

RESEARCH

Open Access



A2M is a potential core gene in intrahepatic cholangiocarcinoma

Guanran Zhang¹, Xuyue Liu¹, Zhengyang Sun², Xiaoning Feng¹, Haiyan Wang³, Jing Hao¹ and Xiaoli Zhang^{1*}

Abstract

Background: Intrahepatic cholangiocarcinoma (ICC) is a type of malignant tumor ranking the second in the incidence of primary liver cancer following hepatocellular carcinoma. Both the morbidity and mortality have been increasing in recent years. Small duct type of ICC has potential therapeutic targets. But overall, the prognosis of patients with ICC is usually very poor.

Methods: To search latent therapeutic targets for ICC, we programmatically selected the five most suitable microarray datasets. Then, we made an analysis of these microarray datasets (GSE26566, GSE31370, GSE32958, GSE45001 and GSE76311) collected from the Gene Expression Omnibus (GEO) database. The GEO2R tool was effective to find out differentially expressed genes (DEGs) between ICC and normal tissue. Gene Ontology (GO) function and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were executed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) v 6.8. The Search Tool for the Retrieval of Interacting Genes (STRING) database was used to analyze protein–protein interaction of these DEGs and protein–protein interaction of these DEGs was modified by Cytoscape3.8.2. Survival analysis was performed using Gene Expression Profiling Interactive Analysis (GEPIA) online analysis tool.

Results: A total of 28 upregulated DEGs and 118 downregulated DEGs were screened out. Then twenty hub genes were selected according to the connectivity degree. The survival analysis results showed that A2M was closely related to the pathogenesis and prognosis of ICC and was a potential therapeutic target for ICC.

Conclusions: According to our study, low A2M expression in ICC compared to normal bile duct tissue was an adverse prognostic factor in ICC patients. The value of A2M in the treatment of ICC needs to be further studied.

Keywords: Intrahepatic cholangiocarcinoma, Hub genes, Gene expression profiling, A2M

Introduction

Intrahepatic cholangiocarcinoma (ICC) is defined as a type of malignant tumor originating from epithelium of secondary bile duct and its branches [1–3]. ICC is the second most familiar primary liver cancer with increasing incidence [4, 5]. Median overall survival (OS) for ICC patients is 12 to 18 months, with 5-year OS rates of less

than 5% [6–8]. Surgical excision plus adjuvant therapy is the main treatment methods at present, but only 15% of patients are qualified for surgery [9] due to its difficulty in detection. MsMab-1 that is an effective antibody against isocitrate dehydrogenase 1/2 (IDH1/2) mutation may be a therapeutic target for small duct type of ICC, but other subtypes still lack therapeutic targets [10]. Local treatments such as thermoablation, stereotactic radiotherapy and chemotherapy might prolong the survival time and improve the quality of life for some patients, but the overall prognosis is poor. Until now, ICC remains difficult to be cured and remains to be urgent to explore new therapeutic targets of ICC.

*Correspondence: zhangxiaoli@sdu.edu.cn

¹ Key Laboratory for Experimental Teratology of Ministry of Education, Department of Histology & Embryology, School of Basic Medical Sciences, Shandong University, Jinan 250012, Shandong, China
Full list of author information is available at the end of the article



In this research, we attempted to discover original prognostic index for ICC patients and struggled for supplying potential therapeutic targets. We analyzed the gene expression profiling data from the Gene Expression Omnibus (GEO) database by bioinformatics technique to dig out the DEGs between normal human tissue and ICC. Then, we performed Gene Ontology function and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis of DEGs. After that, a protein–protein interaction (PPI) network was built on The Search Tool for the Retrieval of Interacting Networks Genes (STRING) database and hub genes of ICC were screened out by cytoscape (3.8.2). DAVID tools were used to make the functional analyses of hub genes. We performed survival analysis of these hub genes using the online tool Gene Expression Profiling Interactive Analysis (GEPIA). Finally, A2M was screened out. All in all, the purpose of this study was to discover the biomarkers that was used for diagnosis, clinical treatment, and monitoring disease progression by analyzing the gene changes that took place during disease progression to improve the comprehension of the mechanism of ICC.

Materials and methods

Data source

The gene expression datasets in this research were obtained from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). All of the datasets were freely downloaded, and no experiments on humans or animals were done by any of the authors.

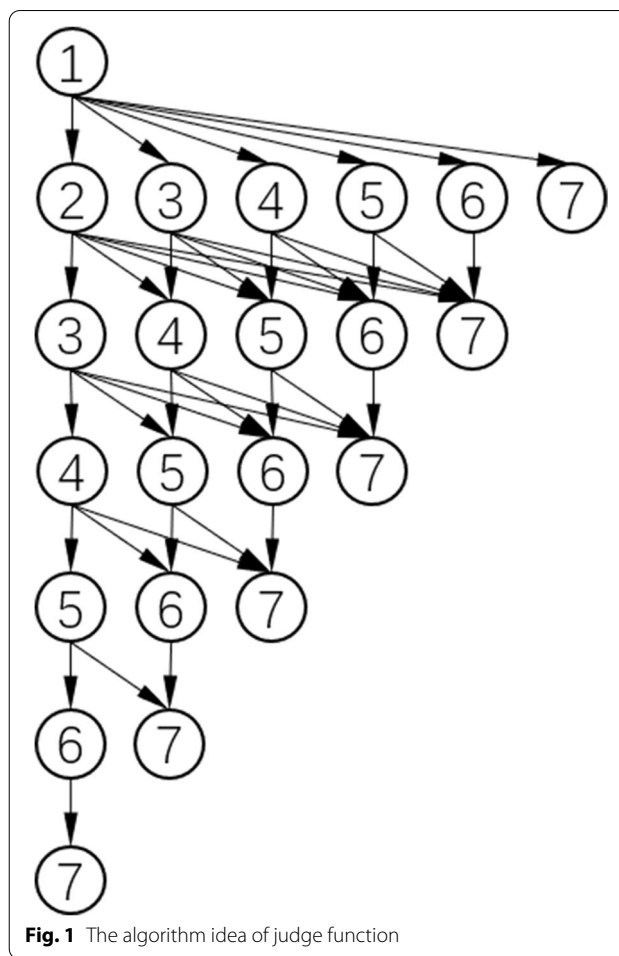
Algorithm idea of judge function

We used Excel to filter the data obtained from GEO database, and used $\text{Padj} < 0.05$ and $|\log\text{FC}| \geq 1.0$ as the standard to obtain DEGs. Then we exported them as text files, so that each line contained only one gene symbol, which were standardized texts.

Digital nodes 1-7 represent 7 standardized texts. We start from any node and proceed along the arrow. We can only advance on the next line, but cannot advance within the same line or return to the previous line. All nodes on the path are all the normalized texts selected by the judge function (Fig. 1).

Comparison program steps

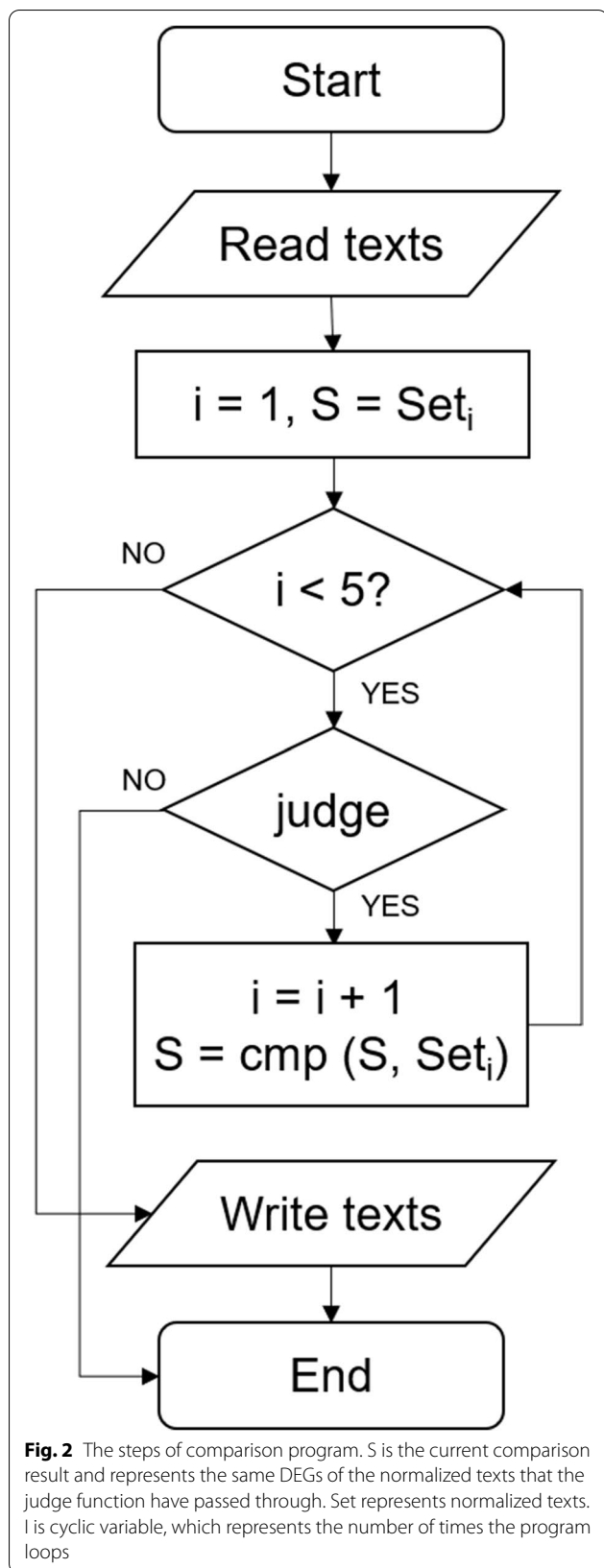
First of all, the comparison program read the standardized texts and determined whether the cyclic variable i was less than 5. If i was less than 5, we used the judge function to select the next dataset, and then used the cmp function to compare the DGEs of the standardized texts, and at the same time i value plus 1. Circulate according to the above method until i becomes 5, and when i was no longer less than 5, the comparison program output the



results (Fig. 2). In addition, when the length of the path generated by the judge function was too short without 5 nodes, the path was invalid, the program ended, and the results were not produced. In conclusion, the results generated by comparison program contained gene symbol of common DEGs.

Data selection

We found 2491 series about human ICC from the database. After careful selection, we screened out seven gene expression profiles (GSE26566, GSE31370, GSE32958, GSE33327, GSE89749, GSE45001 and GSE76311) that met the requirements. Then, we designed a program named comparison to process the seven gene expression datasets. The optimal results were obtained when five of the seven gene expression profiles were selected, and the highest number of differentially expressed genes (DEGs) was 146. Finally, five gene expression profiles (GSE26566, GSE31370, GSE32958, GSE45001 and GSE76311) were selected based on the number of differentially expressed genes (DEGs). Among them, GSE26566 was based on



GPL6104 Platform (Illumina humanRef-8 v2.0 expression beadchip), GSE31370 was based on GPL10558 Platform (Illumina HumanHT-12 V4.0 expression beadchip), GSE32958 was based on GPL6244 Platform ([HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]), GSE45001 was based on GPL14550 Platform (Agilent-028004 SurePrint G3 Human GE 8x60K Microarray) and GSE76311 was based on GPL17586 Platform ([HTA-2_0] Affymetrix Human Transcriptome Array 2.0 [transcript (gene) version]).

Data processing of DEGs

We used the GEO2R analysis tool (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>) to discover the DEGs between ICC and healthy samples and screen out the DEGs based on the value of adjusted P-value (Padj) and fold change (|logFC|). The cross-platform normalization was performed by normalizeBetweenArrays of limma package in Bioconductor. The genes with Padj < 0.05 and |logFC| ≥ 1.0 were selected as DEGs. The DEGs with log FC < -1.0 were defined as downregulated genes, while the DEGs with log FC > 1.0 were defined as upregulated genes. Then we selected the common part of DEGs by drawing the Venn diagrams using the webtool (bioinformatics.psb.ugent.be/webtools/Venn/).

GO and KEGG pathway analysis of DEGs

The results of large dimensions functional enrichment research that gene function is divided into biological process (BP), cellular component (CC) and molecular function (MF) are often obtained by GO analysis. KEGG is a database resource for understanding advanced functions and biological systems from molecular level information, particularly genomic sequencing generated from large molecular datasets. KEGG is one of the most commonly used databases of biological information in the world. We used Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 that was an online bioinformatics database with biological data and analytical tools, providing users with gene and protein functional annotation information to access biological information to get the results of GO analysis and KEGG pathway enrichment analysis of DEGs in this research (<https://david.ncifcrf.gov/>) [11]. The data that false discovery rate (FDR) < 0.01 and gene counts ≥ 10 were taken as statistically meaningful.

PPI network construction and hub genes identification

The Search Tool for the Retrieval of Interacting Networks Genes (STRING) database (<http://string-db.org/>) was used for the PPI networks functional enrichment analysis. The DEGs obtained previously were entered into the STRING database to explore possible PPI relationships.

The minimum required interaction score of PPI pairs was set as 0.400. After that, we used cytoscape software to modify the PPI network (www.cytoscape.org/). We used the plugin cytoHubba of cytoscape software to compute the connectivity degree of each node. Protein nodes with higher degree were more important in the whole PPI network. The top 20 genes with the highest connectivity degree were defined as hub genes. After that, we made the GO and KEGG pathway analysis of the selected hub genes using DAVID online analytical tools. The data of upregulated genes with $FDR < 0.01$ and gene counts ≥ 5 were taken as statistically meaningful and the data of downregulated genes with $FDR < 0.01$ and gene counts ≥ 10 were taken as statistically meaningful.

Survival analysis of hub genes

The Gene Expression Profiling Interactive Analysis (GEPIA) is a TCGA visualization tool developed by Zhang's Lab of Peking University (<http://gepia.cancer-pku.cn/index.html>). We used GEPIA to perform the survival analysis of ICC patients. ICC patients were divided into low group and high group according to the best cutoff values of mRNA expression calculated manually. $P < 0.05$ was considered to be statistically significant for disease free survival analysis and overall survival analysis.

Results

Selection of datasets

According to the type, number and year of cases, we artificially selected seven gene expression profiles (GSE26566, GSE31370, GSE32958, GSE33327, GSE89749, GSE45001 and GSE76311) that met the requirements from 2491 series about human ICC. Five gene expression profiles were screened out from the 7 gene expression profiles with a total of 21 combinations. After that, the comparison program we designed used the backtracking algorithm to traverse the 21 combinations when five gene expression profiles were selected from seven gene expression profiles, and listed the common DEGs of each combination. Then we showed the number of common DEGs that appeared in each combination (Fig. 3).

Identification of DEGs

We used the comparison program to pick out five microarray datasets (GSE26566, GSE31370, GSE32958, GSE45001 and GSE76311) and made the analysis in this study. Among them, GSE26566 included 104 ICC samples and 6 normal samples, GSE31370 contained 6 ICC samples and 5 normal samples, GSE32958 comprised 16 ICC samples and 7 normal samples, GSE45001 involved 10 ICC samples and 10 normal samples and GSE76311 contained 92 ICC samples and 93 normal

samples (Table 1). We obtained DEGs by comparing the gene expression level in ICC samples and normal samples. One thousand ninety-three DEGs that contained 542 upregulated genes and 551 downregulated genes were screened out from GSE26566 using $P < 0.05$ and $|\log FC| \geq 1.0$ as the selection criteria. One thousand six hundred twenty-four DEGs that involved 634 upregulated genes and 990 downregulated genes were picked out from GSE31370. From GSE32958, 2407 DEGs comprising 1172 upregulated genes and 1235 downregulated genes were identified. Two thousand eight hundred twenty-four DEGs were found from GSE45001, including 1492 upregulated genes and 1332 downregulated genes. And 1801 DEGs containing 771 upregulated genes and 1030 downregulated genes were screened out from GSE76311. Figure 4 shows the differential expression of genes in each of the five microarrays. Then, we drew the venn diagram to obtain the intersection of the DEGs in five databases (Fig. 5). Ultimately, we got 146 DEGs containing 28 upregulated genes and 118 downregulated genes of the selected five datasets.

Functional enrichment analyses of DEGs

The results of Gene Ontology function and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis for DEGs were obtained using the DAVID v6.8 (Tables 2 and 3). Gene Ontology includes Molecular Function (MF), biological process (BP) and cellular component (CC). From the results of GO analysis, we can draw a conclusion that the downregulated genes were principally concentrated in BPs and CCs, including oxidation-reduction process, negative regulation of endopeptidase activity, metabolic process, extracellular exosome, blood microparticle, extracellular space, extracellular region, mitochondrion and mitochondrial matrix. As for the molecular function, the downregulated genes were mainly related to the receptor binding. From the results of KEGG pathway enrichment analysis, we drew a conclusion that the downregulated genes were significantly concentrated in metabolic pathways, fatty acid degradation, biosynthesis of antibiotics, complement and coagulation cascades, peroxisome and carbon metabolism. As for the upregulated genes, they were principally concentrated in extracellular matrix organization, cell adhesion, ECM-receptor interaction, Focal adhesion and PI3K-Akt signaling pathway.

PPI network construction and hub genes identification

We used The Search Tool for the Retrieval of Interacting Networks Genes (STRING) to draw a protein-protein interaction (PPI) network and the network was modified by cytoscape software 3.8.2. There were 146 nodes and

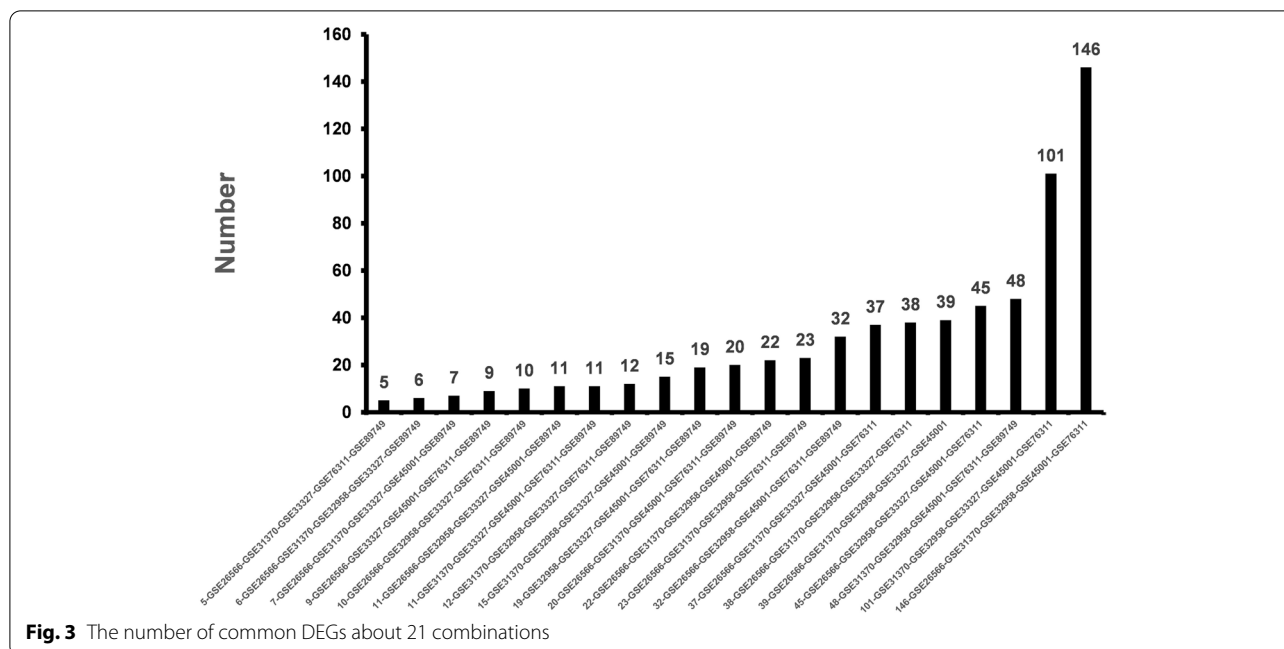


Table 1 Information of the five gene datasets from the GEO database

Dataset ID	ICC	Normal	Total number
GSE26566	104	6	110
GSE31370	6	5	11
GSE32958	16	7	23
GSE45001	10	10	20
GSE76311	92	93	185

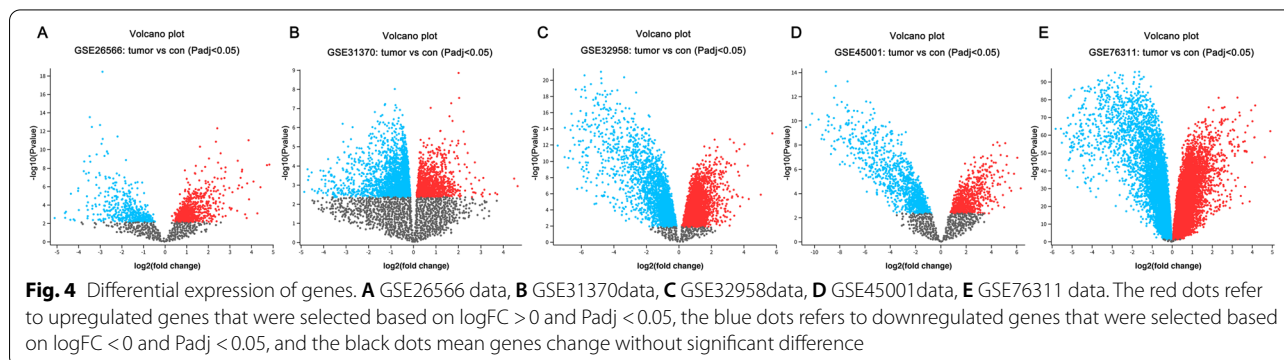
Abbreviations: *GEO* Gene Expression Omnibus, *ICC* intrahepatic cholangiocarcinoma

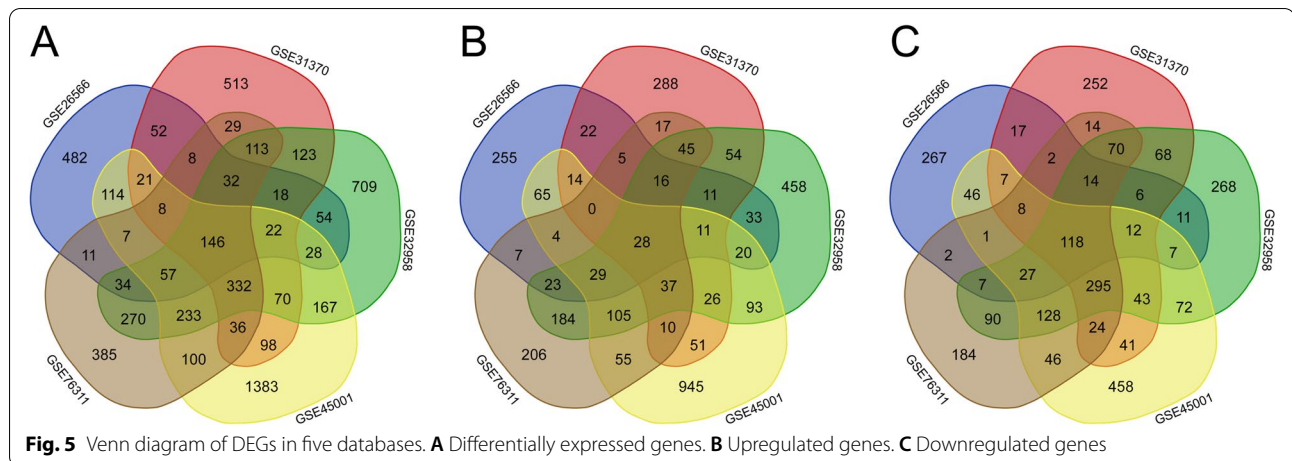
464 edges comprised in the PPI network (Fig. 6). Then, we calculated the degree of connectivity and screened the top 20 genes in the PPI network using cytoHubba App (Table 4). All the twenty hub genes were downregulated in ICC compared with that in normal liver tissue. The

rank of the twenty hub genes were listed in the Table 4. After that, we used DAVID tools to make the functional analyses of hub genes (Table 5). The results of GO analysis showed that hub genes were mainly enriched in CCs, including blood microparticle, extracellular region, extracellular space, extracellular exosome, peroxisomal matrix, platelet alpha granule lumen and intracellular membrane-bounded organelle. BP analysis indicated that the hub genes were enriched in platelet degranulation, negative regulation of endopeptidase activity and receptor-mediated endocytosis. And for the MF, the hub genes were enriched in receptor binding. Besides, the results of KEGG analysis indicated that hub genes were enriched in complement and coagulation cascades and peroxisome.

Survival analysis of twenty hub genes

The Gene Expression Profiling Interactive Analysis (GEPIA) tool was used to determine whether the twenty





potential key genes had the prognostic values. Data of only 36 ICC patients were available for the analysis of disease free survival analysis. The best cutoff value was selected manually. $P < 0.05$ was supposed to be statistically significant, and three of the twenty hub genes were

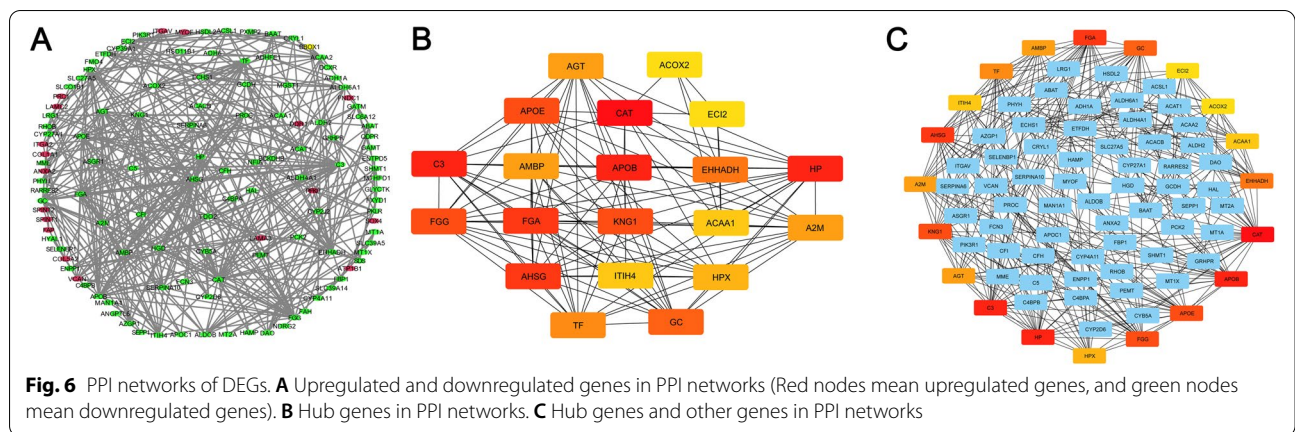
prognostic markers for the disease free survival analysis in ICC patients, but the survival analysis results of AGT and ITIH4 were contradictory to our previous analysis, that was the DEGs analysis results showed that they were downregulated in ICC compared with that in normal bile

Table 2 Results of GO and KEGG pathways enrichment analysis of upregulated genes

Category	Term	Description	Count	FDR
BP term	GO:0030198	Extracellular matrix organization	9	1.80E-08
BP term	GO:0007155	Cell adhesion	11	1.80E-08
KEGG pathway	hsa04512	ECM-receptor interaction	6	5.77E-06
KEGG pathway	hsa04510	Focal adhesion	6	2.11E-04
KEGG pathway	hsa04151	PI3K-Akt signaling pathway	6	0.001699684

Table 3 Results of GO and KEGG pathways enrichment analysis of downregulated genes

Category	Term	Description	Count	FDR
BP term	GO:0055114	Oxidation-reduction process	23	3.10E-08
BP term	GO:0010951	Negative regulation of endopeptidase activity	10	4.37E-05
BP term	GO:0008152	Metabolic process	10	2.83E-04
CC term	GO:0070062	Extracellular exosome	62	1.23E-18
CC term	GO:0072562	Blood microparticle	17	1.20E-13
CC term	GO:0005739	Mitochondrion	28	1.51E-06
CC term	GO:0005615	Extracellular space	28	1.54E-06
CC term	GO:0005576	Extracellular region	28	4.39E-05
CC term	GO:0005759	Mitochondrial matrix	12	1.10E-04
MF term	GO:0005102	Receptor binding	14	1.94E-04
KEGG pathway	hsa01100	Metabolic pathways	44	1.04E-11
KEGG pathway	hsa00071	Fatty acid degradation	11	1.47E-09
KEGG pathway	hsa04610	Complement and coagulation cascades	11	1.78E-07
KEGG pathway	hsa01130	Biosynthesis of antibiotics	16	3.55E-07
KEGG pathway	hsa04146	Peroxisome	10	7.44E-06
KEGG pathway	hsa01200	Carbon metabolism	10	8.53E-05



duct tissue, while the survival analysis showed they were favorable for the ICC patients when they were downregulated. Only A2M was the potential prognostic gene for disease free survival analysis in ICC patients (Fig. 7).

Discussion

Carcinoma of bile duct is divided into ICC and extrahepatic cholangiocarcinoma. According to the general type of tumor, ICC is divided into mass type, periductal infiltration type and intraductal growth type, and its incidence is second only to hepatocellular carcinoma in the primary liver cancer. In the past 20 years, the incidence

of ICC has been on the rise globally [12, 13]. At present, the main indicator of primary liver cancer (PLC) in clinical diagnosis is alpha fetoprotein (AFP), which has been widely used in the general survey and screening of high-risk groups [14]. In China, 30 to 40% of PLC patients are negative for AFP, while ICC patients are almost all negative for AFP [15]. Most ICC patients are in advanced stage when they are diagnosed due to the lack of specific early symptoms of cholangiocarcinoma and negative AFP. ICC is not sensitive to radiotherapy and chemotherapy, and lacks effective targeted drugs [16]. Currently, medical therapy can not significantly improve the therapeutic

Table 4 Rank of the top 20 genes in the PPI network

Rank	Gene symbol	Gene description	Degree
1	CAT	Catalase	28
2	APOB	Apolipoprotein B	24
2	HP	Aaptoglobin	24
2	C3	Complement component 3	24
5	FGA	Fibrinogen alpha chain	23
5	AHSG	α2-HS-glycoprotein	23
7	APOE	Apolipoprotein E	22
7	KNG1	Kininogen 1	22
7	FGG	Fibrinogen gamma chain	22
10	GC	Group-specific component (vitamin D binding protein)	21
11	EHHADH	Enoyl-CoA, hydratase/3-hydroxyacyl CoA dehydrogenase	19
12	TF	Transferrin	18
13	A2M	α2-macroglobulin	17
13	AMBP	Alpha-1-microglobulin	17
13	AGT	Angiotensinogen (serpin peptidase inhibitor, clade A, member 8)	17
16	HPX	Hemopexin	16
17	ITIH4	Inter-alpha-trypsin inhibitor heavy chain family, member 4	15
17	ACAA1	Acetyl-CoA acyltransferase 1	15
19	ACOX2	Acyl-CoA oxidase 2, branched chain	14
19	ECI2	Enoyl-CoA delta isomerase 2	14

effect and prolonged the survival period of patients [17, 18]. Radical resection is the only chance of long-term survival for ICC patients, but the tumor is locally invasive and often metastasizes, especially through the lymphatic system. Because of lacking prognostic marker, the prognosis of ICC patients is still very poor even after surgical resection [19–22]. Most ICC patients have lost the opportunity of radical resection since the diagnosis is too late [23]. For patients with ICC, the lack of early diagnostic index and prognostic indicators, poor efficacy of various treatment methods lead to poor clinical outcome. Hence, it is urgent to explore new diagnostic, therapeutic and prognostic targets of ICC.

The determination of the molecular mechanism about tumorigenesis is very important for the diagnosis and treatment of cancer patients [24]. There is an urgent need to study the pathogenesis of intrahepatic cholangiocarcinoma and the knowledge gained may help us to develop new clinical treatment strategies [25]. In recent years, high-throughput sequencing technology and bioinformatics analysis have been increasingly applied to biological research [26]. Bioinformatics is an interdisciplinary discipline, which uses bioinformatics methods to dig out data at the molecular level, and provides a new way for studying the molecular mechanism of various diseases [27]. Large amounts of data are stored in several common databases such as GEO (<https://www.ncbi.nlm.nih.gov/geo/>)

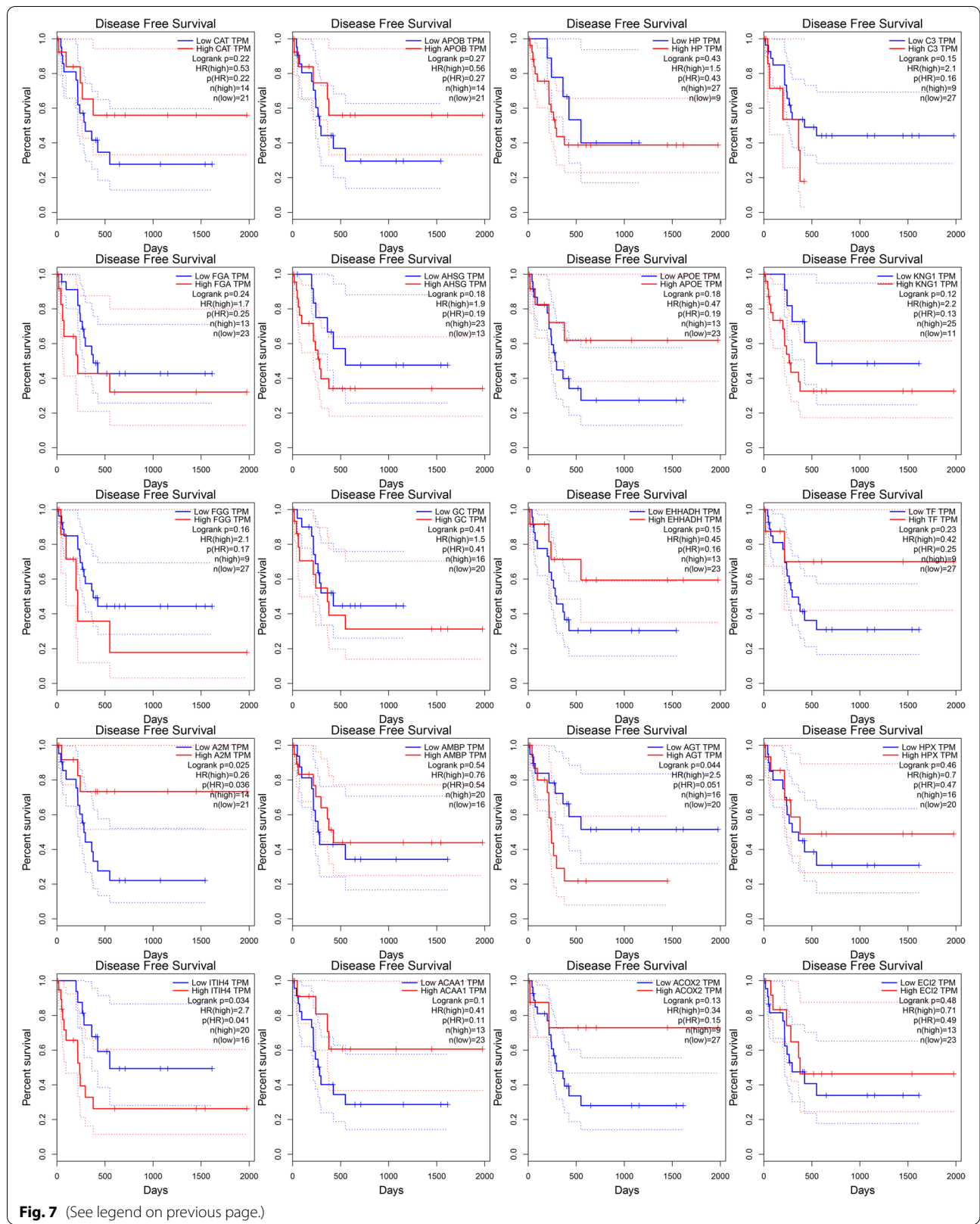
ICGC (<https://dcc.icgc.org/>) and ENCODE (<https://www.encodeproject.org/>) [28]. Compared with a single high throughput screening dataset, integrating some high throughput screening datasets (RNA sequencing and cDNA microarray) is regarded as a better way to improve the reliability of results [29–31]. In this study, an in-silico analysis was performed using bioinformatics methods to screen and identify new molecular targets. We firstly obtained DEGs by comparing ICC tissues with normal samples based on five microarray datasets that were selected by the comparison program from the GEO database. A total of 28 upregulated DEGs and 118 downregulated DEGs were identified and we performed GO and KEGG pathway enrichment analysis of DEGs. The downregulated genes were mainly enriched in BPs and CCs, including oxidation-reduction process, negative regulation of endopeptidase activity, extracellular exosome, metabolic process and so on, and significantly concentrated in the KEGG terms metabolic pathways, fatty acid degradation, biosynthesis of antibiotics, complement and coagulation cascades, peroxisome and carbon metabolism. The upregulated genes were principally concentrated in extracellular matrix organization, cell adhesion, ECM-receptor interaction, focal adhesion and PI3K-Akt signaling pathway. Then, a PPI network was built to research the correlation of the DEGs, and twenty hub genes that were all downregulated in ICC were

Table 5 Functional analyses of hub genes

Category	Term	Description	Count	FDR
BP term	GO:0002576	Platelet degranulation	7	3.68E-07
BP term	GO:0010951	Negative regulation of endopeptidase activity	7	4.89E-07
BP term	GO:0006898	Receptor-mediated endocytosis	5	0.003905848
CC term	GO:0072562	Blood microparticle	14	7.69E-22
CC term	GO:0005576	Extracellular region	15	3.37E-10
CC term	GO:0005615	Extracellular space	14	5.85E-10
CC term	GO:0070062	Extracellular exosome	16	1.78E-08
CC term	GO:0005782	Peroxisomal matrix	5	1.42E-06
CC term	GO:0031093	Platelet alpha granule lumen	5	2.46E-06
CC term	GO:0043231	Intracellular membrane-bounded organelle	6	0.001267314
MF term	GO:0005102	Receptor binding	9	1.63E-07
KEGG pathway	hsa04610	Complement and coagulation cascades	5	2.29E-04
KEGG pathway	hsa04146	Peroxisome	5	2.40E-04

(See figure on next page.)

Fig. 7 Disease free survival analyses of the top twenty hub genes in ICC patients. Abbreviations: CAT, catalase; APOB, apolipoprotein B; HP, haptoglobin; C3, complement component 3; FGA, fibrinogen alpha chain; AHSG, α 2-HS-glycoprotein; APOE, apolipoprotein E; KNG1, kininogen 1; FGG, fibrinogen gamma chain; GC, group-specific component (vitamin D binding protein); EHHADH, enoyl-CoA hydratase/3-hydroxyacyl CoA dehydrogenase; TF, transferrin; A2M, α 2-macroglobulin; AMBP, alpha-1-microglobulin; AGT, angiotensinogen (serpin peptidase inhibitor, clade A, member 8); HPX, hemopexin; ITIH4, inter-alpha-trypsin inhibitor heavy chain family, member 4; ACAA1, acetyl-CoA acyltransferase 1; ACOX2, acyl-CoA oxidase 2, branched chain; ECI2, enoyl-CoA delta isomerase 2



discovered by cytoscape 3.8.2, including CAT, APOB, HP, C3, FGA, AHSB, APOE, KNG1, FGG, GC, EHHADH, TF, A2M, AMBP, AGT, HPX, ITIH4, ACAA1, ACOX2 and ECI2. Finally, we used The Gene Expression Profiling Interactive Analysis (GEPIA) online tool to explore the relationship between hub genes and prognosis of ICC patients. Based on the results, overexpression of A2M was related to favorable prognosis of ICC patients. Therefore, overexpression of A2M might be a favorable prognostic factor of ICC patients.

A2M, also known as α 2-macroglobulin, is a core modulator in controlling protease activity and cell proliferation. α 2-macroglobulin acts as protease inhibitor, hormone, immune modulator and cytokine [32] and some research has demonstrated that α 2-macroglobulin as a kind of macromolecular plasma protein in the blood, α 2-macroglobulin can inactivate a variety of proteases by inhibiting plasmin and kallikrein [33, 34], and can also act as the carrier protein that binds to growth factors, hormones, and cytokines such as platelet derived growth factor (PDGF), basic fibroblast growth factor (bFGF), insulin like growth factor (IGF) and interleukin [35]. There is evidence that A2M can affect TGF- β 1 and other growth regulator ligands after binding to its receptor LRP1 [36]. Fears CY et al. showed that the combination of α 2-macroglobulin and LRP1 also phagocytosed a variety of matrix metalloproteinases, such as MMP-2, thereby inhibiting the migration and invasion of tumor cells [37].

In addition, α 2-macroglobulin is closely related to Alzheimer's disease. Alzheimer's disease (AD) is the most familiar neurodegenerative disease among the elderly people. α 2 macroglobulin that can be synthesized by astrocytes and neurons in the brain is a high affinity binding protein of amyloid β protein (A β), and its 27 amino acids at C-terminal specifically bind with A β peptide to neutralize A β toxicity [38, 39]. A2M can be divided into 6 fragments by restriction enzyme digestion and PCR methods, namely FP1 (aa99-392), FP2 (aa341-590), FP3 (aa591-744), FP4 (aa775-1059), FP5 (aa1030-1279) and FP6 (aa1242-1451). A β binds to FP6 segment with high specificity [40], suggesting that FP6 may become a new direction for the treatment of Alzheimer's disease [41].

Birkenmeier et al. showed that the decline of blood A2M in the elderly people was highly correlated with the incidence of tumors [42] and Lindner et al. reported that A2M binded with its receptor LRP1 to inhibit the Wnt/ β -catenin tumor signaling pathway [43]. Lauer et al. believed that A2M combined with growth factors to inactivate known tumor growth factors, thereby inhibiting tumor growth and invasion [44]. Wood et al. revealed the significance of A2M in the regulation of clock genes [45]. The mechanism of the relation between the

circadian clock and cancer is not clear, but destruction of circadian rhythm is associated with tumorigenesis [46]. There is evidence that A2M regulates tumor cell growth by upregulating PTEN and inhibiting tumour promoting signalling pathways such as PI3K/AKT, SMAD, and A2M is likely to become a new type of therapeutic drug [47]. Increasing the proportion of activated A2M in vivo has been considered for use in the treatment of cancer [48].

In short, the relationship between A2M and ICC has not been fully understood, and it is necessary to further explore the molecular mechanism of A2M and ICC. However, the anticancer effect of A2M suggests that overexpression of A2M might be associated with the better prognosis in patients with ICC, and in our study, the expression of A2M was low in ICC which was closely related to the poor prognosis of ICC. A2M is highly likely to be a therapeutic target for ICC.

Conclusion

We found 146 DEGs containing 28 upregulated genes and 118 downregulated genes between ICC and normal bile duct tissues based on the selected five datasets from the GEO database. Among them, A2M was the potential core gene of ICC. Overexpression of A2M was closely related to better prognosis in ICC patients. The results of our study need further research to confirm. In conclusion, A2M might be a potential target for the treatment of ICC.

Acknowledgements

The authors thank Associate Prof. Jing Liu for statistical consultation of this manuscript.

Authors' contributions

Xiaoli Zhang conceived of the study, Guanran Zhang and Zhengyang Sun carried out the datasets selection and analyzed all the data, Xuyue Liu and Xiaoning Feng helped Guanran Zhang to analyze data and preliminarily review the manuscript; Haiyan Wang and Jing Hao further reviewed and modified the manuscript. Xiaoli Zhang gave the final critical review of the manuscript and acts as the corresponding author. All authors have read and approved the final manuscript.

Authors' information

All authors' information are listed in this manuscript.

Funding

This study was supported by National Natural Science Foundation of China, 81672861; Science and Technology Development Plan of Shandong Province, 2017GSF218029; and Natural Science Foundation of Shandong Province, ZR2019LZL009.

Availability of data and materials

The raw data of this study are freely available from the website <https://www.ncbi.nlm.nih.gov/geo>, and all the analyzed data are included in this manuscript.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The author(s) declare that they have no competing interests.

Author details

¹Key Laboratory for Experimental Teratology of Ministry of Education, Department of Histology & Embryology, School of Basic Medical Sciences, Shandong University, Jinan 250012, Shandong, China. ²School of Information Engineering, Zhengzhou University, Zhengzhou 450001, Henan, China. ³Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China.

Received: 3 March 2021 Accepted: 29 November 2021

Published online: 03 January 2022

References

- Xue B, Wu S, Zheng M, Jiang H, Chen J, Jiang Z, et al. Development and validation of a Radiomic-based model for prediction of intrahepatic Cholangiocarcinoma in patients with intrahepatic Lithiasis complicated by Imagologically diagnosed mass. *Front Oncol*. 2021;10:598253.
- Muñoz-Martínez S, Forner A. The tireless search to improve the prognostic assessment of intrahepatic cholangiocarcinoma: an urgent need. *Liver Int*. 2021;41:252–4.
- Sempoux C, Jibara G, Ward SC, Fan C, Qin L, Roayaie S, et al. Intrahepatic cholangiocarcinoma: new insights in pathology. *Semin Liver Dis*. 2011;31:49–60.
- Banales JM, Marin JJG, Lamarca A, Rodrigues PM, Khan SA, Roberts LR, et al. Cholangiocarcinoma 2020: the next horizon in mechanisms and management. *Nat Rev Gastroenterol Hepatol*. 2020;17:557–88.
- Khan SA, Tavolari S, Brandi G. Cholangiocarcinoma: epidemiology and risk factors. *Liver Int*. 2019;39:19–31.
- Shaib YH, Davila JA, McGlynn K, El-Serag, H.B. Rising incidence of intrahepatic cholangiocarcinoma in the United States: a true increase? *J Hepatol*. 2004;40:472–7.
- Ryerson AB, Ehemann CR, Altekruse SF, Ward JW, Jemal A, Sherman RL, et al. Annual Report to the Nation on the Status of Cancer, 1975–2012, featuring the increasing incidence of liver cancer. *Cancer*. 2016;122:1312–37.
- Lamarca A, Ross P, Wasan HS, Hubner RA, McNamara MG, Lopes A, et al. Advanced intrahepatic Cholangiocarcinoma: post hoc analysis of the ABC-01, –02, and –03 clinical trials. *J Natl Cancer Inst*. 2020;112:200–10.
- Buettner S, Vugt JLV, IJzermans JN, Koerkamp BG. Intrahepatic cholangiocarcinoma: current perspectives. *Oncotargets Ther*. 2017;10:1131–42.
- Ma B, Meng H, Tian Y, Wang Y, Song T, Zhang T, et al. Distinct clinical and prognostic implication of IDH1/2 mutation and other most frequent mutations in large duct and small duct subtypes of intrahepatic cholangiocarcinoma. *BMC Cancer*. 2020;20:318.
- Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaie J, et al. The DAVID gene functional classification tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol*. 2007;8:R183.
- Saha SK, Zhu AX, Fuchs CS, Brooks GA. Forty-year trends in Cholangiocarcinoma incidence in the U.S.: intrahepatic disease on the rise. *Oncologist*. 2016;21:594–9.
- Bridgewater J, Galle PR, Khan SA, Llovet JM, Park J, Patel T, et al. Guidelines for the diagnosis and management of intrahepatic cholangiocarcinoma. *J Hepatol*. 2014;60:1268–89.
- Tang Z, Ye S, Liu Y, Qin L, Sun H, Ye Q, et al. A decade's studies on metastasis of hepatocellular carcinoma. *J Cancer Res Clin Oncol*. 2004;130:187–96.
- Sato Y, Ojima H, Onaya H, Mori T, Hiraoka N, Kishi Y, et al. Histopathological characteristics of hypervascular cholangiocellular carcinoma as an early stage of cholangiocellular carcinoma. *Hepatol Res*. 2014;44:1119–29.
- Blechacz B. Cholangiocarcinoma: current knowledge and new developments. *Gut Liver*. 2017;11:13–26.
- Malka D, Cervera P, Foulon S, Trarbach T, Fouchardière CDL, Boucher E, et al. Gemcitabine and oxaliplatin with or without cetuximab in advanced biliary-tract cancer (BINGO): a randomised, open-label, non-comparative phase 2 trial. *Lancet Oncol*. 2014;15:819–28.
- Ercolani G, Vetrone G, Grazi GL, Aramaki O, Cescon M, Ravaioli M, et al. Intrahepatic cholangiocarcinoma: primary liver resection and aggressive multimodal treatment of recurrence significantly prolong survival. *Ann Surg*. 2010;252:107–14.
- Yoh T, Cauchy F, Roy BL, Seo S, Taura K, Hobeika C, et al. Prognostic value of lymphadenectomy for long-term outcomes in node-negative intrahepatic cholangiocarcinoma: a multicenter study. *Surgery*. 2019;166:975–82.
- Nathan H, Pawlik TM. Staging of intrahepatic cholangiocarcinoma. *Curr Opin Gastroenterol*. 2010;26:269–73.
- Uenishi T, Hirohashi K, Kubo S, Yamamoto T, Yamazaki O, Kinoshita H. Clinicopathological factors predicting outcome after resection of mass-forming intrahepatic cholangiocarcinoma. *Br J Surg*. 2001;88:969–74.
- Inoue K, Makuuchi M, Takayama T, Torzilli G, Yamamoto J, Shimada K, et al. Long-term survival and prognostic factors in the surgical treatment of mass-forming type cholangiocarcinoma. *Surgery*. 2000;127:498–505.
- Weitz J, Rahbari N, Koch M, Büchler MW. The "artery first" approach for resection of pancreatic head cancer. *J Am Coll Surg*. 2010;210:e1–4.
- Chujan S, Suriyo T, Ungtrakul T, Pomyen Y, Satayavivad J. Potential candidate treatment agents for targeting of cholangiocarcinoma identified by gene expression profile analysis. *Biomed Rep*. 2018;9:42–52.
- Ye Z, Zeng Z, Wang D, Lei S, Shen Y, Chen Z. Identification of key genes associated with the progression of intrahepatic cholangiocarcinoma using weighted gene co-expression network analysis. *Oncol Lett*. 2020;20:483–94.
- Shu B, Zhang J, Sethuraman V, Cui G, Yi X, Zhong G. Transcriptome analysis of *Spodoptera frugiperda* Sf9 cells reveals putative apoptosis-related genes and a preliminary apoptosis mechanism induced by azadirachtin. *Sci Rep*. 2017;7:13231.
- Lowe R, Shirley N, Bleackley M, Dolan S, Shafee T. Transcriptomics technologies. *PLoS Comput Biol*. 2017;13:e1005457.
- Sungwan P, Lert-Itthiporn W, Silsirivanit A, Klinhom-On N, Okada S, Wongkham S, et al. Bioinformatics analysis identified CDC20 as a potential drug target for cholangiocarcinoma. *PeerJ*. 2021;9:e11067.
- Yan P, He Y, Xie K, Kong S, Zhao W. In silico analyses for potential key genes associated with gastric cancer. *PeerJ*. 2018;6:e6092.
- Le Y. Screening and identification of key candidate genes and pathways in myelodysplastic syndrome by bioinformatic analysis. *PeerJ*. 2019;7:e8162.
- Huang HM, Jiang X, Hao ML, Shan MJ, Qiu Y, Hu GF, et al. Identification of biomarkers in macrophages of atherosclerosis by microarray analysis. *Lipids Health Dis*. 2019;18:107.
- Perrot G, Langlois B, Devy J, Jeanne A, Verzeaux L, Almagro S, et al. LRP-1-CD44, a new cell surface complex regulating tumor cell adhesion. *Mol Cell Biol*. 2012;32:3293–307.
- Westwood M, Aplin JD, Collinge IA, Gill A, White A, Gibson JM. Alpha 2-macroglobulin: a new component in the insulin-like growth factor/insulin-like growth factor binding protein-1 axis. *J Biol Chem*. 2001;276:41668–78.
- Gopal U, Gonzalez-Gronow M, Pizzo SV. Activated α 2-macroglobulin regulates transcriptional activation of c-MYC target genes through cell surface GRP78 protein. *J Biol Chem*. 2016;291:10904–15.
- Arandjelovic S, Freed TA, Goniás SL. Growth factor-binding sequence in human α 2-macroglobulin targets the receptor-binding site in transforming growth factor- β . *Biochemistry*. 2003;42:6121–7.
- Lillis AP, Duyn LBV, Murphy-Ullrich JE, Strickland DK. LDL receptor-related protein 1: unique tissue-specific functions revealed by selective gene knockout studies. *Physiol Rev*. 2008;88:887–918.
- Fears CY, Grammer JR, Stewart JE, Annis DS, Mosher DF, Bornstein P, et al. Low-density lipoprotein receptor-related protein contributes to the antiangiogenic activity of thrombospondin-2 in a murine glioma model. *Cancer Res*. 2005;65:9338–46.
- Du Y, Bales KR, Dodel RC, Liu X, Glinn MA, Horn JW, et al. Alpha2-macroglobulin attenuates beta-amyloid peptide 1–40 fibril formation and associated neurotoxicity of cultured fetal rat cortical neurons. *J Neurochem*. 1998;70:1182–8.
- Hughes SR, Khorkova O, Goyal S, Knaeblein J, Heroux J, Riedel NG, et al. Alpha2-macroglobulin associates with beta-amyloid peptide and prevents fibril formation. *Proc Natl Acad Sci U S A*. 1998;95:3275–80.

40. Webb DJ, Wen J, Karns LR, Kurilla MG, Gonias SL. Localization of the binding site for transforming growth factor-beta in human alpha2-macroglobulin to a 20-kDa peptide that also contains the bait region. *J Biol Chem*. 1998;273:13339–46.
41. Mettenburg JM, Webb DJ, Gonias SL. Distinct binding sites in the structure of alpha 2-macroglobulin mediate the interaction with beta-amyloid peptide and growth factors. *J Biol Chem*. 2002;277:13338–45.
42. Birkenmeier G, Müller R, Huse K, Forberg J, Gläser C, Hedrich H, et al. Human alpha2-macroglobulin: genotype-phenotype relation. *Exp Neurol*. 2003;184:153–61.
43. Lindner I, Hemdan NYA, Buchold M, Huse K, Bigl M, Oerlecke I, et al. Alpha2-macroglobulin inhibits the malignant properties of astrocytoma cells by impeding beta-catenin signaling. *Cancer Res*. 2010;70:277–87.
44. Lauer D, Müller R, Cott C, Otto A, Naumann M, Birkenmeier G. Modulation of growth factor binding properties of alpha2-macroglobulin by enzyme therapy. *Cancer Chemother Pharmacol*. 2001;47:S4–9.
45. Wood PA, Yang X, Hrushesky WJM. Clock genes and cancer. *Integr Cancer Ther*. 2009;8:303–8.
46. Fu L, Lee CC. The circadian clock: pacemaker and tumour suppressor. *Nat Rev Cancer*. 2003;3:350–61.
47. Kurz S, Thieme R, Amberg R, Groth M, Jahnke H, Pieroh P, et al. Correction: the anti-tumorigenic activity of A2M—a lesson from the naked mole-rat. *PLoS One*. 2018;13:e0195169.
48. Cuéllar JM, Cuéllar VG, Scuderi GJ. α 2-macroglobulin: autologous protease inhibition technology. *Phys Med Rehabil Clin N Am*. 2016;27:909–18.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

