


Absent in melanoma 1-like (AIM1L) serves as a novel candidate for overall survival in hepatocellular carcinoma

Wenliang Zhou^a, Yuan Zhang^b, Shixi Zhang^{a*}, and Zongguo Yang ^{b*}

^aDepartment of Infectious Diseases, Shangqiu Municipal Hospital, Shangqiu, He'nan, China; ^bDepartment of Integrative Medicine, Shanghai Public Health Clinical Center, Fudan University, Shanghai, China

ABSTRACT

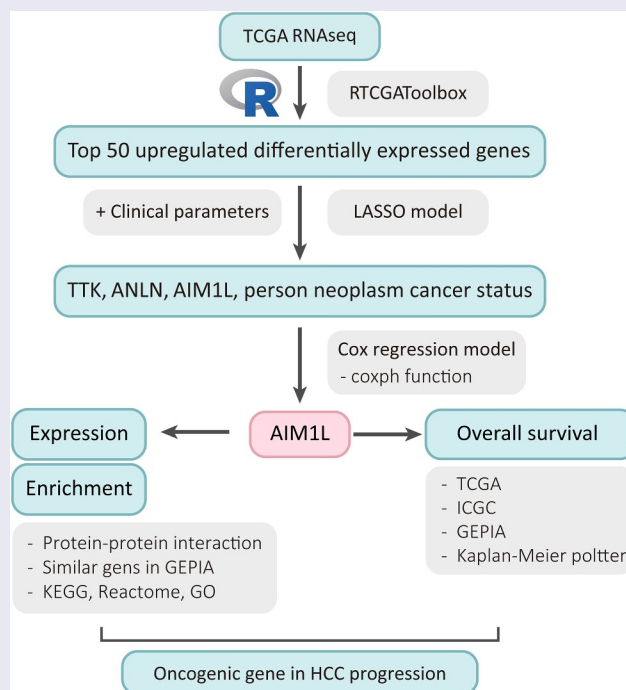
Identifying biomarkers for hepatocellular carcinoma (HCC) survival is of great importance for the early detection, monitoring, and predicting for prognosis. This study aimed to investigate the candidate biomarkers for predicting overall survival (OS) in HCC patients. Using RTCGAToolbox, top 50 upregulated differentially expressed genes (DEGs) were identified. The least absolute shrinkage and selection operator (LASSO) and Cox models were used to select powerful candidate genes, and log rank method was used to address the survivor functions of potential biomarkers. Selected by LASSO model, ANLN, TTK, AIM1L and person neoplasm cancer status might be candidate parameters associated with OS in HCC patients. After adjusting person neoplasm cancer status, ANLN and TTK levels in Cox model, AIM1L was identified as a risk factor for predicting OS in HCC patients (HR = 1.5, $P = 0.037$). Validated in The Cancer Genome Atlas (TCGA), International Cancer Genome Consortium (ICGC) and Gene Expression Omnibus (GEO) series, AIM1L was significantly overexpressed in tumor tissues compared to nontumor tissues (all $P < 0.0001$). HCC patients with high AIM1L in tumor tissues had significantly unfavorable OS compared to those with low AIM1L in TCGA, ICGC, Gene Expression Profiling Interactive Analysis (GEPIA) and Kaplan-Meier Plotter datasets (all $P < 0.05$). Conclusively, AIM1L is upregulated in tumor samples and serves as a novel candidate for predicting unfavorable OS in HCC patients.

ARTICLE HISTORY


Received 30 March 2021
Revised 1 June 2021
Accepted 2 June 2021

KEYWORDS

AIM1L; hepatocellular carcinoma; overall survival; absent in melanoma 1-like



*CONTACT Zongguo Yang  yangzongguo@shphc.org.cn  Department of Integrative Medicine, Shanghai Public Health Clinical Center, Fudan University, 2901 Caolang Road, Shanghai, 201508, China; Shixi Zhang  shixizhang1977@163.com  Shangqiu Municipal Hospital, No. 1 Yingbin Road, Shangqiu 476100, He'nan Province China

 Supplemental data for this article can be accessed [here](#).

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

Hepatocellular carcinoma (HCC) is one of the most frequently occurring human malignancies and the second leading cause of cancer-related deaths [1]. During the past few decades, the incidence of HCC continued to increase, and it will be estimated to rise over the next 10 to 20 years [2–4]. Even recent advances in the first-line systemic treatment of showed effective results for HCC patients, the overall prognosis of this population is still dull [5–7]. The liver cancers including HCC related mortality has marked increased by more than 2% annually since 2007 [2]. To identify potential biomarkers for HCC survival should be of great importance for the early detection, monitoring, and predicting for prognosis, which is also helpful for the understanding the pathological characteristics in HCC patients.

Several genetic candidates have been used in HCC surveillance, diagnosis, prognosis, and treatment responses [8–10]. With the development of high-throughput technologies and gene chips, and next-generation sequencing, the genomic pattern of HCC has been determined which greatly improved our understanding of genetic and epigenetic changes and their interaction in the HCC aggressiveness [8,11–14]. All these approaches have become fast approaches to identify differentially expressed genes (DEGs) and functional pathways, and led to a dramatic increase in the accessibility of molecular insights at multiple biological levels involved in the HCC development [13,15–18]. Recently, these bioinformatics data repositories have rapidly evolved into an essential aid for molecular hepatology [15]. The availability of genome sequencing data from liver tumors provides us with valuable resources, which is vital to help us to facilitate the identification of promising biomarkers or therapeutic targets for HCC population [14].

In our study, the top 50 upregulated DEGs between tumor and nontumor samples in HCC patients were identified using RTCGAToolbox package in The Cancer Genome Atlas (TCGA) dataset [19]. The least absolute shrinkage and selection operator (LASSO) model and Cox proportional hazards regression model were used to

investigate and validate the potential candidates [20]. The aim of this study is to screen promising candidates for predicting overall survival (OS) in HCC patients with integrative bioinformatic approaches.

2 Materials and methods

2.1 Data source

Using the ‘getFirehoseData’ function in the RTCGAToolbox package [19], gene expression data calculated by RNAseq from HCC patients were downloaded from the Firehose project when dataset was set as ‘LIHC’, runDate was set as ‘20,160,128’, and RNAseqGene means ‘TRUE’. To assess the DEGs between the normal and tumor samples in HCC patients in the TCGA dataset, the ‘getDiffExpressedGenes’ function was addressed with criteria P value < 0.05 , adjusted P value < 0.05 and $\log_{2}FC \geq 2$. Top 50 upregulated DEGs was obtained for heatmap performance when ‘hmTopUpN’ equals to 50 and ‘hmTopDownN’ equals to 0.

2.2 Patients

The clinical and gene expression data with Z scores of Liver Hepatocellular Carcinoma (LIHC, TCGA, PanCancer Atlas) dataset was obtained from cBioPortal for cancer genomics [21,22]. After restricting tumor pathological type as HCC and matching gene expression levels with clinical data, 366 HCC patients were included in this study. Clinico-pathological characteristics including age, gender, American Joint Committee on Cancer (AJCC) staging, new tumor even after initial treatment, pathological TNM stages, person neoplasm cancer status, race, radiation therapy, and weight were available. Every participant provided verified informed consent, as declared in the original dataset.

In the International Cancer Genome Consortium (ICGC, <https://daco.icgc.org/>) database, liver cancer project LIRI-JP with HCC subtype was included in this study [23]. 260 HCC donors were available. After matching the gene expression data and survival information, and

removing subjects with OS less than 20 days, 235 cases were included in the final analysis.

2.3 Gene expression analysis

Raw.CEL files of the microarray datasets in Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) were downloaded and normalized by quantile method of Robust Multichip Analysis from the R affy package [24]. k-Nearest Neighbor method by impute function was used to impute the missing gene expression data [25]. Gene expression levels between the normal and tumor samples were calculated by the Limma package [26].

In the TCGA dataset, mRNA normalized counts data of LIHC derived from RNAseq Htseq platform were downloaded from Genomic Data Commons Data Portal (<https://cancergenome.nih.gov/>). The edgeR package in R program were used to identify gene expression levels between tumor and nontumor tissues [27,28].

In the ICGC dataset, mRNA normalized counts data of 237 tumor samples and 197 normal liver samples in LIRI-JP project, which was derived from Illumina HiSeq platform, were downloaded from ICGC data portal (<https://dcc.icgc.org/>) [23].

2.4 LASSO model establishment

LASSO regression model was used to determine the most powerful prognostic markers for OS in HCC patients [20]. In the TCGA dataset, parameters including age, gender, tumor status, AJCC staging, weight, radiation therapy, race, pathological status, and top 50 upregulated DEGs were included in the LASSO model. ‘glmnet’ and ‘survival’ packages were used for LASSO model establishment with family equals to ‘cox’ and alpha equals to 1. The model was validated with 5-fold cross-validation. Both ‘lambda.1se’ and ‘lambda.min’ were used to assess the coefficient of parameters [29].

2.5 Functional enrichment of candidate genes

Protein-protein interaction (PPI) analysis of the candidate genes were addressed in STRING [30] and STITCH databases [31]. Top 100 similar genes of the

candidate genes were respectively searched from LIHC tumor, LIHC normal and Genotype-Tissue Expression (GTEx) liver datasets in GEPIA database [32]. All these AIM1L-related genes were screened by edgeR package with $|\log FC| > 1$, and adjusted P value < 0.05 [27,28], differentially expressed genes were enrolled in the Gene Set Enrichment Analysis (GSEA) [33]. Investigate gene sets module in the Molecular Signatures Database (MSigDB) v7.4 in GSEA were used for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, Reactome, and Gene Ontology (GO) enrichment [34,35]. Top ten genesets were included with a false discovery rate (FDR) q -value less than 0.05.

2.6 Statistical analysis

Differences of gene levels between the individual groups were analyzed using student t test or Mann-Whitney U test based on data types. Parameters enrolled in LASSO model were included in Cox proportional hazards regression model. Results were reported as hazard ratios (HR) with 95% confidence interval (95% CI). Log rank method was used to address the survivor functions of candidate genes for OS in HCC patients. The Kaplan Meier plotter [36] and Gene Expression Profiling Interactive Analysis (GEPIA) [32] online services were also used to confirm the survival relationship. Stata software version 16.0 (STATA Corp., Texas, USA) was used. A two-tailed $P < 0.05$ were considered significance for all tests.

3 Results

The LASSO model showed that TTK, ANLN, AIM1L and person neoplasm cancer status should be underlying candidates of OS in HCC. Using integrated bioinformatic methods, we found that AIM1L was at a low level in normal livers and significantly overexpressed in HCC tumor tissues. Cox regression model and Kaplan-Meier analysis in multiple datasets indicated that high levels of AIM1L in tumors contributed to unfavorable OS in HCC patients. Functional enrichment revealed that AIM1L-related genes might involve in cell proliferation and cell migration. We assumed that AIM1L exerts oncogenic roles in HCC progression.

3.1 Top 50 upregulated candidate genes identified by RTCGAToolbox

Detected by the RTCGAToolbox package in R program, top 50 upregulated DEGs in tumor tissues were identified. The heatmap of these DEGs between tumor and nontumor samples were presented in Figure 1.

3.2 Selection of powerful biomarkers for OS in HCC patients by LASSO and Cox

In the TCGA dataset, the top 50 upregulated DEGs, together with clinico-pathological

characteristics including age, gender, AJCC staging, new tumor even after initial treatment, pathological TNM stages, person neoplasm cancer status, race, radiation therapy, and weight were enrolled in LASSO model (Figure 2a). After the 5-fold cross validation, parameters including TTK, ANLN, AIM1L and person neoplasm cancer status were recruited to be underlying candidates of OS in HCC patients when λ took the minimum value (Figure 2a). The regression coefficient plot of factors by LASSO was shown in Figure 2b.

To fit a Cox proportional hazards regression model, the 'coxph' function in R program was used. As summarized in Figure 3, AIM1L serves

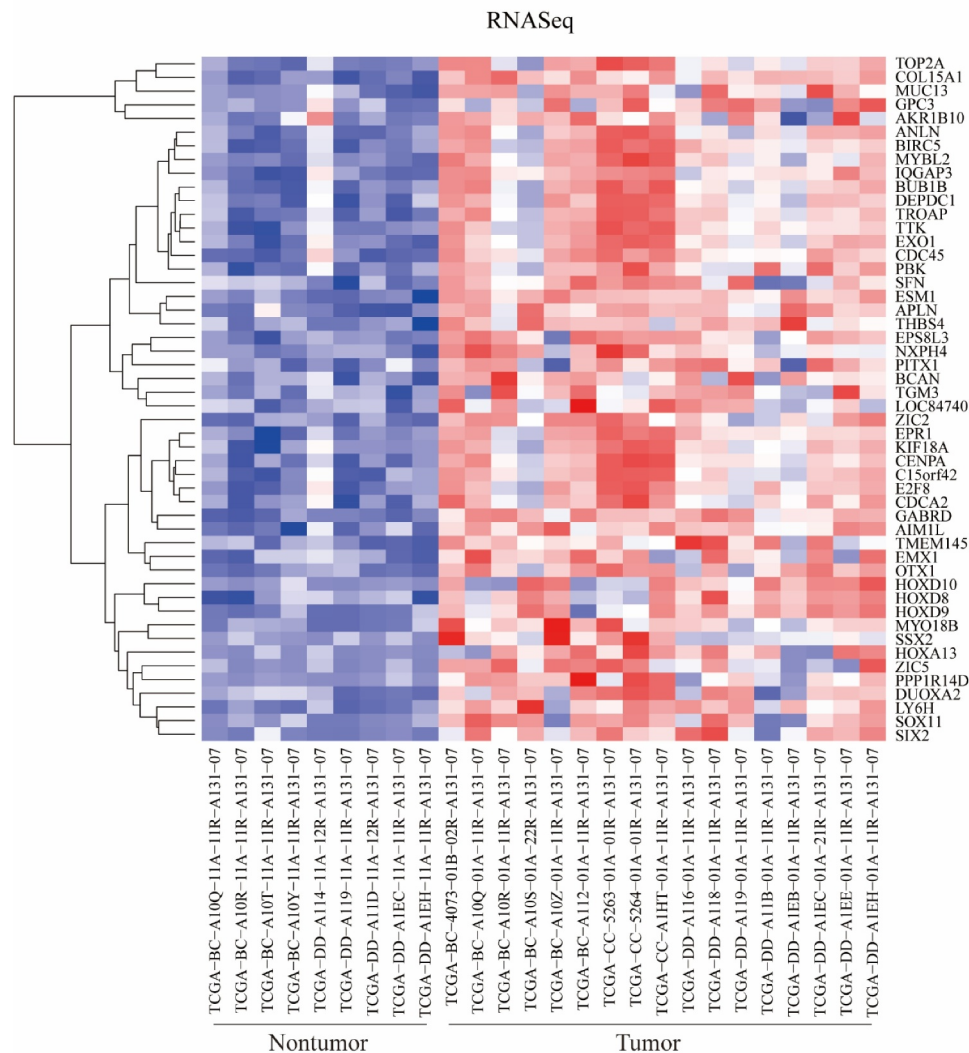


Figure 1. Differentially expressed genes (DEGs) between tumor and nontumor tissues in HCC patients in TCGA dataset were screened the 'getDiffExpressedGenes' function in RTCGAToolbox package with criteria P value < 0.05 , adjusted P value < 0.05 and $\log_{2}FC \geq 2$. Top 50 upregulated DEGs was obtained for heatmap performance when 'hmTopUpN' equals to 50 and 'hmTopDownN' equals to 0.

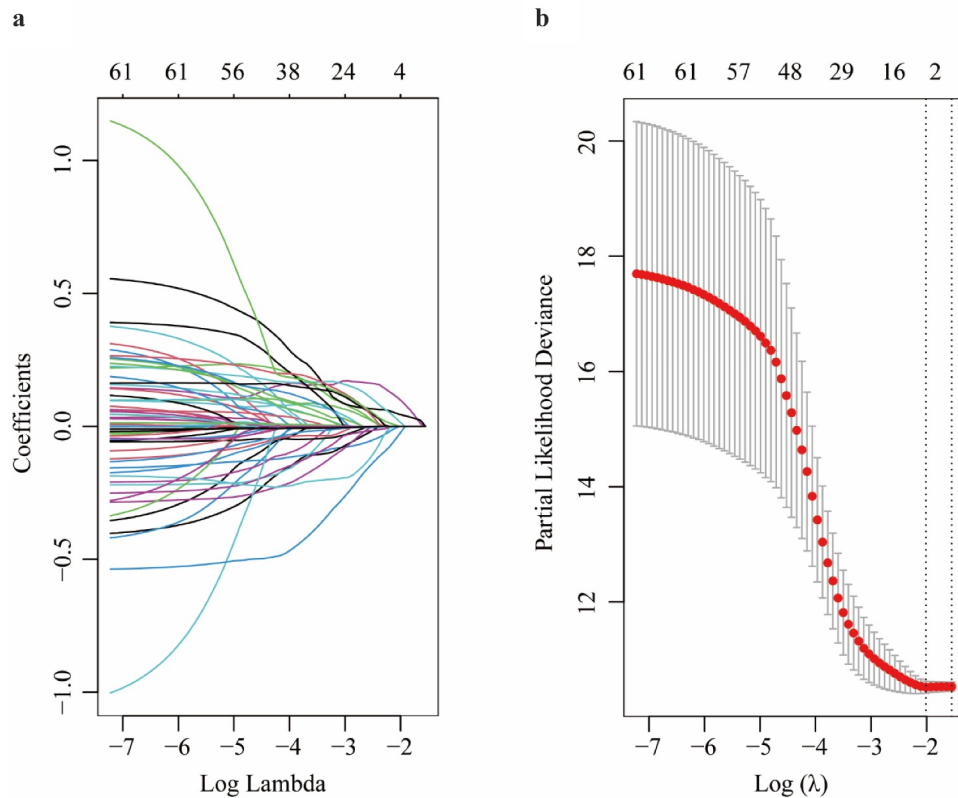


Figure 2. Parameter selection through LASSO regression (a) and elucidation of LASSO coefficient profiles for selected factors (b).

as an independent prognosis predictor for OS in HCC patients after adjusting TTK, ANLN and person neoplasm cancer status (HR = 1.5, 95% CI = 1.02–2.1, $P = 0.037$, Figure 3).

3.3 AIM1L expression

In Fagerberg's report, RNAseq was performed to determine tissue-specificity of protein-coding genes of tissue samples from 95 human individuals representing 27 different tissues [37]. In Consensus dataset, normalized expression (NX) of the candidate from three transcriptomics datasets, namely, Human Protein Atlas (HPA), the Genotype-Tissue Expression (GTEx) project and the Functional Annotation of Mammalian Genomes 5 (FANTOM5) project, was calculated. As shown in Figure 4, AIM1L was expressed at low levels in normal liver tissues in Fagerberg's dataset and Consensus dataset (Figure 4a and b). In addition, single cell RNA sequencing indicated that AIM1L was not detected in Hepatocytes, Cholangiocytes, Endothelial cells, Erythroid cells,

Ito cells, Kupffer cells, T cells and B cells in liver tissue (Figure 4c).

Next, we compared the AIM1L expression between tumor and nontumor samples. In the TCGA dataset, AIM1L mRNA was significantly upregulated in tumor tissues compared to nontumor tissues ($P < 0.0001$, Figure 5a). In the 50 paired normal and tumor samples extracted from TCGA dataset, AIM1L was also significantly overexpressed in tumor samples compared to normal livers ($P < 0.0001$, Figure 5b). Additionally, AIM1L mRNA was significantly upregulated in tumor tissues compared to nontumor tissues in LIRI-JP project from ICGC ($P < 0.0001$, Figure 5c) and GEO series including GSE45436 [38], GSE55092 [39], GSE84402 [40], GSE101685 [41], GSE14323 [42], GSE112790 [43], and GSE121248 [44] (all $P < 0.0001$, Figure 5d). In a rat model naturally occurring hepatocarcinogenesis induced by oxidative stress [45], AIM1L was significantly upregulated in liver cancer tissues compared to normal liver tissues ($P < 0.01$, Figure 5e).

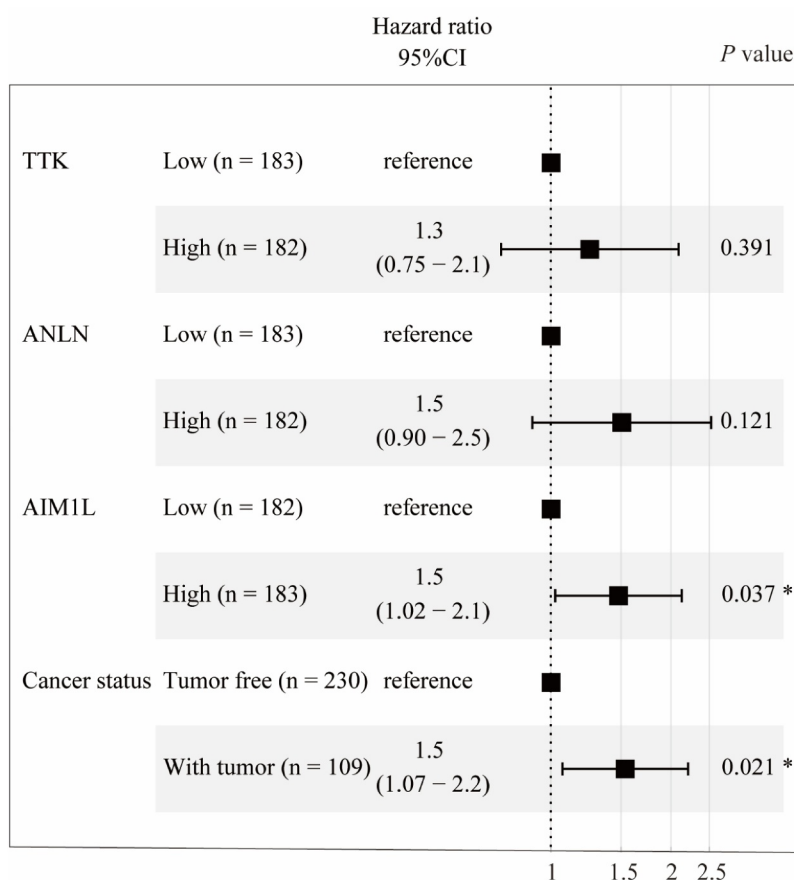


Figure 3. Cox proportional hazards regression model using the ‘coxph’ function in R program was established. Four parameters including TTK, ANLN, AIM1L, and cancer status screened by LASSO model were included in Cox regression model for OS in HCC patients. After adjusting TTK and ANLN, AIM1L and cancer status were significantly associated with OS in HCC patients (both $P < 0.05$).

3.4 High AIM1L accounts for unfavorable OS in HCC patients

As described in Figure 6, HCC patients from the TCGA dataset with high AIM1L in tumor tissues had significantly unfavorable OS compared to those with low AIM1L ($P = 0.041$, Figure 6a). In the LIRI-JP project from the ICGC database, HCC patients with high AIM1L in tumor tissues had significantly poorer OS than those with low AIM1L ($P = 0.046$, Figure 6b).

In the GEPIA database, high AIM1L might contribute to significantly worse OS in HCC patients (HR = 1.5, $P = 0.02$, Figure 6c). Similarly, HCC patients with high AIM1L levels in tumor tissues had higher risk for unfavorable OS compared to those with low AIM1L in the Kaplan-Meier Plotter database (HR = 1.64, 95% CI = 1.16–2.33, $P = 0.005$, Figure 6d). Considered results from

Cox model and Kaplan-Meier analysis, we assumed that AIM1L serves as a novel biomarker for predicting unfavorable OS in HCC patients.

3.5 Functional enrichment of AIM1L-related genes

The PPI of AIM1L was presented in Supplementary Figure S1. The similar genes of AIM1L in LIHC tumor, LIHC normal and GTEx datasets in GEPIA platform were summarized in Supplementary Table S1. Whether these AIM1L-related genes were differentially expressed between tumor and nontumor tissues were screened using the edgeR package in R program. Totally, 144 AIM1L-related DEGs were selected in the functional enrichment analysis (Supplementary Figure S2). As described in Figure 7, AIM1L-related genes

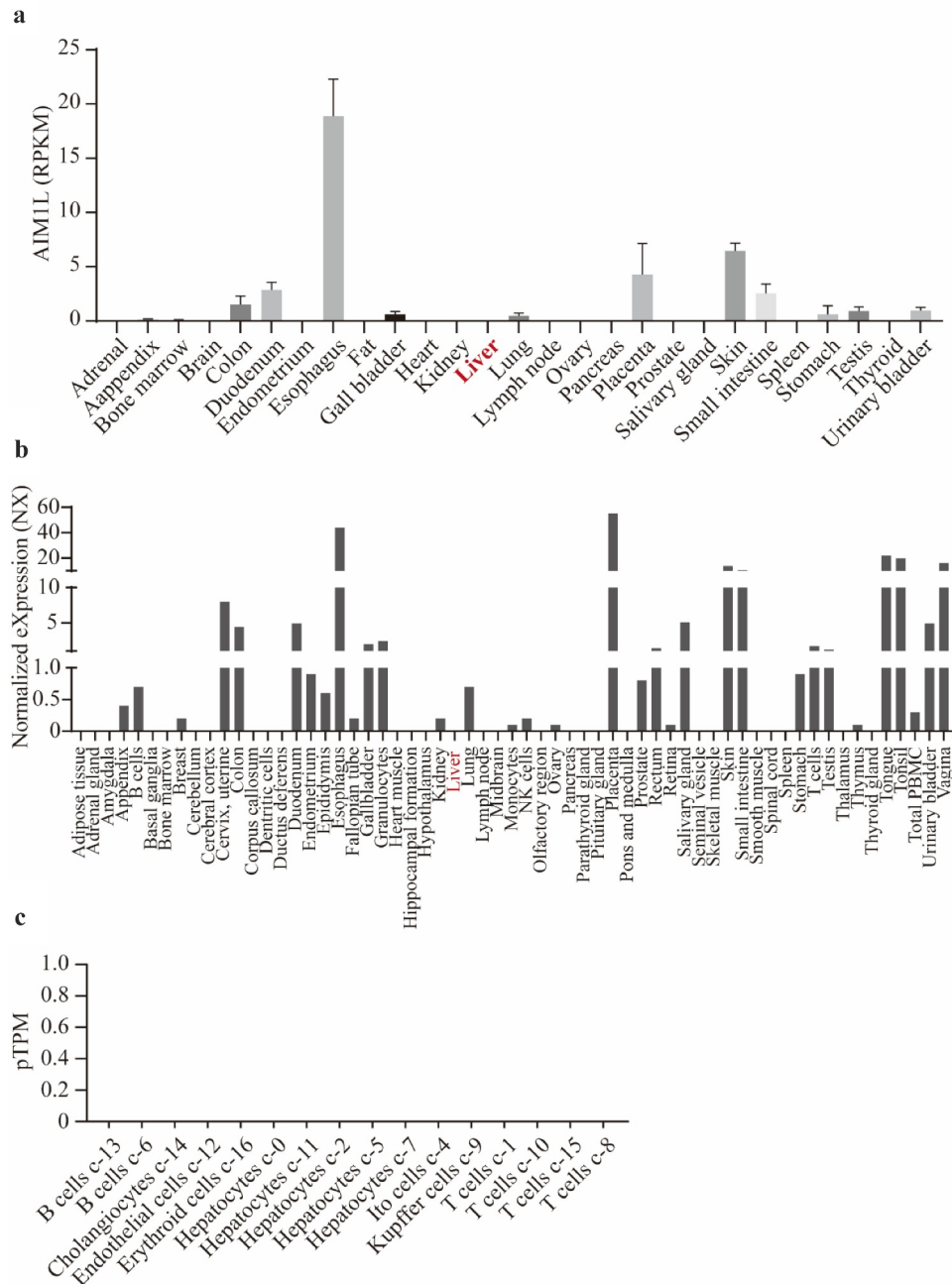


Figure 4. Through searching AIM1L in PubMed (gene ID: 55057), the gene expression levels in 27 normal tissues from Fagerberg's dataset were obtained. In Consensus dataset in GEPIA database, normalized expression (NX) of AIM1L from three transcriptomics datasets, namely, Human Protein Atlas (HPA), the Genotype-Tissue Expression (GTEx) project and the Functional Annotation of Mammalian Genomes 5 (FANTOM5) project, was calculated. AIM1L is at low levels in normal liver both in Fagerberg's dataset (a) and Consensus dataset from HPA (b). In HPA database, single cell RNA sequencing indicated that AIM1L was not detected in Hepatocytes, Cholangiocytes, Endothelial cells, Erythroid cells, Ito cells, Kupffer cells, T cells and B cells in liver tissue (c).

involved in the process of extracellular matrix (ECM) organization, ECM-receptor interaction, calcium signaling pathway, focal adhesion, regulation of action cytoskeleton, and neuroactive

ligand-receptor interaction (all FDR < 0.05, Figure 7a). GO enrichment revealed that AIM1L-related genes implicated in multiple biological processes (BP) including cell proliferation and cell

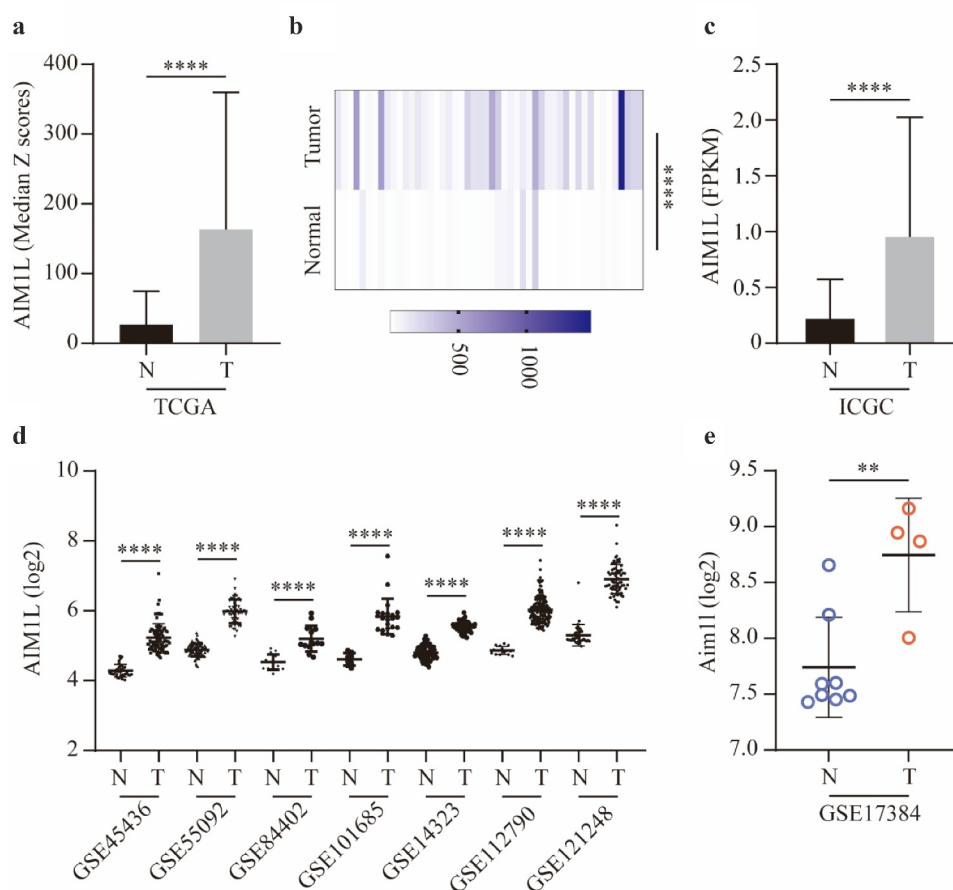


Figure 5. AIM1L expression between tumor and nontumor samples in HCC patients were compared. AIM1L was significantly upregulated tumor tissues in TCGA ($P < 0.0001$, A), in 50 paired tumor samples from TCGA ($P < 0.0001$, B), in LIRI-JP project from ICGC ($P < 0.0001$, C) and in GEO series including GSE45436, GSE55092, GSE84402, GSE101685, GSE14323, GSE112790, and GSE121248 (all $P < 0.0001$, D). In a rat model naturally occurring hepatotumorigenesis induced by oxidative stress, AIM1L was significantly upregulated in liver cancer tissues compared to normal liver tissues ($P < 0.01$, E).

migration (Figure 7b). The GO cellular component (CC), and molecular function (MF) of AIM1L-related genes were also summarized in Figure 7b (all FDR < 0.05).

4 Discussion

Previously, AIM1L mRNA overexpression was identified in several cancer cell lines including prostate cancer, bladder inverted papilloma, bladder cancer, colon cancer, pancreas cancer, ovarian cancer, endometrial cancer, and breast cancer. While only placenta and testis exhibited high AIM1L mRNA expression among noncancerous tissues [46]. A pharmacogenomics study indicated that the risk of nausea and vomiting in opioid-treated cancer patients has a genetic component.

Whole exome sequencing of DNA pools revealed that six single nucleotide polymorphisms in some genes including AIM1L were associated with nausea and vomiting in opioid-treated cancer patients [47]. Unfortunately, impact of AIM1L on tumorigenesis is rarely investigated. Hence, the roles of AIM1L in the human cancers including HCC population are still essentially needed to be elucidated.

AIM1L mRNA was not detected in normal liver tissues [46], which is consistent with our results. AIM1L mRNA was significantly upregulated in tumor samples according to the findings from our bioinformatic study enrolled several public datasets. In addition, high AIM1L mRNA levels were significantly correlated with OS in HCC patients, both in TCGA and ICGC databases.

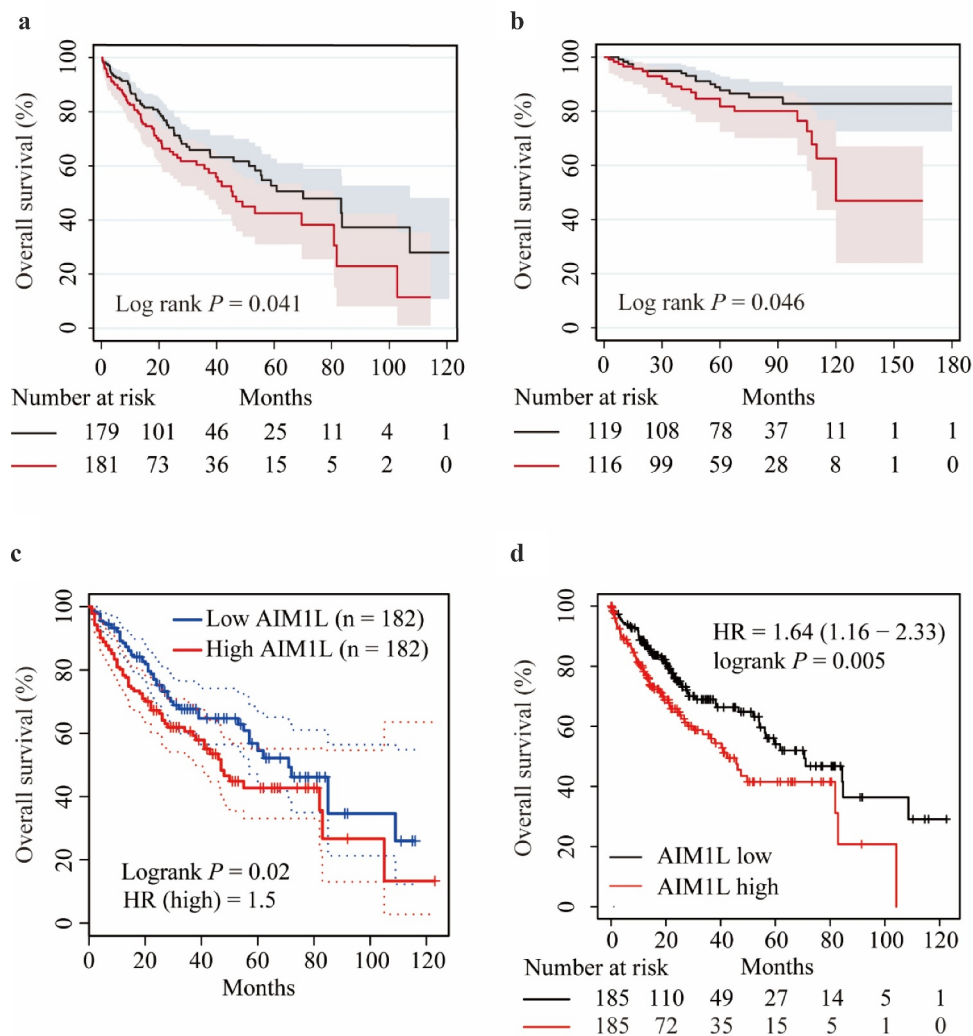
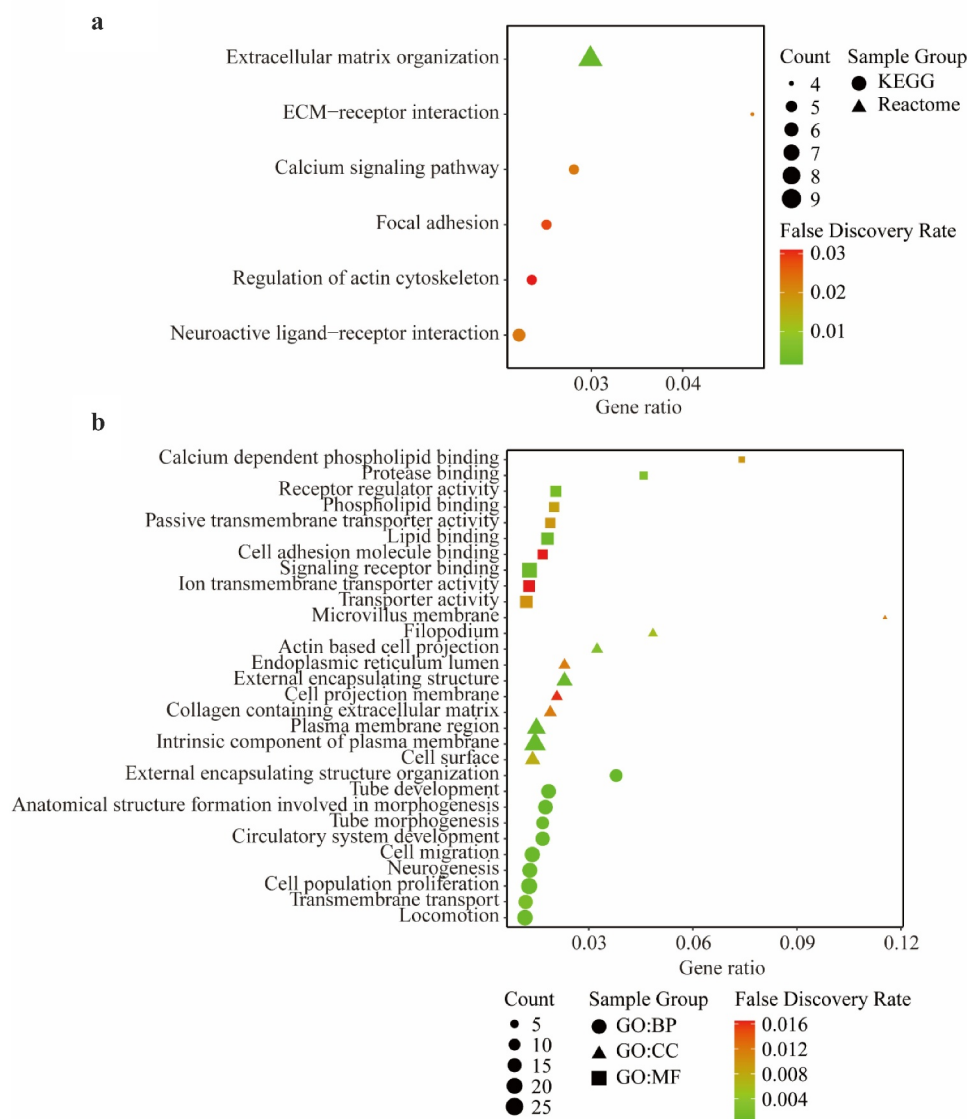


Figure 6. Associations between AIM1L and HCC survival. Kapan-Meier plot method indicated that HCC patients with AIM1L overexpression in tumor tissues had unfavorable OS compared to those with low levels of AIM1L in TCGA dataset (Log rank $P = 0.041$, A), ICGC dataset (Log rank $P = 0.046$, B), GEPIA (HR = 1.5, Log rank $P = 0.02$, C) and Kapan-Meier plotter (HR = 1.64, 95% CI = 1.16–2.33, Log rank $P = 0.041$, D) datasets.

Our functional enrichment of AIM1L-related genes indicated that AIM1L and its-related genes might involve in cell proliferation, cell migration, and signaling pathways, namely, extracellular matrix organization, ECM-receptor interaction, calcium signaling pathway, etc. Since tumorigenesis roles of AIM1L in human cancers have not been reported in details, absent in melanoma 1 (AIM1) has been studies in various human malignancies [48–50]. AIM1 has been shown to be highly overexpressed in prostate cancer tissues, and cultured androgen-independent prostate cancer cells, indicating that AIM1 might be a potential therapeutic target for treatment of prostate cancer

[51]. Among 93 bladder cancers samples and 26 nonmalignant tissues, the frequencies of AIM1 methylation were significantly higher in tumors (84%) than that in nontumor tissues (27%) [52]. The AIM1 methylation was also shown to be correlated with nasopharyngeal carcinoma compared to controls [53]. And AIM1 methylation was significantly associated with invasive tumors [52]. Metastatic melanoma had higher frequency of AIM1 promoter hypermethylation than primary melanomas. Melanomas AIM1 methylation was correlated with disease-free survival (DFS) and OS in Stage I/II patients. Circulating methylated AIM1 was also detected in melanoma patients'



serum and was predictive of OS in Stage IV patients [49]. On the other hand, advanced prostate cancers usually own AIM1 deletion and reduced expression. AIM1 depletion in prostate epithelial cells increases cell migration and invasion, and anchorage-independent growth. AIM1 could also inhibit pro-invasive properties in benign prostate epithelium. Moreover, AIM1 downregulation results in high risk of metastatic dissemination in vivo [48]. AIM1 has also been identified as a potential suppressor candidate of human malignant melanoma [50]. Considered

controversial results of AIM1 in human cancers and little literatures of AIM1L available, we assumed that further research focusing AIM1L on tumorigenesis in cancers including HCC should be urgently conducted.

This bioinformatic study has some limitations. First, no experiments were performed to address the effects of AIM1L on hepatoma cellular functions. Second, no our own follow-up data of HCC patients were available, the predictive values of AIM1L for OS in HCC patients were not validated in prospective cohorts. Third, this analysis was

conducted at mRNA level, links between AIM1L protein and HCC prognosis was not investigated.

5 Conclusion

AIM1L mRNA is at a low level in normal liver tissues, and is upregulated in tumor samples. High levels of AIM1L in tumors contributed to unfavorable OS in HCC patients. The AIM1L-related genes might involve in cell proliferation and cell migration. based on our results, we cautiously draw a hypothesis that AIM1L is an oncogenic gene and promotes cancer progression in HCC patients. Considered few reports of AIM1L on HCC tumor pathological phenotypes, we suggest that experimental and clinical research of the impact of AIM1L on the development of HCC should be addressed in future.

Highlights

- AIM1L is low in normal liver and high in liver tumors.
- High AIM1L in tumors accounts for unfavorable OS in HCC patients.
- Enrichment analysis indicated that AIM1L might involve in cell proliferation and cell migration.

Acknowledgements

This work was supported by National Natural Science Foundation of China (81803901) and National Science and Technology Major Program (2018ZX10303502). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contribution

Wenliang Zhou: Writing - Original draft preparation, Data curation, Methodology. Yuan Zhang: Data curation, Methodology. Shixi Zhang: Supervision, Methodology, Writing- Reviewing and Editing, Funding acquisition. Zongguo Yang: Supervision, Conceptualization, Methodology, Writing- Reviewing and Editing, Funding acquisition.

Disclosure of potential conflicts of interest

No potential conflict of interest was reported by the author(s).

Data availability

Datasets of the current study are available from the NCBI Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>), The Cancer Genome Atlas (<https://portal.gdc.cancer.gov/>), and the International Cancer Genome Consortium (<https://daco.icgc.org/>) databases. All the datasets were available from the corresponding authors with reasonable request.

Funding

This work was supported by the National Natural Science Foundation of China [81803901]; National Science and Technology Major Program [2018ZX10303502].

ORCID

Zongguo Yang  <http://orcid.org/0000-0002-6623-4841>

References

- [1] Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136:E359–86.
- [2] Mejia JC, Pasko J. Primary Liver Cancers: intrahepatic Cholangiocarcinoma and Hepatocellular Carcinoma. *Surg Clin North Am*. 2020;100:535–549.
- [3] Omata M, Cheng AL, Kokudo N, et al. Asia-Pacific clinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update. *Hepatol Int*. 2017;11:317–370.
- [4] Petrick JL, Kelly SP, Altekruse SF, et al. Future of Hepatocellular Carcinoma Incidence in the United States Forecast Through 2030. *J Clin Oncol*. 2016;34:1787–1794.
- [5] Amaro CP, Tam VC. Management of hepatocellular carcinoma after progression on first-line systemic treatment: defining the optimal sequencing strategy in second line and beyond. *Curr Oncol*. 2020;27: S173–S80.
- [6] Lee HW, Cho KJ, Park JY. Current Status and Future Direction of Immunotherapy in Hepatocellular Carcinoma: what Do the Data Suggest? *Immune Netw*. 2020;20:e11.
- [7] Cardarelli-Leite L, Hadjivassiliou A, Klass D, et al. Current locoregional therapies and treatment strategies in hepatocellular carcinoma. *Curr Oncol*. 2020;27: S144–S51.

- [8] Pan Y, Chen H, Yu J. Biomarkers in Hepatocellular Carcinoma: current Status and Future Perspectives. *Biomedicines*. 2020;8(12):576.
- [9] Abreu P, Ferreira R, Mineli V, et al. Alternative Biomarkers to Predict Tumor Biology in Hepatocellular Carcinoma. *Anticancer Res*. 2020;40(12):6573–6784.
- [10] Kubota N, Fujiwara N, Hoshida Y. Clinical and Molecular Prediction of Hepatocellular Carcinoma Risk. *J Clin Med*. 2020;9(12):3843.
- [11] Xie S, Jiang X, Zhang J, et al. Identification of significant gene and pathways involved in HBV-related hepatocellular carcinoma by bioinformatics analysis. *PeerJ*. 2019;7:e7408.
- [12] Zhuang L, Yang Z, Meng Z. Upregulation of BUB1B, CCNB1, CDC7, CDC20, and MCM3 in Tumor Tissues Predicted Worse Overall Survival and Disease-Free Survival in Hepatocellular Carcinoma Patients. *Biomed Res Int*. 2018;2018:7897346.
- [13] Roh SW, Abell GC, Kim KH, et al. Comparing microarrays and next-generation sequencing technologies for microbial ecology research. *Trends Biotechnol*. 2010;28:291–299.
- [14] Huang W, Skanderup AJ, Lee CG. Advances in genomic hepatocellular carcinoma research. *Gigascience*. 2018;7. DOI:10.1093/gigascience/giy135
- [15] Teufel A. Bioinformatics and database resources in hepatology. *J Hepatol*. 2015;62:712–719.
- [16] Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. *Science*. 2013;339:1546–1558.
- [17] Wang T, Chen B, Meng T, et al. Identification and immunoprofiling of key prognostic genes in the tumor microenvironment of hepatocellular carcinoma. *Bioengineered*. 2021;12:1555–1575.
- [18] Xiong C, Wang G, Bai D. A novel prognostic models for identifying the risk of hepatocellular carcinoma based on epithelial-mesenchymal transition-associated genes. *Bioengineered*. 2020;11:1034–1046.
- [19] Samur MK. RTCGAToolbox: a new tool for exporting TCGA Firehose data. *PLoS One*. 2014;9:e106397.
- [20] Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med*. 1997;16:385–395.
- [21] Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multi-dimensional cancer genomics data. *Cancer Discov*. 2012;2:401–404.
- [22] Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013;6:pl1.
- [23] Shiraishi Y, Fujimoto A, Furuta M, et al. Integrated analysis of whole genome and transcriptome sequencing reveals diverse transcriptomic aberrations driven by somatic genomic changes in liver cancers. *PLoS One*. 2014;9:e114263.
- [24] Gautier L, Cope L, Bolstad BM, et al. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*. 2004;20:307–315.
- [25] Dong X, Lin L, Zhang R, et al. TOBMI: trans-omics block missing data imputation using a k-nearest neighbor weighted approach. *Bioinformatics*. 2019;35:1278–1283.
- [26] Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
- [27] Nikolayeva O, Robinson MD. edgeR for differential RNA-seq and ChIP-seq analysis: an application to stem cell biology. *Methods Mol Biol*. 2014;1150:45–79.
- [28] Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26:139–140.
- [29] Engebretsen S, Bohlin J. Statistical predictions with glmnet. *Clin Epigenetics*. 2019;11:123.
- [30] Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47:D607–D13.
- [31] Szklarczyk D, Santos A, Von Mering C, et al. STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. *Nucleic Acids Res*. 2016;44:D380–4.
- [32] Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017;45:W98–W102.
- [33] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102:15545–15550.
- [34] Liberzon A, Birger C, Thorvaldsdottir H, et al. The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Syst*. 2015;1(6):417–425.
- [35] Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27:1739–1740.
- [36] Menyhart O, Nagy A, Gyorffy B. Determining consistent prognostic biomarkers of overall survival and vascular invasion in hepatocellular carcinoma. *R Soc Open Sci*. 2018;5:181006.
- [37] Fagerberg L, Hallstrom BM, Oksvold P, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics*. 2014;13:397–406.
- [38] Wang HW, Hsieh TH, Huang SY, et al. Forfeited hepatogenesis program and increased embryonic stem cell traits in young hepatocellular carcinoma (HCC) comparing to elderly HCC. *BMC Genomics*. 2013;14:736.

- [39] Melis M, Diaz G, Kleiner DE, et al. Viral expression and molecular profiling in liver tissue versus microdissected hepatocytes in hepatitis B virus-associated hepatocellular carcinoma. *J Transl Med.* 2014;12:230.
- [40] Wang H, Huo X, Yang XR, et al. STAT3-mediated upregulation of lncRNA HOXD-AS1 as a ceRNA facilitates liver cancer metastasis by regulating SOX4. *Mol Cancer.* 2017;16:136.
- [41] Lee YS. Gene expression profile of hepatocellular carcinoma patients in Taiwan. *Gene Expression Omnibus.* 2017; Series GSE101685. : <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi>
- [42] Mas VR, Maluf DG, Archer KJ, et al. Genes involved in viral carcinogenesis and tumor initiation in hepatitis C virus-induced hepatocellular carcinoma. *Mol Med.* 2009;15:85–94.
- [43] Shimada S, Mogushi K, Akiyama Y, et al. Comprehensive molecular and immunological characterization of hepatocellular carcinoma. *EBioMedicine.* 2019;40:457–470.
- [44] Wang SM, Ooi LL, Hui KM. Identification and validation of a novel gene signature associated with the recurrence of human hepatocellular carcinoma. *Clin Cancer Res.* 2007;13:6275–6283.
- [45] Tsubota A, Matsumoto K, Mogushi K, et al. IQGAP1 and vimentin are key regulator genes in naturally occurring hepatotumorigenesis induced by oxidative stress. *Carcinogenesis.* 2010;31:504–511.
- [46] Kamata Y, Kuhara A, Iwamoto T, et al. Identification of HLA class I-binding peptides derived from unique cancer-associated proteins by mass spectrometric analysis. *Anticancer Res.* 2013;33:1853–1859.
- [47] Colombo F, Pintarelli G, Galvan A, et al. Identification of genetic polymorphisms modulating nausea and vomiting in two series of opioid-treated cancer patients. *Sci Rep.* 2020;10:542.
- [48] Haffner MC, Esopi DM, Chaux A, et al. AIM1 is an actin-binding protein that suppresses cell migration and micrometastatic dissemination. *Nat Commun.* 2017;8:142.
- [49] Hoshimoto S, Kuo CT, Chong KK, et al. AIM1 and LINE-1 epigenetic aberrations in tumor and serum relate to melanoma progression and disease outcome. *J Invest Dermatol.* 2012;132:1689–1697.
- [50] Ray ME, Wistow G, Su YA, et al. AIM1, a novel non-lens member of the betagamma-crystallin superfamily, is associated with the control of tumorigenicity in human malignant melanoma. *Proc Natl Acad Sci U S A.* 1997;94:3229–3234.
- [51] Vainio P, Mpindi JP, Kohonen P, et al. High-throughput transcriptomic and RNAi analysis identifies AIM1, ERGIC1, TMED3 and TPX2 as potential drug targets in prostate cancer. *PLoS One.* 2012;7:e39801.
- [52] Brait M, Begum S, Carvalho AL, et al. Aberrant promoter methylation of multiple genes during pathogenesis of bladder cancer. *Cancer Epidemiol Biomarkers Prev.* 2008;17:2786–2794.
- [53] Loyo M, Brait M, Kim MS, et al. A survey of methylated candidate tumor suppressor genes in nasopharyngeal carcinoma. *Int J Cancer.* 2011;128:1393–1403.