma and their role in apoptosis. World J Gastroenterol. 2016 Oct;7(37):8361–74. 22(.
36. Donnelly N, Gorman AM, Gupta S, Samali A. The eIF2α kinases: their structures and functions. Cell Mol Life Sci. 2013 Oct;70(19):3493–511.
37. Tallóczy Z, Jiang W, Virgin HW, Leib DA, Scheuner D, Kaufman RJ, et al. Regulation of starvation- and virus-induced autophagy by the eIF2alpha kinase signaling pathway. Proc Natl Acad Sci USA. 2002 Jan;8(1):190–5. 99(.
38. Bretin A, Carrière J, Dalmasso G, Bergougnoux A, B'chir W, Maurin A-C, et al. Activation of the EIF2AK4-EIF2A/eIF2α-ATF4 pathway triggers autophagy response to Crohn disease-associated adherent-invasive Escherichia coli infection. Autophagy. 2016 Mar 17;12(5):770–83.
39. Moretti L, Cha YI, Niermann KJ, Lu B. Switch between apoptosis and autophagy: radiation-induced endoplasmic reticulum stress? Cell Cycle. 2007 Apr 1;6(7):793–8.
40. Yamazaki M, Araki K, Shibata A, Mishina M. Molecular cloning of a cDNA encoding a novel member of the mouse glutamate receptor channel family. Biochem Biophys Res Commun. 1992 Mar 16;183(2):886–92.

41. Fallin MD, Lasseter VK, Avramopoulos D, Nicodemus KK, Wolyniec PS, McGrath JA, et al. Bipolar I disorder and schizophrenia: a 440-single-nucleotide polymorphism screen of 64 candidate genes among Ashkenazi Jewish case-parent trios. Am J Hum Genet. 2005 Dec;77(6):918–36.
42. Livide G, Patriarchi T, Amenduni M, Amabile S, Yasui D, Calcagno E, et al. GluD1 is a common altered player in neuronal differentiation from both MECP2-mutated and CDKL5-mutated iPS cells. Eur J Hum Genet. 2015 Feb;23(2):195–201.
43. Takai Y, Sasaki T, Matozaki T. Small GTP-binding proteins. Physiol Rev. 2001 Jan;81(1):153–208.
44. Petrosyan A, Cheng P-W, Clemens DL, Casey CA. Downregulation of the small GTPase SAR1A: a key event underlying alcohol-induced Golgi fragmentation in hepatocytes. Sci Rep [Internet]. 2015 Nov 26 [cited 2019 Nov 26];5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4660820/.
45. Pankiv S, Clausen TH, Lamark T, Brech A, Bruun J-A, Outzen H, et al. p62/SQSTM1 binds directly to Atg8/LC3 to facilitate degradation of ubiquitinated protein aggregates by autophagy. J Biol Chem. 2007 Aug 17;282(33):24131–45.
46. Mathew R, Karp C, Beaudoin B, Vuong N, Chen G, Chen H-Y, et al. Autophagy Suppresses Tumorigenesis Through Elimination of p62. Cell. 2009 Jun 12;137(6):1062–75.

Tables

Table 1 Drug sensitivities analysis of the five key ARGs in LIHC in the GDSC database

	Drug name	Drug ID	Synonyms	Target	Target pathway	Cor	P-value
A	OSI-930	298		KIT, VEGFR, PDGFR	RTK signaling	-0.295	0.030595896
	SL 0101-1	1039		RSK, AURKB, PIM3	ERK MAPK signaling	-0.256	0.045780241
	selumetinib	1062		MEK1, MEK2	ERK MAPK signaling	-0.36	0.014806353
	EHT 1864	1069		Rac GTPases	cytoskeleton	-0.385	0.010997744
	HG-5-113-01	1142		LOK, LTK, TRCB, ABL(T315I)	ABL signaling	0.988	0.001107525
	HG-5-88-01	1143		EGFR, ADCK4	EGFR signaling	0.921	0.002897228
	CCT018159	1170		HSP90	other	-0.258	0.044870223
	Trametinib	1372	GSK1120212	MEK1, MEK2	ERK MAPK signaling	-0.278	0.036580048
	selumetinib	1498		MEK1, MEK2	ERK MAPK signaling	-0.435	0.005884497
	RDEA119	1526		MEK1, MEK2	ERK MAPK signaling	-0.322	0.022820074
1	GSK1070916	226		AURKB	mitosis	-0.584	0.026060806
	GSK429286A	230		ROCK2	cytoskeleton	-0.518	0.048848921
	AICAR	1001	N1-(b-D-Ribofuranosyl)-5-aminoimidazole-4-carboxamide	AMPK agonist	metabolism	-0.683	0.009096679
	Methotrexate	1008		Dihydrofolate reductase (DHFR)	DNA replication	-0.524	0.046247672
	PD-0325901	1060	PD-0325901	MEK1, MEK2	ERK MAPK signaling	0.579	0.022900437
	XMD11-85h	1158		BRSK2, FLT4, MARK4, PRKCD, RET, SPRK1	other	0.498	0.049803199
	ZG-10	1161		IRAK1	other	0.539	0.033980264
	XMD8-92	1164		ERK5	other	0.767	0.002676975
	QL-VIII-58	1166		mTOR, ATR	TOR signaling	0.655	0.010193816
	PFI-1	1219		BRD2, BRD3, BRD4	chromatin other	0.507	0.045882005
	Trametinib	1372	GSK1120212	MEK1, MEK2	ERK MAPK signaling	0.607	0.017151605
	selumetinib	1498		MEK1, MEK2	ERK MAPK signaling	0.655	0.010193816
	RDEA119	1526		MEK1, MEK2	ERK MAPK signaling	0.499	0.049354077
l	Mitomycin C	136	Mytozytrex,NSC-26980	DNA crosslinker	DNA replication	0.672	0.039483588
	LFM-A13	192	DDE-28	BTK	other	-0.421	0.022404029
	OSI-930	298		KIT, VEGFR, PDGFR	RTK signaling	-0.543	0.005637357
	AICAR	1001	N1-(b-D-Ribofuranosyl)-5-aminoimidazole-4-carboxamide	AMPK agonist	metabolism	-0.429	0.020600244
	Methotrexate	1008		Dihydrofolate reductase (DHFR)	DNA replication	-0.704	0.00065483
	AZD6482	1066	WO2009093972	PI3K beta (P3C2B)	PI3K signaling	-0.472	0.012916124
	XMD11-85h	1158		BRSK2, FLT4, MARK4, PRKCD, RET, SPRK1	other	0.817	0.008408638
	ZG-10	1161		IRAK1	other	0.839	0.006476792
	XMD8-92	1164		ERK5	other	0.966	0.001250116
	QL-VIII-58	1166		mTOR, ATR	TOR signaling	0.883	0.003762365
	CCT018159	1170		HSP90	other	-0.366	0.038929729
	Bleomycin (50 uM)	1378		DNA damage	DNA replication	-0.374	0.036018954
A	PF-562271	158	KIN001-205	FAK	cytoskeleton	0.533	0.016339843

	JQ12	164	JQ12	HDAC	chromain histone acetylation	0.585	0.008638969
	FH535	173	FH535	unknown	other	0.431	0.049959577
	Tipifarnib	204		Farnesyl-transferase (FNTA)	other	0.453	0.039839286
	Bosutinib	1019	SKI-606	SRC, ABL, TEC	ABL signaling	0.549	0.013496219
	CEP-701	1024	CEP-701	FLT3, JAK2, NTRK1, RET	RTK signaling	0.528	0.017330455
	JNK Inhibitor VIII	1043	JNK Inhibitor VIII	JNK	JNK and p38 signaling	-0.542	0.024718826
	PD-0332991	1054	Palbociclib Isethionate	CDK4, CDK6	cell cycle	0.528	0.017330455
	HG-5-113-01	1142		LOK, LTK, TRCB, ABL(T315I)	ABL signaling	-0.523	0.030492325
	XMD11-85h	1158		BRSK2, FLT4, MARK4, PRKCD, RET, SPRK1	other	-0.614	0.010557622
	Tamoxifen	1199		ER	other	-0.523	0.030492325
	QL-XII-61	1203		BMX, BTK	other	-0.96	5.06E-05
	YK 4-279	1239		RNA helicase A	other	0.445	0.043297466
M1	GW-2580	193	GW-2580	cFMS	RTK signaling	-0.458	0.016334934
	AICAR	1001	N1-(b-D-Ribofuranosyl)-5-aminoimidazole-4-carboxamide	AMPK agonist	metabolism	-0.855	9.00E-05
	Cisplatin	1005	cis-Diammineplatinum(II) dichloride	DNA crosslinker	DNA replication	-0.351	0.046250063
	Cytarabine	1006	Ara-Cytidine,Arabinosyl Cytosine,U-19920	DNA synthesis	DNA replication	-0.442	0.019268282
	ATRA	1009	Tretinoin	Retinoic acid and retinoid X receptor agonist	other	-0.379	0.035734704
	SL 0101-1	1039		RSK, AURKB, PIM3	ERK MAPK signaling	0.766	0.023291098
	PD-0332991	1054	Palbociclib Isethionate	CDK4, CDK6	cell cycle	-0.657	0.001574464
	selumetinib	1062		MEK1, MEK2	ERK MAPK signaling	0.698	0.044517742
	XMD11-85h	1158		BRSK2, FLT4, MARK4, PRKCD, RET, SPRK1	other	0.716	0.037722355
	QL-XII-61	1203		BMX, BTK	other	0.755	0.025970206
	CHIR-99021	1241	CT 99021	GSK3B	WNT signaling	-0.535	0.007036208
	Talazoparib	1259	BMN-673	PARP1, PARP2	Genome integrity	-0.443	0.019072299
	rTRAIL	1261		DR4, DR5	apoptosis regulation	-0.657	0.001574464
	Olaparib	1495	KU0059436, AZD2281	PARP1, PARP2	Genome integrity	-0.391	0.031895533

Figures

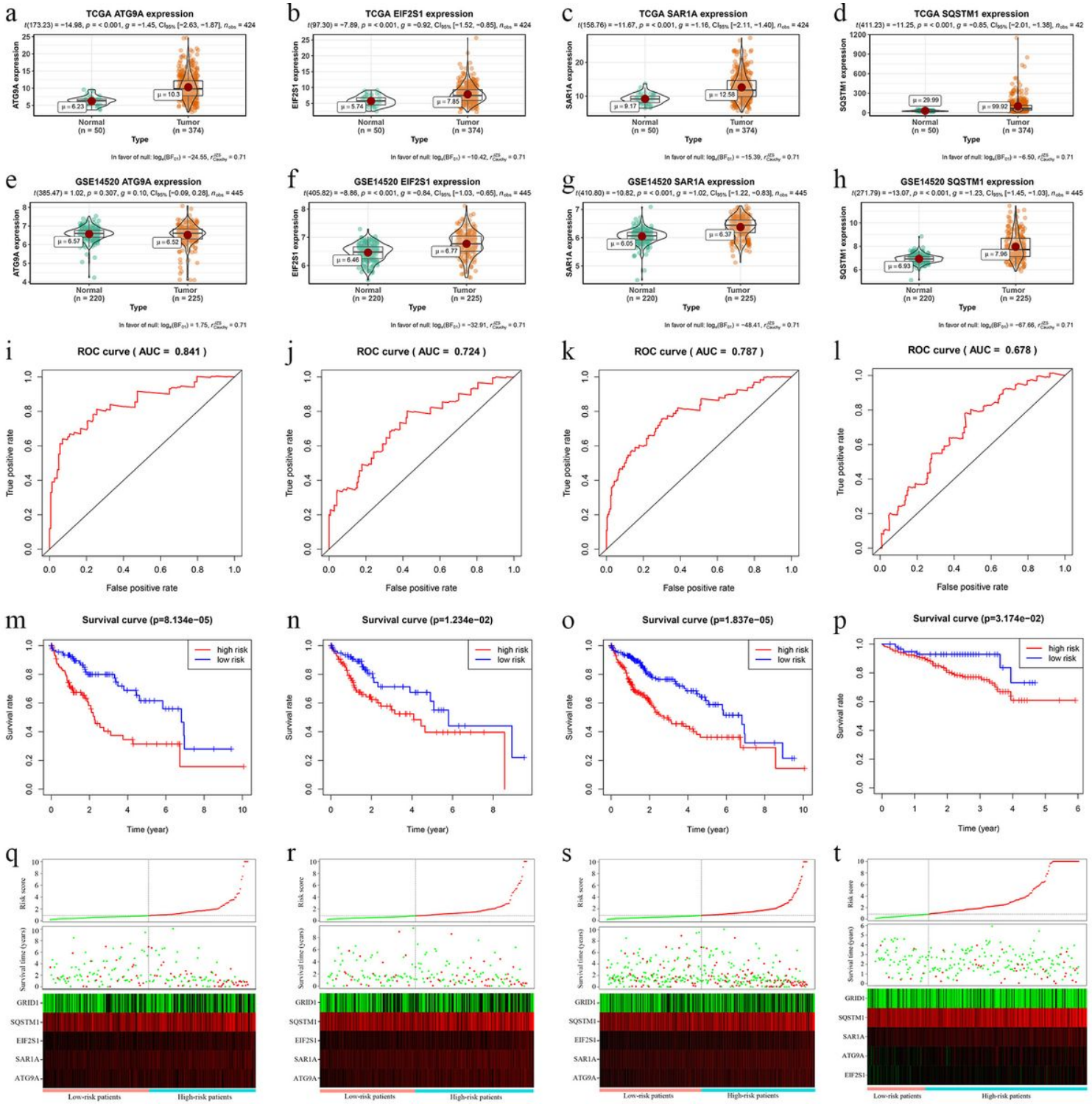


Figure 1

The verifications of the PI model. a-d: the expression patterns of genes ATG9A (a), EIF2S1 (b), SAR1A(c) and SQSTM1(d) in LIHC and non-cancer liver tissues in the whole TCGA dataset. e-h: the expression patterns of genes ATG9A (e), EIF2S1 (f), SAR1A(g) and SQSTM1(h) in LIHC and non-cancer liver tissues in GSE14520 dataset. (i-l): the ROC curves in the training group (i), the testing group (j), the whole TCGA dataset (k) and the ICGC database (l). (m-p): the Kaplan-Meier survival curves of high-risk and low-risk groups in the training group (m), the testing group (n), the whole TCGA dataset (o) and the ICGC database

(p). In the risk score analysis of LIHC patients in the training group (q), the testing group (r), the whole TCGA dataset (s) and the ICGC database (t), the distribution of risk score of LIHC patients was ranged in the upper panel; the survival state of LIHC patients was ranged in the middle panel, with red pot as dead and green pot as alive; and the expression patterns of the five key ARGs in our PI model was ranged in the bottom panel. p < 0.05 was viewed as statistically significant. ICGC: International Cancer Genome Consortium; PI: prognostic index; ROC: receiver operating characteristic; AUC: the area under the ROC curve.

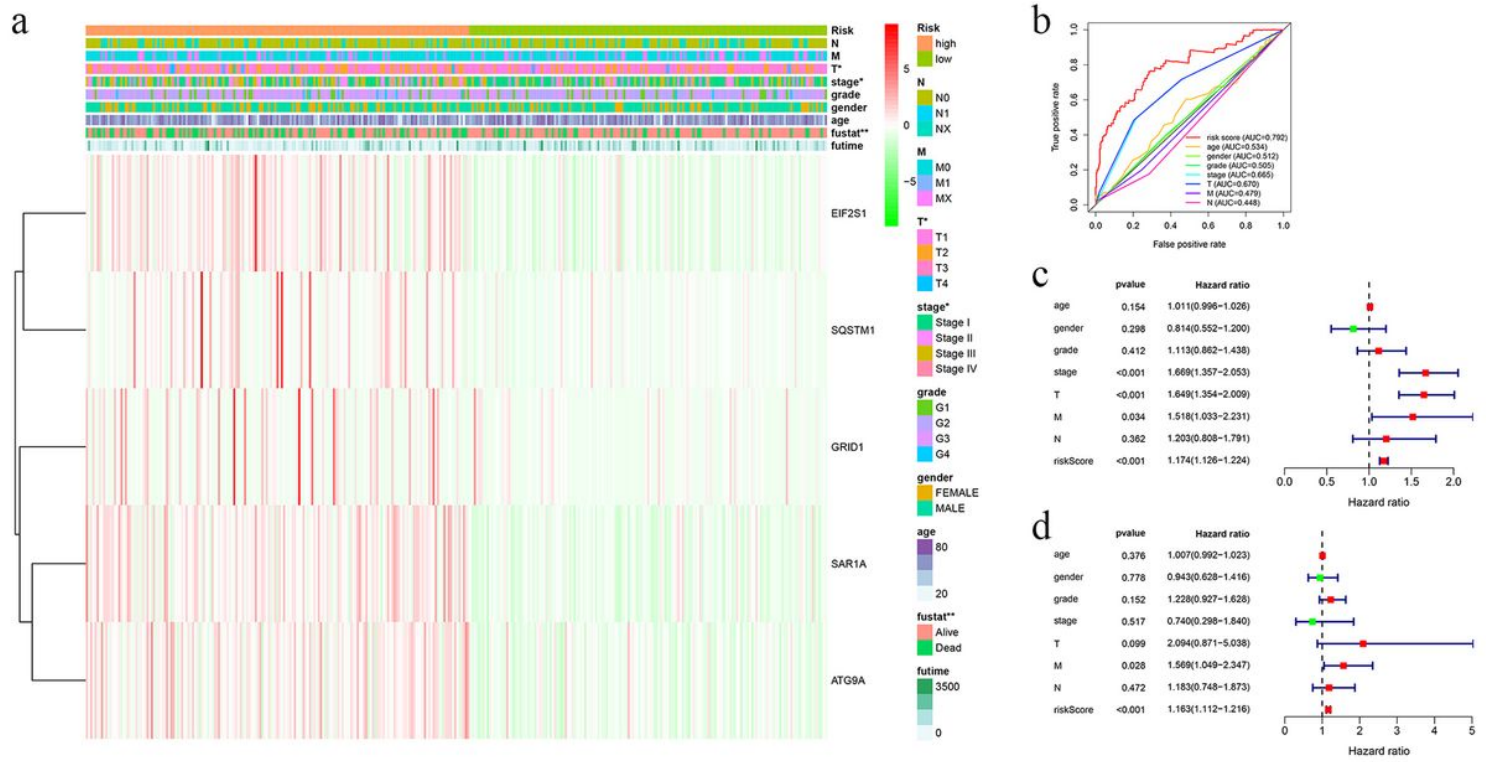


Figure 2

Comparison between the PI model and common clinicopathologic features in LIHC patients. a: the heatmap of the distribution of the risk score and some common clinicopathologic features in the whole TCGA database, including age, gender, grade, stage, T, N, M, survival state and survival time. b: the comparisons of the ROC curves of our PI model and some common clinicopathologic features. c: the univariate independent prognostic analysis. d: the multivariate independent prognostic analysis. p < 0.05 was viewed as statistically significant; *: p <0.05; **: p <0.01. T: primary tumor staging; N: regional lymph node staging; M: distant metastasis staging; ROC: receiver operating characteristic; AUC: the area under the ROC curve.

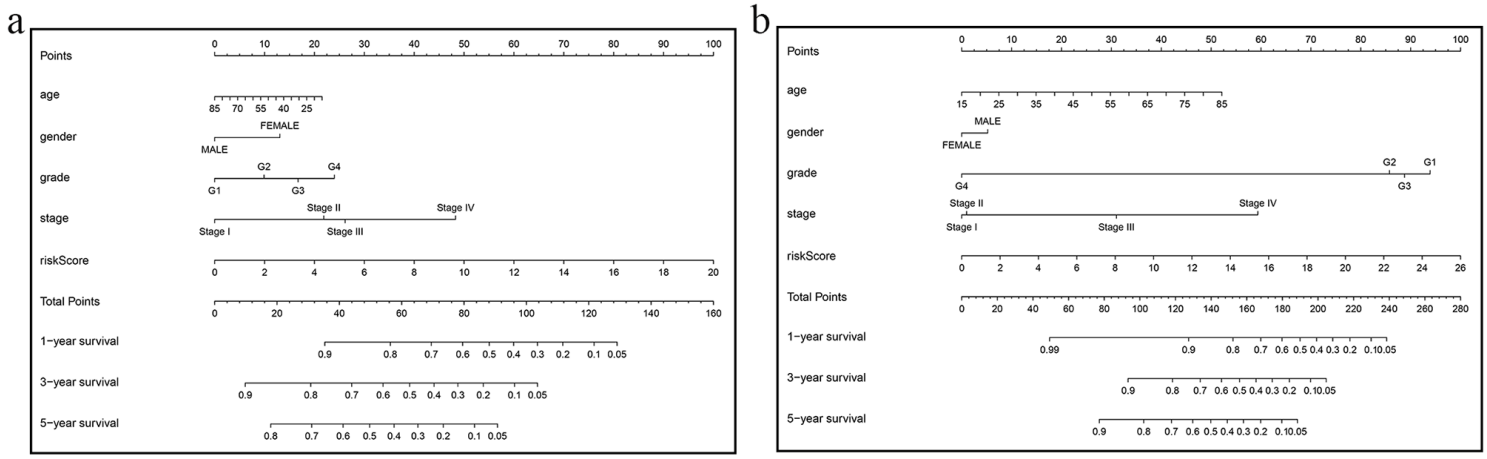


Figure 3

Nomograms to predict the 1,3,5-year survival rate of individual LIHC patient. a: nomogram based on the training group; b: nomogram based on the testing group.

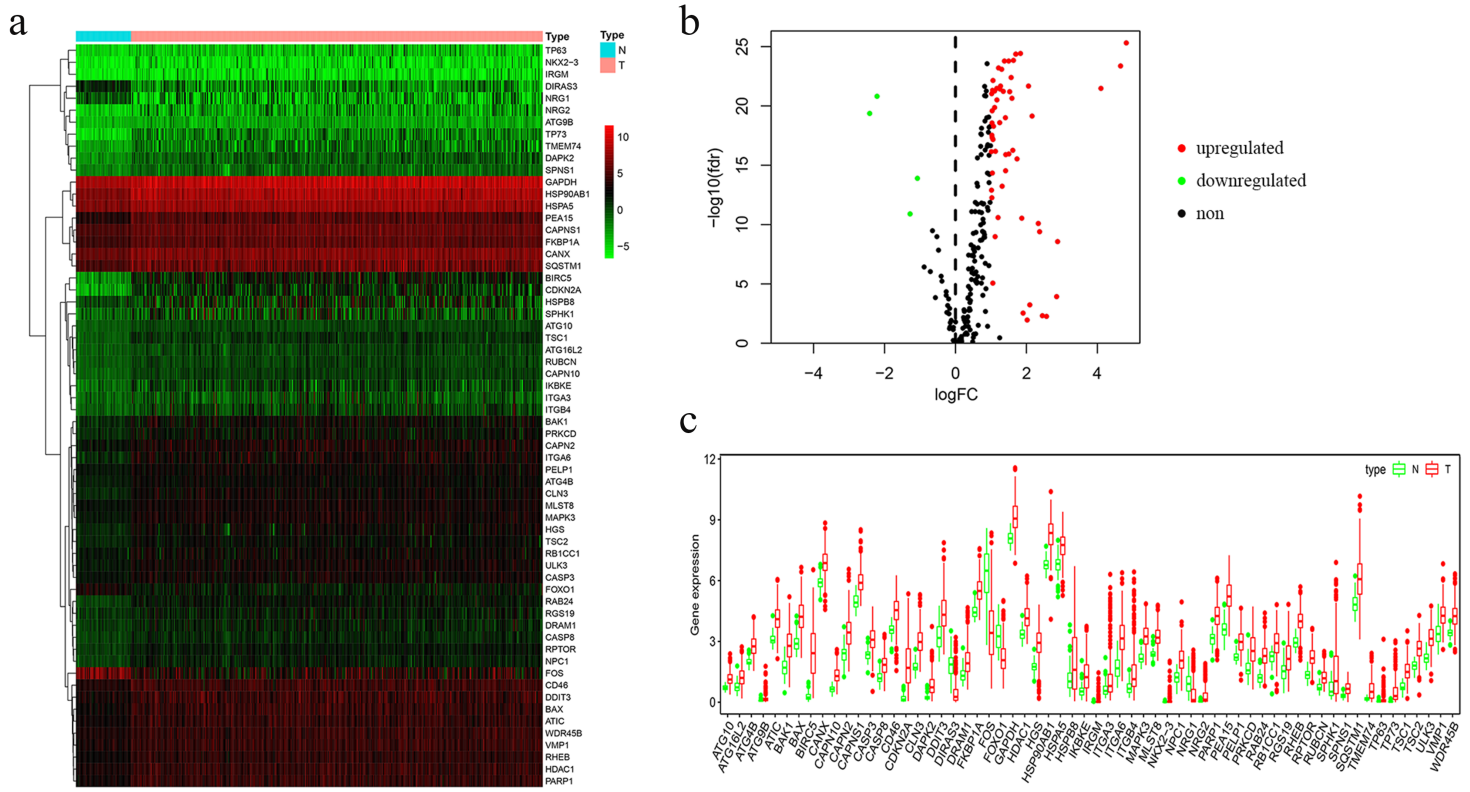
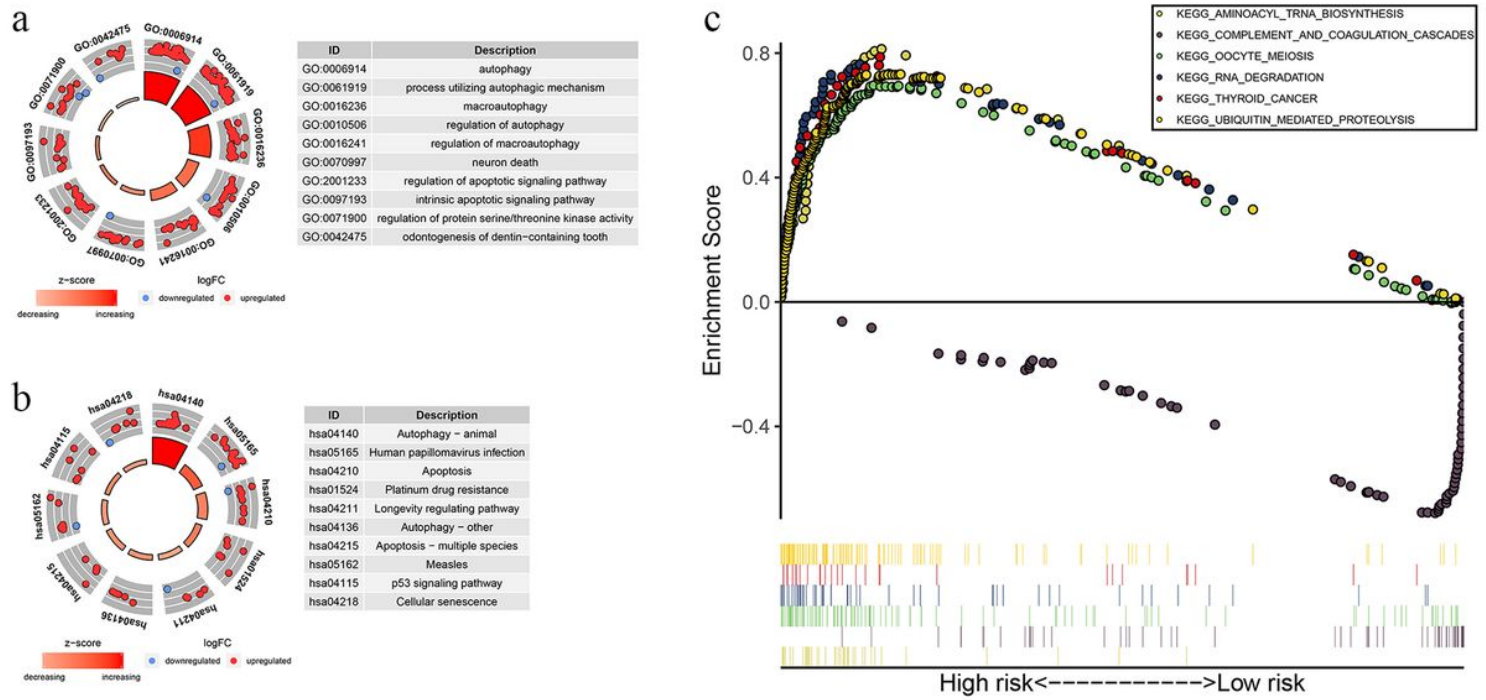


Figure 4

The expression patterns of 62 differentially expressed ARGs in LIHC and normal tissues. a: the heat map of 62 differentially expressed ARGs in LIHC and normal tissues. The blue type was on behalf of the non-tumor sample, and the pink type was on behalf of LIHC sample. The expression pattern of ARGs was shown from red to green: red denoted up-regulated and green denoted down-regulated. b: the volcano figure of 62 differentially expressed ARGs in LIHC and normal tissues. The horizontal axis is logFC, and the vertical axis is -log10(FDR). The red dots represent upregulated genes, the green dots downregulated

genes; and the black dots represent genes without statistical expressed difference between LIHC and normal tissues. c: the boxplot of expression level of 62 differentially expressed ARGs in LIHC and normal tissues. The x-axes showed the 62 differentially expressed ARGs and the y-axes showed their expression levels. The green type was on behalf of the non-tumor sample, and the red type was on behalf of LIHC sample. FC: fold change; FDR: false discovery rate.



ID	Description
GO:0006914	autophagy
GO:0061919	process utilizing autophagic mechanism
GO:0016236	macroautophagy
GO:0010506	regulation of autophagy
GO:0016241	regulation of macroautophagy
GO:0070997	neuron death
GO:2001233	regulation of apoptotic signaling pathway
GO:0097193	intrinsic apoptotic signaling pathway
GO:0071900	regulation of protein serine/threonine kinase activity
GO:0042475	odontogenesis of dentin-containing tooth

ID	Description
hsa04140	Autophagy - animal
hsa05165	Human papillomavirus infection
hsa04210	Apoptosis
hsa01524	Platinum drug resistance
hsa04211	Longevity regulating pathway
hsa04136	Autophagy - other
hsa04215	Apoptosis - multiple species
hsa05162	Measles
hsa04115	p53 signaling pathway
hsa04218	Cellular senescence

Figure 5

The functional enrichment analyses. a: the circle graph of the top 10 GO terms. b: the circle graph of the top 10 KEGG terms. In the circle graphs, the inner ring showed the z-score of the corresponding pathway which reflected their increase or decrease amplitude, and the outer ring showed the logFC scatter plots of assigned genes in each term. Red dots represented upregulation and blue dots downregulation. c: GSEA analysis of KEGG pathways in the training group. The KEGG pathway with ES > 0 is activated in high-risk group, while the KEGG pathway with ES < 0 is activated in low risk group. Each dot represents one gene involved. GO: gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; FC: fold change; GSEA: Gene Set Enrichment Analysis.

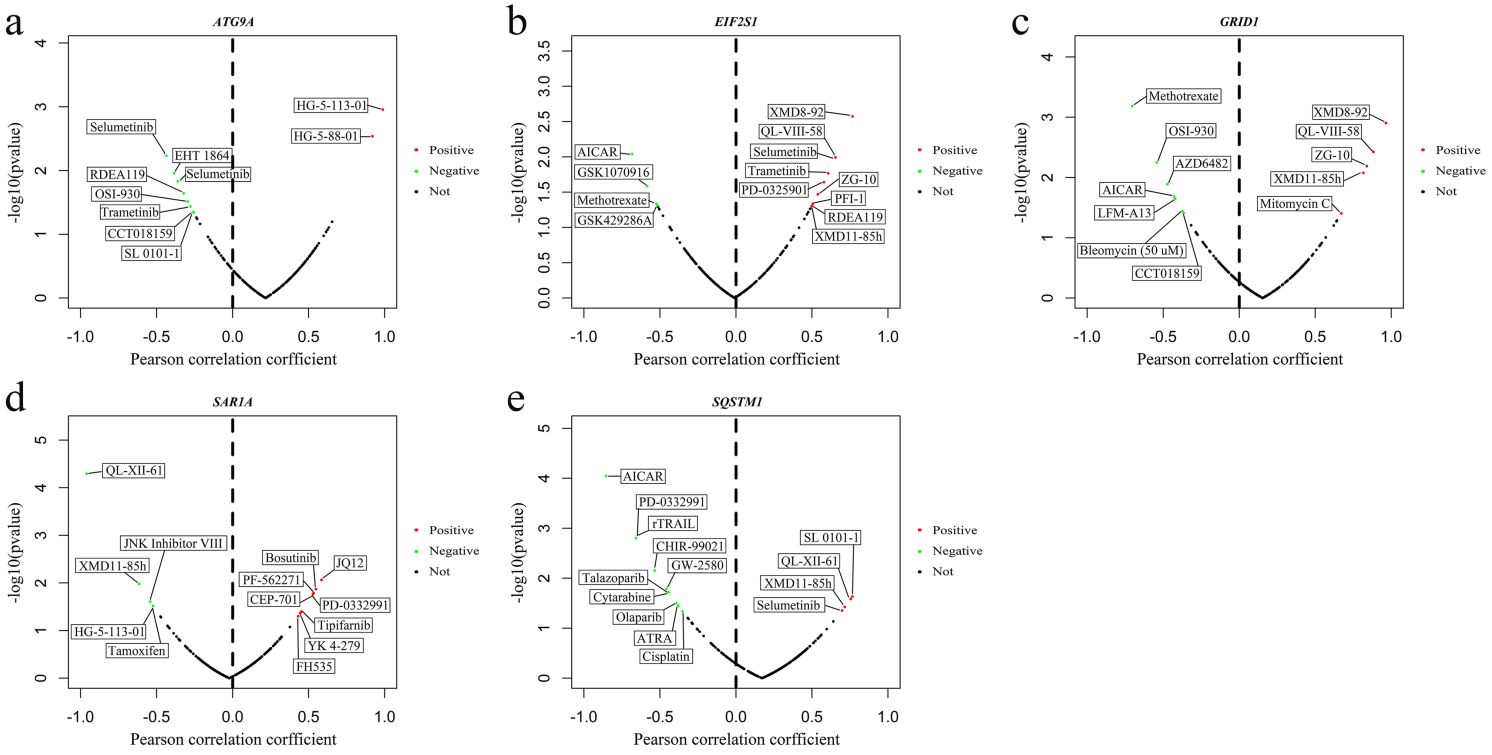


Figure 6

Drug sensitivity analyses in the GDSC database. (a-e) respectively displayed the correlation between the expression of genes ATG9A, EIF2S1, GRID1, SAR1A and SQSTM1 and the drug sensitivities in LIHC cell lines. The drug of which the half maximal inhibitory concentration (IC50) value was negatively correlated to the expression of the according gene ($p < 0.05$) was represented by the green dot; and the drug of which the IC50 value was positively correlated to the expression of the according gene ($p < 0.05$) was represented by the red dot; and the drug of which the IC50 value was not statistically correlated to the expression of the according gene ($p > 0.05$) was represented by the black dot. GDSC: Drug Sensitivity in Cancer.

Supplementary Files

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

- Additionalfile6.xlsx
- Additionalfile3.xlsx
- Additionalfile9.xlsx
- Additionalfile10.xlsx
- Additionalfile6.xlsx
- Additionalfile9.xlsx
- Additionalfile10.xlsx
- Additionalfile3.xlsx

- Additionalfile1.xlsx
- Additionalfile7.pdf
- Additionalfile7.pdf
- Additionalfile5.pdf
- Additionalfile2.pdf
- Additionalfile4.pdf
- Additionalfile5.pdf
- Additionalfile2.pdf
- Additionalfile4.pdf
- Additionalfile8.pdf
- Additionalfile8.pdf
- Additionalfile1.xlsx