

SCIENTIFIC REPORTS



OPEN

Biostatistics mining associated method identifies AKR1B10 enhancing hepatocellular carcinoma cell growth and degenerated by miR-383-5p

Junqing Wang^{1,2,3}, Yunyun Zhou⁴, Xiaochun Fei⁵, Xuehua Chen^{2,3} & Yongjun Chen¹

Previous studies have reported that the aberrantly expressed AKR1B10 is associated with many cancer development, however the functional roles of AKR1B10 and its regulatory mechanisms in hepatocellular carcinoma (HCC) have been limited studied. In this project, we identified AKR1B10 functional as an oncogene in HCC through tumor/normal human tissue comparison from both GEO microarray and TCGA RNAseq dataset. Further experimental validations from three HCC cell lines (SMMC-7721, HePG2 and HeP3B) also suggested the ontogenetic functions of AKR1B10 in HCC tumor growth. By knocking down AKR1B10 through shRNA in HCC HeP3B cells, we showed it significantly induced cell cycle arrest and inhibited cell growth. Interestingly, integrative analysis of TCGA RNAseq data and miRNA-seq data predicted that miR-383-5p, a novel post-transcriptional tumor suppressor, is negatively associated with AKR1B10 expression. To further investigate the role of miR-383-5p in regulating AKR1B10 in HCC, we performed Dual-luciferase reporter assay experiments. Results showed that miR-383-5p is an upstream modulator targeting AKR1B10 in the post-transcriptional stage. Thus, we report AKR1B10 modulated regulated by miR-383-5p, promotes HCC tumor progress, and could be potentially a therapeutic target for precision medicine in HCC.

World widely, hepatocellular carcinoma (HCC) presents a high incidence rate among human malignancies, ranking a fifth of morbidity and second of malignancy-related mortality^{1,2}. Biomarkers aberrantly expressed in HCC at either mRNA or protein stages have been gradually illustrated for HCC tumorigenesis. However, comprehensive and integrative analysis for identifying new therapeutic targets and understanding their functional roles in tumor progression are critically important in the development of novel HCC cancer therapies for precision medicine.

Data mining the public database from Gene Expression Omnibus (GEO) database and the Cancer Genome Atlas (TCGA) database provides opportunities for detecting characteristic gene expressions in certain malignancies. In this study, we screened differentially expressed genes (DEGs) by tumor/normal comparison for GEO datasets and explored the significantly enriched Gene Ontology and KEGG pathway in HCC. Then we focused on 3 up-regulated and 11 down-regulated DEGs all shared among in both GEO microarray and TCGA RNAseq data for down streaming analysis. The Search Tool for the Retrieval of Interacting Genes (STRING) database was applied to construct the protein-protein interaction (PPI) network for the shared DEGs and their interacted genes. Among our shared DEGs, Aldo-Keto Reductase Family 1 Member B10 (AKR1B10) was observed as an

¹Department of Surgery, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197, Rui Jin Er Road, Shanghai, 20025, People's Republic of China. ²Shanghai Institute of Digestive Surgery, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197, Rui Jin Er Road, Shanghai, 20025, People's Republic of China. ³Shanghai Key Laboratory of Gastric Neoplasms, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197, Rui Jin Er Road, Shanghai, 20025, People's Republic of China. ⁴Department of Data Science, University of Mississippi Medical Center, Jackson, MS, 39216, USA. ⁵Department of Pathology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197, Rui Jin Er Road, Shanghai, 20025, People's Republic of China. Junqing Wang and Yunyun Zhou contributed equally to this work. Correspondence and requests for materials should be addressed to J.W. (email: wangjunqingmd@hotmail.com) or Y.C. (email: cylj10651@rjh.com.cn)

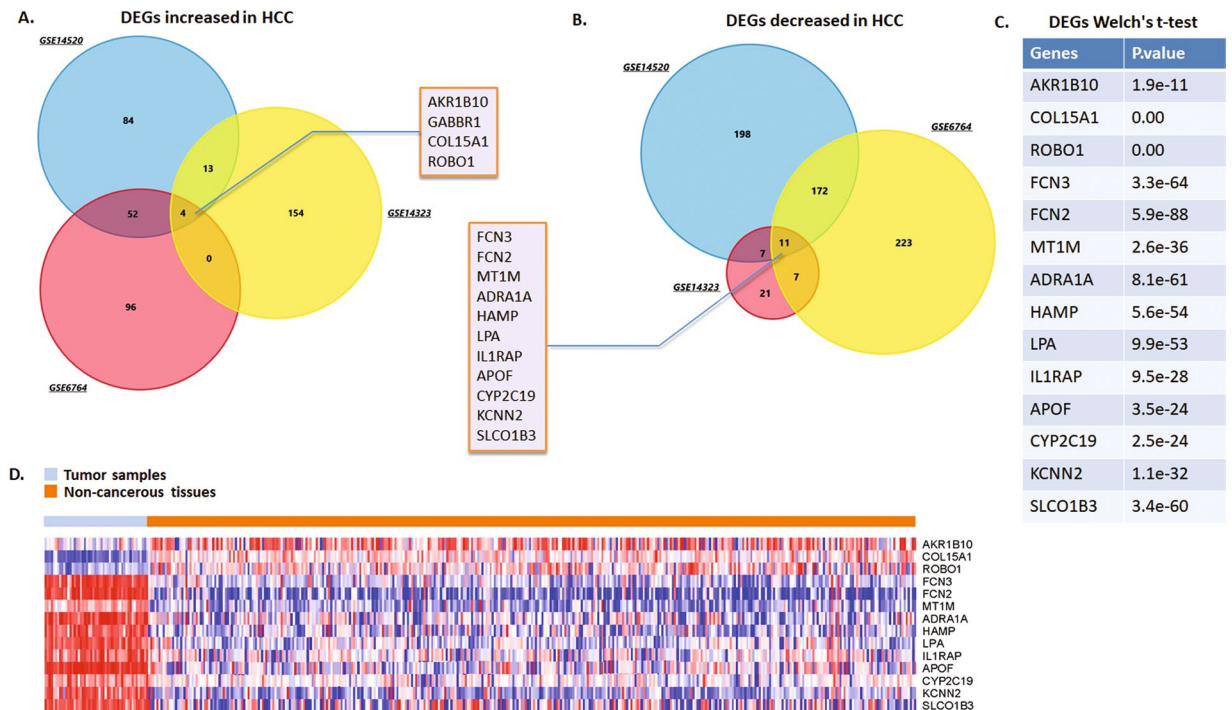


Figure 1. DEGs identified from NCBI GEO datasets **(A)** Venn chart of the over-expressing genes in GSE14520, GSE14323 and GSE6764 liver cancer datasets, which overlapped according to the analysis of FunRich software. Four DEGs extremely over-expressed in the three datasets were selected, including AKR1B10, GABBR1, COL15A1, AKR1B100. **(B)** Venn chart of the decreasing genes in three live cancer datasets. Eleven DEGs extremely decreased were selected, including FCN3, FCN2, MT1M, ADRA1A, HAMP, LPA, IL1RAP, APOF, YP2C19, KCNN2 and SLCO1B3. **(C)** The identified overlapped 15 DEGs from microarray were further tested in TCGA mRNA expression from RNAseq data by Welch's t-test with $P < 0.05$ as the cutoff for tumor and normal samples comparison. **(D)** Heatmap generated through TCGA database exploration and data mining. Expressions of the 15 DEGs above were presented by hot spots with gradient from red as high expression to blue as low expression.

oncogene participating in the progress of HCC. We explored and verified the expression status of AKR1B10 in both HCC cell lines and tissue samples from our medical center. Knock-down of AKR1B10 in HCC Hep3B cells significantly induced cell cycle arrest and cell growth obstruction.

AKR1B10 is a member of the Aldo/keto reductase super-family, which mainly reduce the aliphatic and aromatic aldehydes efficiently³. Accumulating evidence reveals that AKR1B10 functions as a pivotal promoter of human cancers in multiple tissues and organs, including breast, lung, liver, endometrium and gastrointestinal mucosa³⁻⁶. However, AKR1B10 in HCC has been controversially discussed, and AKR1B10 was described as either a negative prognostic factor or an independent promoting factor of HCC⁷⁻¹⁰. To investigate the AKR1B10 functions, in our study, we performed integrative analysis for TCGA mRNA and miRNA data and built a differentially expressed miRNA-mRNA regulatory network in HCC. Results suggested AKR1B10 is interacted with microRNA-383-5p (miR-383-5p). According to further validation, AKR1B10 was proved to be directly degenerated by miR-383-5p, a tumor suppressor gene in HCC. Herein, we suggest AKR1B10 as a biomarker promoting HCC targeted by miR-383-5p directly. Our findings provide us potential therapeutic targets in metabolism disorder related HCC treatment.

Results

Differentially expressed mRNA associated with HCC and pathway analysis. Based on the criterion of $\log_2FC \geq 1.5$ and P value $< 1.0E-04$ for selecting DEGs, 403 genes presented amplification in the HCC tissues compared to the non-cancerous samples. Moreover, other 639 genes decreased determined by the same criterion were set. We noticed significant enrichment of the processes closely related to cancer genesis and progress. Results suggested the most significant pathway is the cellular aldehyde metabolic process (P -Value: 2.46e-07) (Supp. Table 1). We further focused on 4 over-expressed genes (AKR1B10, GABBR1, COL15A1, ROBO10) out of 403 candidates (Fig. 1A), and 11 down-expressed genes (FCN3, FCN2, MT1M, ADRA1A, HAMP, LPA, IL1RAP, APOF, YP2C19, KCNN2, and SLCO1B3) shared among the three datasets (Fig. 1B). The identified overlapped 15 DEGs from microarray were further tested in TCGA mRNA expression from RNAseq data by Welch's t-test with $P < 0.05$ as the cutoff for tumor and normal samples comparison. Only the P value of GABBR1 gene did not pass the significant cutoff. Therefore we ignored this gene from further analysis (Fig. 1C). The mRNA expression pattern of the 14 DEGs in tumor and normal liver tissues was demonstrated by heatmap in Fig. 1D, through TCGA database exploration of 369 liver hepatocellular carcinoma samples and 49 adjacent specimens.

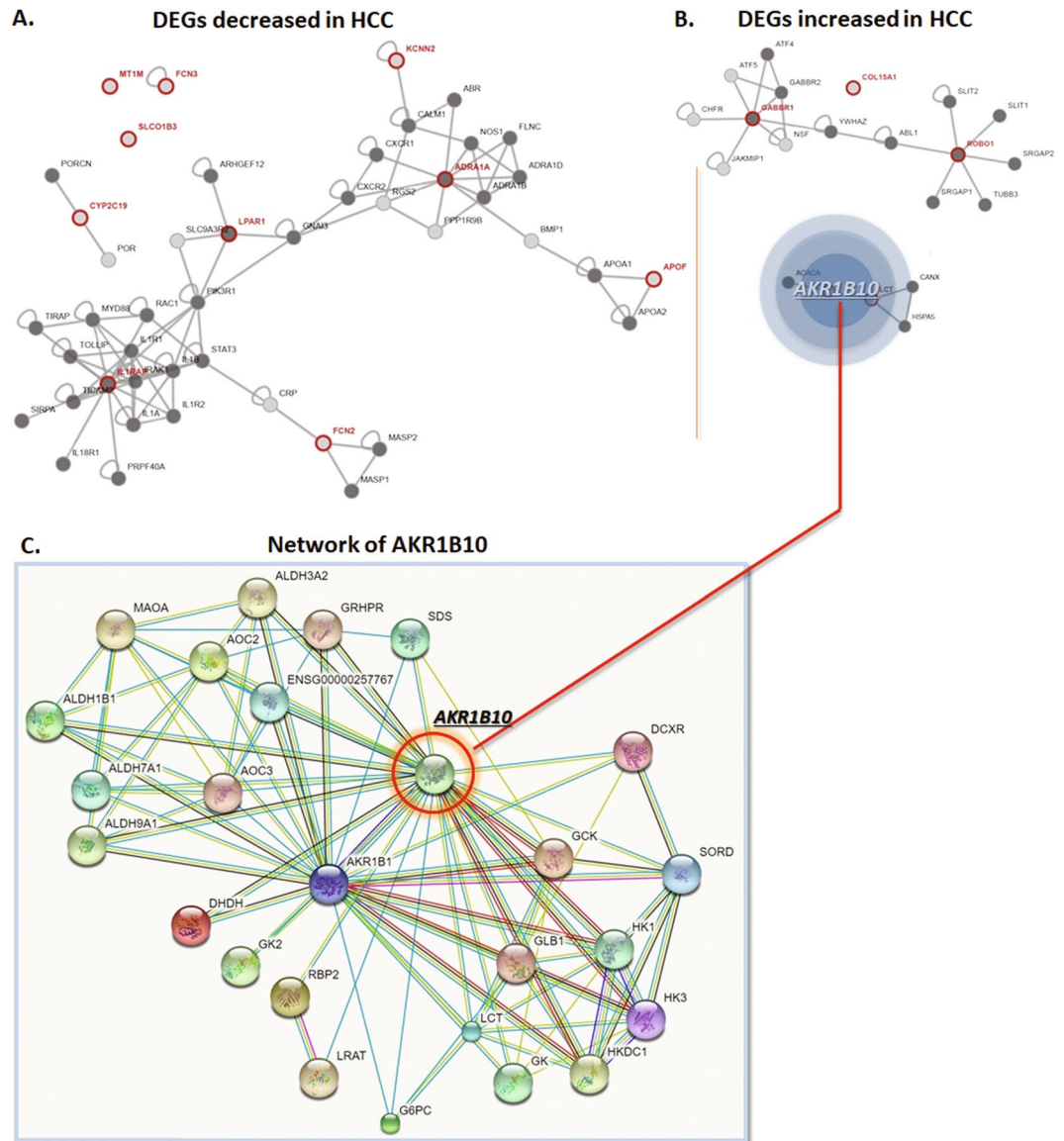


Figure 2. Protein-protein Interaction of the DEGs and selection of AKR1B10. **(A)** Online tool of STRING database analysis of the PPI network for the 11 DEGs decreased in liver cancer datasets. **(B)** Online tool of STRING database analysis of the PPI network for the 4 DEGs over-expressed in liver cancer datasets. **(C)** Amplification of the network for PPI associated with AKR1B10, which demonstrates a complex network between AKR1B10 and proteins critically functions in tumorigenesis and process.

Protein-protein interaction network for the overlapped DEGs from STRING database, and Cytoscape software was shown in Fig. 2A,B. The detailed statistic summaries for the interacted proteins of the DE mRNAs with the combined score cutoff 0.4 predicted by STRING database was shown in Supp. Table 2. Interestingly, AKR1B10, a member of aldo/keto reductase super-family, has been reported as key molecular reducing aliphatic and aromatic aldehydes for cell metabolism¹¹. This process is involved in physiological reactions such as the reduction of isoprenyl aldehydes, conversion of retinal to retinol, and bio-transformation of pro-carcinogens¹². The PPI network for AKR1B10 was shown in Fig. 2C.

Gene expression pattern of AKR1B10 in the HCC tissues and cell lines. As Fig. 3A showed, the initial exploration of the three databases demonstrated that AKR1B10 presented a significantly higher level of expression in tumor tissues compared to the normal liver tissues ($P = 3.50E-11$ for GSE14323, $P = 8.17E-7$ for GSE14520 and $P = 6.98E-9$ for GSE6764).

To validate the discoveries from the database mining, we explored the expression pattern of AKR1B10 by using IHC assay in our 53 paired HCC tumor and non-cancerous tissue samples. The cases were evaluated according to the expression intensity of AKR1B10, and were separated into low and high AKR1B10 expression group. For the tumor samples, 69.81% (37/53) of the cases were observed highly expressed AKR1B10, and only a portion of 30.19% (16/53) cases presented low AKR1B10 expression. For the non-cancerous specimens paired,

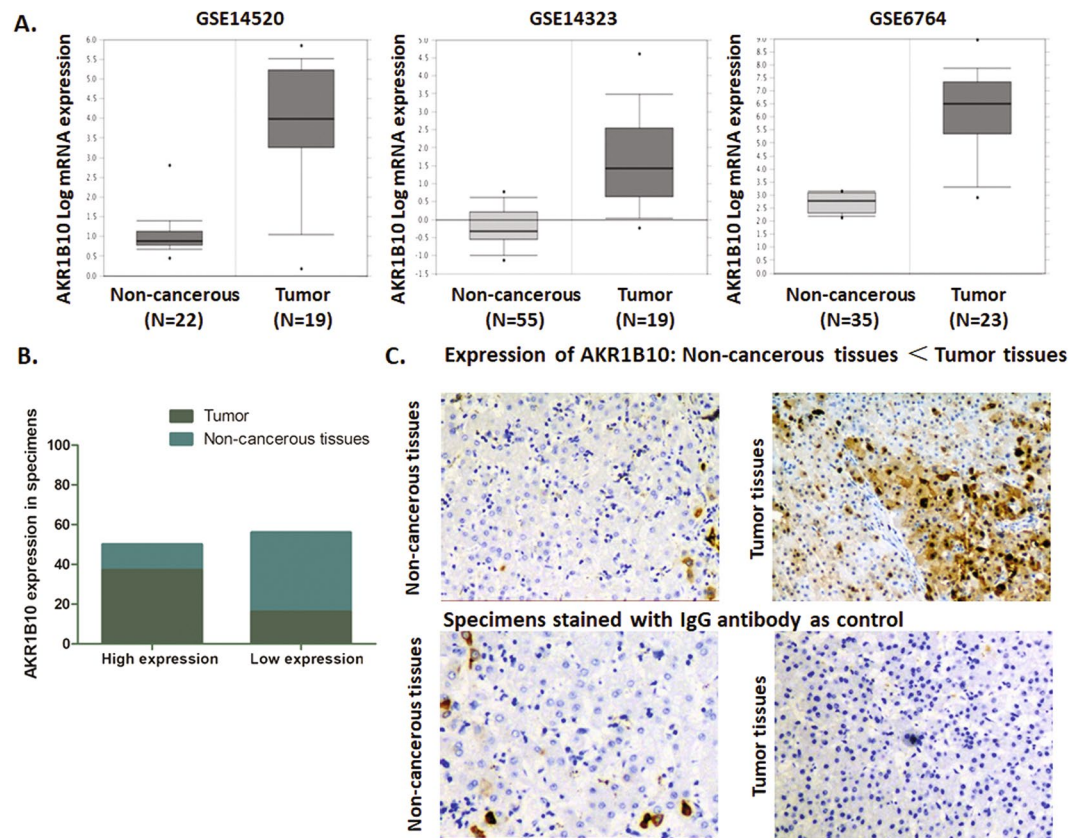


Figure 3. AKR1B10 expression profile in NCBI GEO datasets and HCC patients' specimens (A) Comparison of AKR1B10 mRNA expression status respectively in GSE14520, GSE14323 and GSE6764 liver cancer datasets through *t*-test. AKR1B10 was significantly higher expressed in HCC tissues than that in the non-cancerous tissues ($P = 3.50E-11$ for GSE14323, $P = 8.17E-7$ for GSE14520 and $P = 6.98E-9$ for GSE6764). (B) Expression of AKR1B10 was high in 69.81% out of 53 HCC tumor specimens (37/53). On the contrary, most of the non-cancerous tissues presented a relative lower AKR1B10 expression. The expression of AKR1B10 in tumor tissues was frequently and significantly higher than adjacent non-cancerous tissues ($P < 0.05$). (C) Representative graph immunohistochemistry analysis carried out on tissue microarray (200x).

highly expressed AKR1B10 was only observed in a portion of 24.53% (13/53) cases, with a 75.47% (40/53) portion presents in low AKR1B10 expression (Fig. 3B,C).

We also observed a similar pattern in three HCC cell lines (SMMC-7721, HeP3B and HePG2), that the mRNA and protein expression of AKR1B10 is significantly increased compared to the control QSG-7701 cells ($P < 0.05$) (Fig. 4A,B).

Knock-down of AKR1B10 suppresses cell proliferation and arrests the cell cycle in HeP3B cells. HCC HeP3B cells, presenting a high level of AKR1B10, was selected and transfected with pGU6/Neo vectors to knock down the expression of AKR1B10. Transfection effect was verified through qRT-PCR and Western blot analysis (Fig. 5A,B).

The CCK8 assay demonstrated a significant depression of cell proliferation in AKR1B10 knock down HeP3B cells as P value < 0.05 for d1 and d4, and P value < 0.01 for d2 to d3 (Fig. 5C). Simultaneously, we conducted flow cytometry analysis and found that the cell cycle of HeP3B cells was significantly arrested at G0/G1 phase when AKR1B10 knocked down (Fig. 5D–E). The percentage of HeP3B cells in G0/G1 phase was raised from 49.40% to 61.90% ($P < 0.01$). The S phase was decreased from 35.82% to 30.59%, and the G2/M phase was decreased from 14.78% to 7.51%. These results indicate that knock-down of AKR1B10 in HeP3B cells significant obstructs the tumor cell growth.

Differentially expressed miRNA associated with HCC and pathway analysis. As the Heatmap showed in Supp. Figure 2A, we identified 79 differentially expressed miRNAs, 47 down-expressed, and 32 up-expressed miRNAs, from TCGA HCC miRNA-Seq data based on the cutoff threshold (P value $< 1.0E-04$ and $\log_2 FC \geq 2$). The detailed statistics summary for the 79 DE miRNAs is shown in Supp. Table 3. The functional enrichment analysis for the 79 DE miRNAs was conducted by Funrich Software shown in Supp. Figure 2B–D. Interestingly, consistent with the significant pathway analysis from DE mRNAs, we found the most significant pathway for miRNAs is also related to cellular metabolism (P -Value: < 0.001) and transport process (P -Value: < 0.001) in a biological process. The results make sense and are consistent with many other scientific

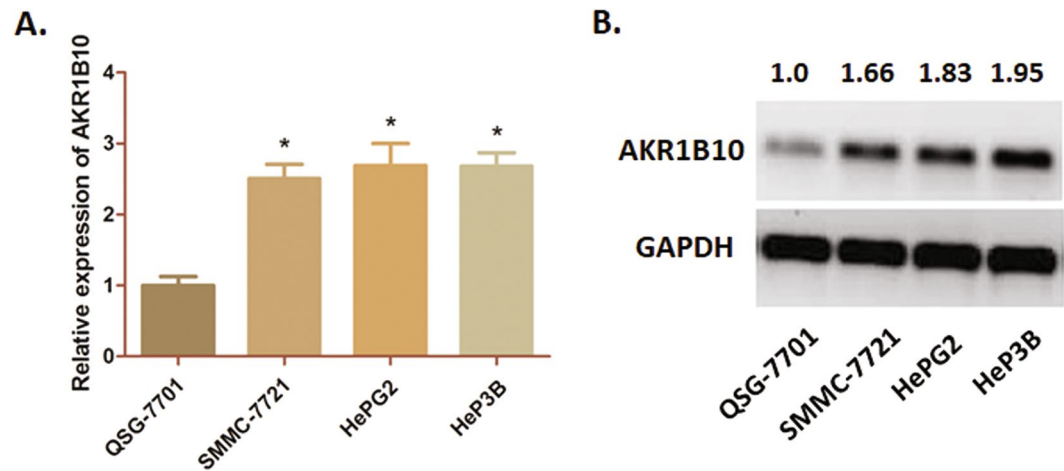


Figure 4. AKR1B10 expression profile in HCC cell lines (A) QRT-PCR assay was conducted in three HCC cell lines (SMMC-7721, HeP3B and HePG2) compared with QSG-7701 cells. The mRNA expression of AKR1B10 was significantly higher in HCC cells than that in the control one (* $P < 0.05$). (B) Western blot analysis was carried out, and the blot was cropped and matched according to the columns from the same gel. The protein expression of AKR1B10 was significantly higher in HCC cells than that in the control QSG-7701 cells. Numbers above the blot indicate normalized protein amounts relative to the negative control, as determined by densitometry.

findings since liver cancer was believed to be a type of disease due to metabolism disorder. Our results also suggested mRNA and miRNA interact each other in the metabolism process for the HCC development.

The integrative analysis identified differential miRNA regulatory network in HCC. To evaluate the regulatory relationship between differentially expressed miRNA and mRNA expressions in HCC, we conducted integrative network analysis. The integrative network for DE mRNAs and their interacted genes regulated by down-expressed miRNA and up-expressed miRNAs were shown in Supp. Figure 3A,B respectively. The final prediction of “core” network for cross-talked DE miRNA and mRNA in HCC development was shown in Supp. Figure 4A. We can see that AKR1B10 was predicted to be the target genes of miR-383-5p.

Our results suggested that miRNA deregulation is associated with several steps of cancer initiation and progression. Each miRNA can potentially directly or indirectly regulate the expression of many genes, and a single gene can be targeted by multiple miRNAs. For example, as shown in the Supp. Figure 4A, IL1RAP gene, targeted by several miRNAs, has been reported as an effective therapeutic target in many cancer^{13,14}; while miR-130a-5p targeting both ROBO1 and COL15A1, has been reported to be associated with cancer cell migration and implied a useful biomarker in chemotherapy.

The interacted miRNA-mRNA pairs can also be prognostic and predictive biomarkers. As the Kaplan-Meier curve based on HCC survival analysis shown in Supp. Figure 4B–C, high expression of AKR1B10 is associated with poor prognosis, and low expression of miR-383-5p is associated with poor prognosis.

MiR-383-5p post-transcriptional degenerates AKR1B10 in HeP3B cells. The expression status of miR-383-5p in differential human malignancies was detected through dbDEMC (Version 2.0) software, and indicated a relatively low expression of miR-383-5p in most of the human tumors including HCC (Fig. 6A). This finding is consistent with our differential expression analysis for miRNAs of TCGA HCC, which further confirmed that miR-383-5p is an inhibitor in HCC cells. As we predicted miR-383-5p as the upstream regulators of AKR1B10 from our network integrative analysis, we performed a qRT-PCR assay to verify our predictions. Results suggested that miR-383-5p in HeP3B cell lines were significantly lower expressed than that in QSG-7701 cells (Fig. 6B).

The potential interaction between miR-383-5p and 3'-UTR of AKR1B10 mRNA was illustrated as Fig. 6C. The Dual-luciferase reporter assay was practiced according to the estimated interacting sequences. The expression of AKR1B10 mRNA and protein appeared to be significantly decreased when miR-383-5p was introduced into HeP3B cells (Fig. 6D,E). Up-regulation of miR-383-5p in HeP3B cells (HeP3B/miR-383-5p) declined luciferase signal, which demonstrated a direct binding and modulation of miR-383-5p on AKR1B10 (Fig. 6F).

Discussion

As one of the highly aggressive malignancies of human beings, HCC causes suboptimal outcome of overall survival¹⁵. A large portion of the HCC patients were diagnosed in advanced stages, with no possibility for radical resection or presented poor response to chemotherapy¹⁶. Currently, molecule-targeted therapy provides us possible therapeutic strategy in anti-cancer treatment. While, as for HCC, there seems no effective molecule-targeted medicine as first-line treatment, except for sorafenib^{17,18}. NCBI GEO and TCGA databases, which are containing numerous information of gene expression in multiple human malignancies, provide us the possibility for integrative investigating and seeking for new targets with potential value in cancer treatment.

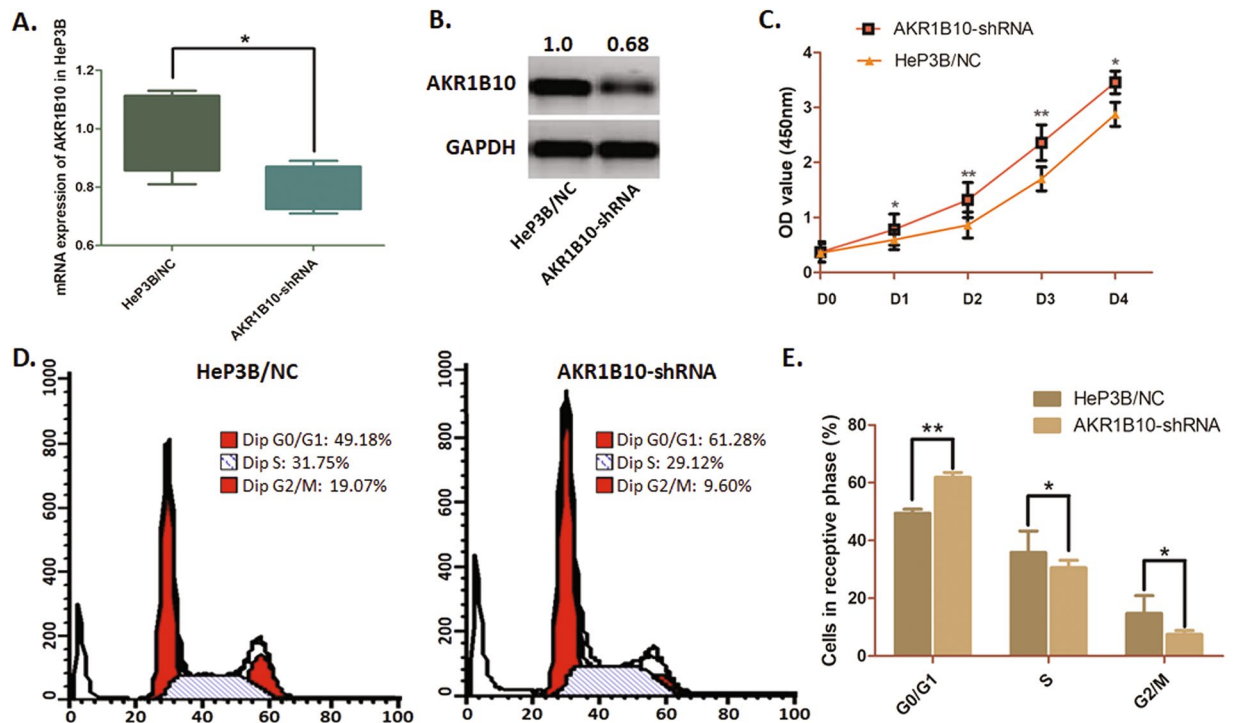


Figure 5. Knock-down AKR1B10 in HeP3B cells impairs cell proliferation and arrests cell cycle. (A) QRT-PCR assay verified the significant expression decline of AKR1B10 at mRNA status in HeP3B cells through shRNA transfection by using pGU6/Neo vectors (*P < 0.05). (B) Western-blot analysis indicated a significant decrease of AKR1B10 at protein status in HeP3B cells through shRNA transfection by using pGU6/Neo vectors. The blot was cropped and matched according to the columns from the same gel. (C) The effect of AKR1B10 on cell proliferation was determined by WST assay. The cell proliferation of HeP3B cells was significantly suppressed when AKR1B10 knocked down (*P < 0.05 for d1 and d4; **P < 0.01 for d2 and d3). (D) Representative histograms describing cell cycle profiles of HeP3B cells. The cell cycle was significantly arrested in G0/G1 phase when AKR1B10 knocked down. (E) Proportion of cells in various phases of the cell cycle. The results are means of three independent experiments (SD) (**P < 0.01, *P < 0.05).

In this study, we explored the HCC datasets with patients' information and gene expression profile from NCBI GEO database and TCGA database. DEGs with extremely differential expression profiles in HCC tissues compared with normal liver tissues were set into cohorts for further study. As we observed, 403 genes amplified and 639 ones decreased in HCC specimens were selected from three NCBI GEO datasets (GSE14323, GSE14520 and GSE6764). The results obtained by overlapping these genes provided us 15 genes aberrantly expressed in HCC tissues.

GO and KEGG enrichment pathway relative to the DEGs was detected. A series of pathways involving in tumorigenesis and progress were observed, including cellular aldehyde metabolic process. The oncogenetic functions of AKR1B10 have been reported in many other types of cancer by previous studies. For instance, the inhibition of AKR1B10 significantly suppressed the tumor progress in gastrointestinal cancer, pancreatic cancer, breast cancer etc^{19–21}, and over-expressed AKR1B10 induced cisplatin resistance enhancement by positively regulating peroxisome proliferator activated receptor- γ (PPAR γ)⁶. Some of the studies also suggested the tumor suppressor effect of AKR1B10 and its clinical relevance^{22,23}. Although several studies mentioned the over-expression characteristic of AKR1B10 associated with HCC progression^{24–26}, limited studies for the investigations of the regulatory mechanisms for the AKR1B10 interactions with other molecules are continuously needed so as to systematically interpret its functions in HCC development in depth.

In this study, we initially explored differentially expressed genes (DEGs) through patients' tumor and normal comparison from three NCBI GEO microarray datasets, which are GSE14323, GSE14520, and GSE6764. We explored all the DEGs functional pathways from system biology perspective. Results suggest the top one hit pathway for HCC development is metabolism related. Then, we validated the expression of DEGs in TCGA RNAseq data and focused on the investigation on the 14 DEGs shared in both three GEO microarray datasets and TCGA RNAseq dataset, which contains 3 up-regulated and 11 down-regulated genes. AKR1B10 is identified as one of the top hit over-expressed genes which is associated with metabolism disorder in HCC development.

Further experimental validations of the AKR1B10 expression and functions in either mRNA status or protein status was carried out on both human specimen and cell lines from our medical center. A similar pattern of AKR1B10 expression was observed. IHC assay presented a significantly higher expression of AKR1B10 in HCC tumor tissues compared to the adjacent non-cancerous tissues. Three HCC cell lines we detected also showed higher AKR1B10 than the control QSG-7701 cells. Then we explored the functions of AKR1B10 in HCC

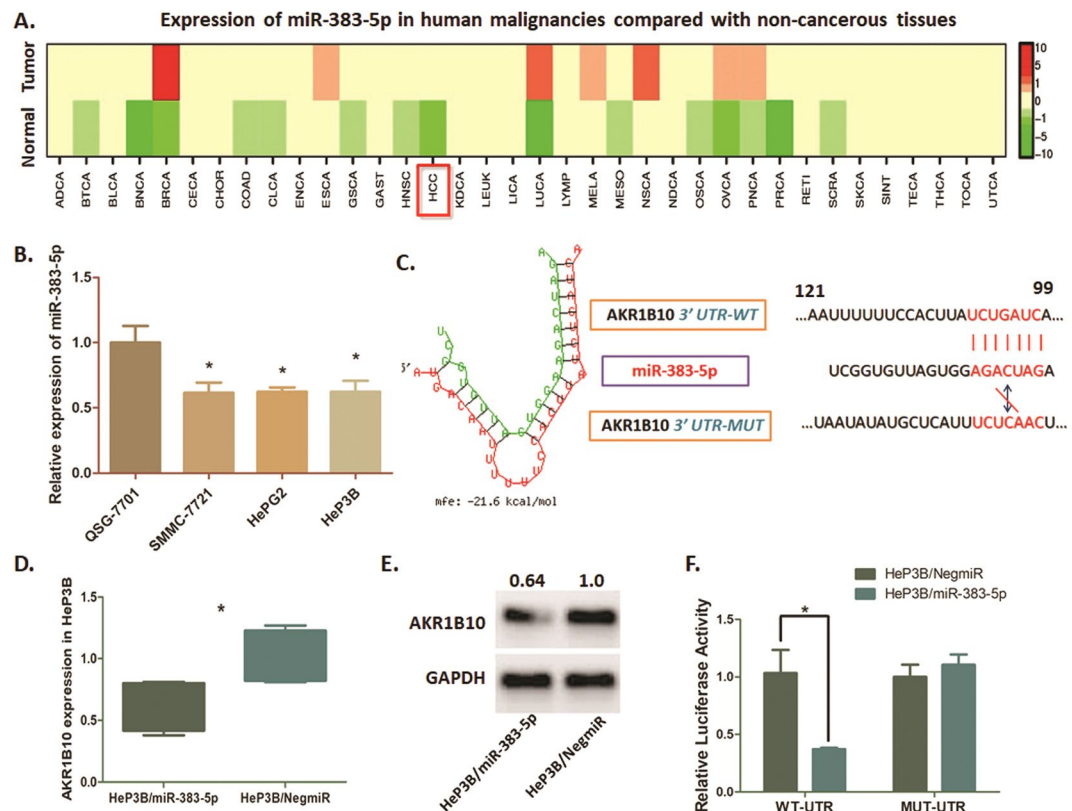


Figure 6. Degeneration of AKR1B10 in HeP3B cells directly regulated by miR-383-5p. **(A)** Expression profile heatmap of miR-383-5p in differential human malignancies generated by dbDEMC (Version 2.0) software. The bar indicates the expression level of objective genes by using color change, from green for low expression to red for high expression. MiR-383-5p is declined in HCC tissues compared with normal liver tissues. **(B)** MiR-383-5p expression detected by qRT-PCR. MiR-383-5p was significantly lower in HCC cells than that in QSG-7701 cells (* $P < 0.05$). **(C)** The left graph shows the minimum free energy (MFE) of the combination between miR-383-5p and AKR1B10 mRNA 3'-UTR, and the absolute number of this combination is 21.6 kcal/mol. The right graph is the scheme for the predicted miR-383-5p binding site in the wild type AKR1B10 mRNA 3'UTR (3'UTR-WT), and in the mutant construct (3'UTR-MUT). **(D)** QRT-PCR assay demonstrating the effect of miR-383-5p on AKR1B10 mRNA status in HeP3B cells. AKR1B10 expression was significantly suppressed by over-expressing miR-383-5p in HeP3B cells. **(E)** Western blot analysis shows the degeneration of AKR1B10 expression at protein status in HeP3B cells. AKR1B10 protein was significantly decreased after introducing miR-383-5p into HeP3B cells. The blot was cropped and matched according to the columns from the same gel. **(F)** The direct interaction between miR-383-5p and AKR1B10 was detected by Dual-luciferase reporter assay. Over-expression of miR-383-5p in HeP3B cells (HeP3B/miR-383-5p) significantly decreased the luciferase signal of AKR1B10/pMIR/WT compared with the negative control (HeP3B/NigmiR) (* $P < 0.05$). Mutation of the putative miR-383-5p-binding site abolished this suppressive effect.

tumorigenesis process by knocking down the AKR1B10 through shRNA in HCC HeP3B cells. As we observed, the mRNA and protein expression of AKR1B10 was significantly declined after transfection. Along with the expression impairing of AKR1B10, the cell growth of HeP3B cells presented to be meeting with a remarkable stagnation, and the cell cycle was observably arrested in G0/G1 phase. The findings above indicate that AKR1B10 functions to be a promotor inducing HeP3B cell growth in tumor process.

The molecular mechanism conferring to AKR1B10 amplification is valuable for our intensive understanding on HCC progress. Knowledgeably, post-transcriptional regulation is a pivotal approach leading to tumor promotion or suppression²⁷. To elucidate the regulatory functions of AKR1B10, we performed integrative network analysis from differentially expressed miRNA and mRNAs from TCGA HCC RNAseq data. We used the interactions data from miRWalk2.0 database to generate a global miRNAs and their target genes network. According to this analysis, we discovered miR-383-5p is the upstream molecular and interacted with AKR1B10 which was highly related to the metabolism pathway the as one of the most critical factors in the development of HCC by GO enrichment analysis and pathway analysis.

The miR-383-5p has been reported as a tumor suppressor in various human cancer including HCC^{28–31}. For example, miR-383-5p inhibited tumor cell mobility through regulating EPAS1 in lung cancer cells^{28–31}, the decrease of miR-383-5p is suggested as a potential independent biomarker for poor prognosis in intestinal subtype of gastric cancer²⁶, and the rescue of miR-383-5p expression in HCC cells prompted a better response

outcome by the sorafenib targeting therapy³². Our results from bioinformatics analysis and experimental validations in this project are consistent with their findings.

Our qRT-PCR assay demonstrated that miR-383-5p expression was significantly decreased in HCC HeP3B cell line. The interactions of miR-383-5p and AKR1B10 were tested by the Dual-luciferase reporter assay. The post-transcriptional regulatory relationship between miR-383-5p and 3'-UTR of AKR1B10 mRNA was evidenced by the decreased luciferase signal by building their direct interaction. Logically, degeneration of AKR1B10 in HCC cells could partly explain the tumor inhibiting mechanism that miR-383-5p induced in tumors. The survival analysis suggested the inverse association of expression alteration for miR-383-5p and AKR1B10 with survival outcome. This clinical evidence again confirmed the negative correlation between miR-383-5p and AKR1B10. Even, for the targeting therapy with sorafenib, the rescue of miR-383-5p expression in HCC cells prompts an approving outcome.

As summarized, integrative bioinformatics analysis through NCBI GEO database and TCGA database identified AKR1B10 as an oncogene interacted with miR-383-5p and suggested their involvement of metabolism pathway in HCC development. *In vitro* experiment combining with clinical specimen study verified the high expression of AKR1B10 in HCC specimens and cell lines. Knock-down AKR1B10 significantly declined HeP3B cells' ability of growth, which indicated AKR1B10 as a promotor in HCC process. Tumor suppressing miR-383-5p was verified as the up-stream regulator modulating AKR1B10 expression in the process of HCC development.

All these information suggested that AKR1B10 might be one of the pivotal molecules involved in HCC tumorigenesis and progress. More experimental validations to investigate other predicted miRNA regulatory functions can be performed in the future.

Methods and Materials Statements.

- (1) The experimental protocols were approved by Ruijin Hospital, Shanghai Jiao Tong University School of Medicine.
- (2) The methods applied in this study described below were carried out in accordance with relevant guidelines and regulations.
- (3) We declare that the informed consent was obtained from all the subjects. For any subject under 18 years old, the consent was obtained from a parent and/or the legal guardian.

Data resource and Description. The overall workflow of this study design is summarized as shown in Supp. Figure 1. Three expression microarray datasets GSE14323, GSE14520 and GSE6764 containing HCC tumor and normal samples were downloaded from the NCBI GEO dataset (<https://www.ncbi.nlm.nih.gov/geo/>). Platforms of GEO datasets for GSE14323 and GSE14520 dataset are GPL3921 (Affymetrix HT Human Genome U133A Array), and for GSE6764 is GPL570 (Affymetrix Human Genome U133 Plus 2.0 Array) (Agilent Technologies, Santa Clara, CA, USA). The GSE14323 dataset contains 74 samples with 55 HCC tumor and 19 normal samples; The GSE14520 dataset contains 44 samples with 22 tumor and 19 normal samples; and, the GSE6764 dataset contains 58 samples with 35 tumor and 23 normal samples.

TCGA liver hepatocellular carcinoma mRNA and miRNA mature strand RSEM normalized counts data derived from RNAseq Illumina HiSeq platform were downloaded from UCSC Xena database (<https://xena-browser.net/>). TCGA RNAseq data contains 418 samples with 369 primary tumor and 49 solid tissue normal samples. The miRNA mature ID was annotated by R miRBaseConverter package³³.

Bioinformatics analysis for identifying differentially expressed mRNAs in HCC microarray and their functional analysis. Raw .CEL files of the microarray from each GEO dataset were normalized by quantile method of Robust Multichip Analysis (RMA) from R affy package³⁴. Significant analysis for DEGs was performed by tumor and normal comparison from R samr package^{35,36}. The cutoff value for DEGs selection in each dataset is based on the following criteria: $\log_2 FC \geq 1.5$ and P value $< 1.0E-04$.

A combined DEGs list from all the three datasets, which is comprised of 403 upregulated and 639 down-regulated genes, were identified as a union. Gene Ontology and KEGG functional pathway enrichment were conducted by Database for Annotation Visualization and Integrated Discovery tools (DAVID).

We used Venn Gram to show the 4 up-regulated and 11 down-regulated genes shared among the three microarray datasets. The results of Protein-Protein Interaction analysis (PPI) for 15 DE mRNAs were performed by STRING database (<http://string-db.org>).

Surgical specimens and cell culture. Fifty-three HCC specimens from patients performed partial hepatectomy without preoperative therapy were detected by using immunohistochemistry analysis in our medical center during 2015–2016, paired with adjacent non-cancerous tissues.

Three HCC cell lines (SMMC-7721, HePG2 and HeP3B) were purchased from Shanghai Institutes for Biological Sciences, Chinese Academy of Science (Shanghai, China), along with the normal human hepatic cell line QSG-7701. Hep3B cells, which mimic miR-383-5p high-expression were constructed as we previously described for Dual-luciferase report assay, and the control ones were set^{37,38}. Cells culture medium were set, at 37°C with an atmosphere of 5% CO₂ in the humidified cell incubator, as the RPMI 1640, which was supplemented with 10% heat-inactivated fetal bovine serum (FBS), 100 ug/ml streptomycin and 100U/ml Penicillin.

Immunohistochemistry analysis and Western blot analysis. Antibodies against AKR1B10 were introduced (Abcam, USA) following the customer instruction. IgG antibody was utilized as control in IHC assay.

Two professional pathologists were assigned for inspecting the tissues blindly. RIPA buffer containing Protease Inhibitor Cocktail (Pierce, USA) was used to lyse the cells, and the protein concentration was measured by BCA Protein Assay Kit (Pierce, USA). Proteins were electrophoresed and electrotransferred. Antibody against AKR1B10 (1:1000) and GAPDH (1:5000) were probed. Horseradish peroxidase-conjugated secondary antibody was used for further probe. Protein quantity was detected by using GAPDH as a loading control.

RNA isolation and Real-time qPCR assay. TRIzol reagent (Invitrogen, USA) was used for total RNA was extracted from cell lines. The first-strand cDNA was synthesized using High-Capacity cDNA Reverse Transcription Kit (ABI, USA). RT-primers of AKR1B10 mRNAs were synthesized as follows: 5'-GACCCCTTGTGAGGAAAGCC-3' (forward) and 5'-ATTGCAACACGTTACAGGCC-3' (reverse) (Sangon Biotech Company, Shanghai, China). TaqMan Gene Expression Assays was utilized for Real-time quantitative polymerase chain reaction (qRT-PCR) (ABI, USA).

Cell transfection. HeP3B cells were transfected by pGU6/Neo vectors (GenePharma, Shanghai, China) containing shRNA suppressing AKR1B10 translation or non-containing ones. We cultured and selected the cells in medium containing 400 µg/ml G418 (Santa Cruz, USA). Stably transfected cells above were validated by qRT-PCR and Western blot analysis compared with the negative control cells. All cells were cultured and maintained in medium containing 200 µg/ml G418. MiR-383-5p was introduced into HeP3B cells through the mimic method for further Dual-luciferase Reporter Assay.

Cell proliferation assay and cell cycle analysis. 1×10^6 HeP3B cells stable transfected or the negative control cells were cultured in 96-well microtiter plates in triplicate and incubated for 5 days at 37 °C with an atmosphere of 5% CO₂. OD was measured by using microplate computer software (Bio-Rad Laboratories, USA) according to the protocol of CCK8 Assay Kit (Dojindo, Japan). The curves of cell proliferation were plotted.

Cells above were treated by ethanol fixation, RNase A treatment, and Propidium Iodide staining, and then were detected under flow cytometry by FACSCalibur (Becton Dickinson, USA). Cell populations at the G0/G1, S and G2/M phases were quantified by Modfit software (Becton Dickinson, USA) excluding a calculation of cell debris and fixation artifacts.

DEGs analysis for TCGA mRNA and miRNA expression and the functional analysis. The identified overlapped 15 DEGs from microarray were further tested in TCGA mRNA expression from RNAseq data by Welch's t-test with $P < 0.05$ as the cutoff for tumor and normal samples comparison. Only the P value of GABBR1 gene did not pass the significant cutoff. Therefore we removed this gene from further analysis.

2172 mature miRNAs strand expression profiling was carried out from TCGA tumor and normal comparison. After filtering miRNAs that were not expressed in most of the samples, 1270 miRNAs were used for further differential expression analyses by R samr package^{35,36} with default settings ($P < 1.0E-04$). Any miRNAs with less than 2 Fold change were further removed from the downstream analysis.

The miRNA –mRNA network integrative analysis. To further investigate the DE miRNAs regulatory mechanism in HCC, we generated an interaction network by integrating TCGA HCC DE miRNA-mRNA co-expression network and miRNA-target genes (TG) network based on database searching. For the construction of the TCGA miRNA-mRNA co-expression network, we built the affinity matrix based on the similarity cutoff of miRNA-mRNA pairs connectivity. We set the cutoff for selecting the significant miRNA and their co-expressed mRNA pairs based on the following criteria: Pearson correlation P -Value less than $P < 1.0E-04$. For building the miRNA-target genes (TG) epigenetic network, we mapped our TCGA DE miRNA to the miRWalk2.0 database, which curated the connectivity of miRNA with their targeted genes from 12 existing miRNA database such as MirTarbase, miRanda and TargetScan, to identify target genes of DE miRNAs from the TCGA HCC patient data. Totally 10753 genes mapping to the target genes of 79 DE miRNA were identified. Then we merged the co-expression DE miRNA-mRNA network with the miRNA-TG network by an iGraph tool from R package. Finally, the “core” network was refined using R package SpidermiR tool by only keeping the target genes overlapped with 14 DE mRNAs and their interacted genes predicted by STRING database with the cutoff value of combined score more than 0.4. The final HCC-specific miRNA-mRNA regulatory network was constructed by merging the co-expression weighted network derived from DEGs of TCGA HCC dataset and the miRNA-target genes (TG) network derived from database searching.

Experimental validation for Direct interaction between AKR1B10 mRNA and miR-383-5p. According to the analysis, microRNA-383-5p (miR-383-5p) appears directly binding to the 3' UTR of AKR1B10 mRNA. Dual-luciferase report assay was conducted to verify the exact binding between miR-383-5p and AKR1B10. A 198 bp sequence from the 3'-UTR of AKR1B10 mRNA containing the potential miR-383-5p binding site was set as follow: 5'-aggtgctgttttagacattttctgtatgttcaactaggatcagaatcacagaaaagcatggcttgaataaggaatgacaattttccattatctgatcagaacaaatgtttattaagcatcagaactctgccaacactgaggatgtaagatcaataaaaaataataatcataaccaaaaaaaa-3'. The corresponding mutant sequence were constructed by Sangon Biotech Company (Shanghai, China) as follow: 5'-gtgctccagatattgtcttattatgtcttctactagttgcaactgtaaaactctgtataccttgcgtagtaaatgcataagtcataatattgctcatttgcactgtagataagataaattacccttgacatagttgccg atctcagtgcaagaatacaactaaatataatattaaattgaaatcgatctatatatt-3'. The above sequences were respectively cloned into pMIR-REPORT luciferase vectors (Promega, USA), which contains Firefly luciferase. Moreover, the pRL-TK vectors, which contains Renilla luciferase was applied as a control. Hep3B cells transfected with miR-383-5p mimics or the negative control were co-transfected with the vectors. Luciferase activity was measured by using Dual-Glo Luciferase assay system (Promega, USA) 48 h post-transfection.

Survival analysis and statistics analysis for the experiment data. Survival analysis of AKR1B10 was evaluated through the application of the kmplot tool (<http://www.kmplot.com>) including 364 liver cancer patients with overall survival information. The survival analysis for miR-383-5p expression is from 614 liver cancer patients curated by kmplot with default settings. For the expression of both AKR1B10 and miR-383-5p, the best performing threshold was used for the Univariate Cox regression analysis after screening the lower and upper expression quartiles. The hazard ratio with 95% confidence and *P*-Value from the log-rank test were calculated by kmplot tool.

The experimental data were statistically analyzed by using SPSS 18.0. Fisher's exact test was applied to categorical outcomes; AVONA and t-test were performed on continuous outcomes from the experiments for *P* values calculation. A *P* value < 0.05 was considered as significant statistically.

References

- Torre, L. A. *et al.* Global cancer statistics, 2012. *CA: a cancer journal for clinicians* **65**, 87–108, <https://doi.org/10.3322/caac.21262> (2015).
- Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2015. *CA: a cancer journal for clinicians* **65**, 5–29, <https://doi.org/10.3322/caac.21254> (2015).
- Sinreih, M. *et al.* STAR and AKR1B10 are down-regulated in high-grade endometrial cancer. *The Journal of steroid biochemistry and molecular biology*, <https://doi.org/10.1016/j.jsbmb.2017.02.015> (2017).
- Suzuki, A., Matsushima, H., Satoh, A., Hayashi, H. & Sotobata, I. Prognostic implications of cardiac scintigraphic parameters obtained in the early phase of acute myocardial infarction. *Clinical cardiology* **11**, 370–376 (1988).
- Jung, Y. J. *et al.* AKR1B10-inhibitory Selaginella tamariscina extract and amentoflavone decrease the growth of A549 human lung cancer cells *in vitro* and *in vivo*. *Journal of ethnopharmacology* **202**, 78–84, <https://doi.org/10.1016/j.jep.2017.03.010> (2017).
- Matsunaga, T. *et al.* Aldo-keto reductase 1B10 promotes development of cisplatin resistance in gastrointestinal cancer cells through down-regulating peroxisome proliferator-activated receptor-gamma-dependent mechanism. *Chemico-biological interactions* **256**, 142–153, <https://doi.org/10.1016/j.cbi.2016.07.008> (2016).
- Sonohara, F. *et al.* Prognostic significance of AKR1B10 gene expression in hepatocellular carcinoma and surrounding non-tumorous liver tissue. *Oncology letters* **12**, 4821–4828, <https://doi.org/10.3892/ol.2016.5240> (2016).
- Matkowskyj, K. A. *et al.* Aldoketoreductase family 1B10 (AKR1B10) as a biomarker to distinguish hepatocellular carcinoma from benign liver lesions. *Human pathology* **45**, 834–843, <https://doi.org/10.1016/j.humpath.2013.12.002> (2014).
- Ha, S. Y. *et al.* High expression of aldo-keto reductase 1B10 is an independent predictor of favorable prognosis in patients with hepatocellular carcinoma. *Gut and liver* **8**, 648–654, <https://doi.org/10.5009/gnl13406> (2014).
- Murata, A. *et al.* Pretreatment AKR1B10 expression predicts the risk of hepatocellular carcinoma development after hepatitis C virus eradication. *World J Gastroenterol* **22**, 7569–7578, <https://doi.org/10.3748/wjg.v22.i33.7569> (2016).
- Cousido-Siah, A. *et al.* Structural analysis of sulindac as an inhibitor of aldose reductase and AKR1B10. *Chemico-biological interactions* **234**, 290–296, <https://doi.org/10.1016/j.cbi.2014.12.018> (2015).
- Zemanova, L. *et al.* Flavones Inhibit the Activity of AKR1B10, a Promising Therapeutic Target for Cancer Treatment. *Journal of natural products* **78**, 2666–2674, <https://doi.org/10.1021/acs.jnatprod.5b00616> (2015).
- Ågerstam, H. *et al.* Antibodies targeting human IL1RAP (IL1R3) show therapeutic effects in xenograft models of acute myeloid leukemia. *Proceedings of the National Academy of Sciences* **112**, 10786–10791 (2015).
- Liu, F. *et al.* SRSF10-mediated IL1RAP alternative splicing regulates cervical cancer oncogenesis via mIL1RAP-NF-κB-CD47 axis. *Oncogene*, **1** (2018).
- Yang, J. D. & Roberts, L. R. Hepatocellular carcinoma: A global view. *Nature reviews. Gastroenterology & hepatology* **7**, 448–458, <https://doi.org/10.1038/nrgastro.2010.100> (2010).
- Farazi, P. A. & DePinho, R. A. Hepatocellular carcinoma pathogenesis: from genes to environment. *Nature reviews. Cancer* **6**, 674–687, <https://doi.org/10.1038/nrc1934> (2006).
- Singh, S., Singh, P. P., Roberts, L. R. & Sanchez, W. Chemopreventive strategies in hepatocellular carcinoma. *Nature reviews. Gastroenterology & hepatology* **11**, 45–54, <https://doi.org/10.1038/nrgastro.2013.143> (2014).
- Cheng, S., Wei, X. & Wu, M. Effective ways to improve the prognosis of advanced stage (BCLC stage C) hepatocellular carcinoma. *Zhonghua Wai Ke Za Zhi* **53**, 324–327 (2015).
- He, Y. C. *et al.* Overexpression of AKR1B10 in nasopharyngeal carcinoma as a potential biomarker. *Cancer Biomark* **16**, 127–135, <https://doi.org/10.3233/CBM-150548> (2016).
- Chung, Y. T. *et al.* Overexpression and oncogenic function of aldo-keto reductase family 1B10 (AKR1B10) in pancreatic carcinoma. *Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc* **25**, 758–766, <https://doi.org/10.1038/modpathol.2011.191> (2012).
- Li, J. *et al.* AKR1B10 promotes breast cancer cell migration and invasion via activation of ERK signaling. *Oncotarget* **8**, 33694–33703, <https://doi.org/10.18632/oncotarget.16624> (2017).
- Ohashi, T., Idogawa, M., Sasaki, Y., Suzuki, H. & Tokino, T. AKR1B10, a transcriptional target of p53, is downregulated in colorectal cancers associated with poor prognosis. *Mol Cancer Res* **11**, 1554–1563, <https://doi.org/10.1158/1541-7786.MCR-13-0330-T> (2013).
- Yao, H. B. *et al.* AKR1B10, a good prognostic indicator in gastric cancer. *European journal of surgical oncology: the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology* **40**, 318–324, <https://doi.org/10.1016/j.ejso.2013.12.014> (2014).
- Roessler, S. *et al.* A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. *Cancer Res* **70**, 10202–10212, <https://doi.org/10.1158/0008-5472.CAN-10-2607> (2010).
- Chen, X. *et al.* Gene expression patterns in human liver cancers. *Mol Biol Cell* **13**, 1929–1939, <https://doi.org/10.1091/mbc.02-02-0023> (2002).
- Guichard, C. *et al.* Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat Genet* **44**, 694–698, <https://doi.org/10.1038/ng.2256> (2012).
- Kim, J., Yao, F., Xiao, Z., Sun, Y. & Ma, L. MicroRNAs and metastasis: small RNAs play big roles. *Cancer metastasis reviews*, <https://doi.org/10.1007/s10555-017-9712-y> (2017).
- Ma, H., Liu, B., Wang, S. & Liu, J. MicroRNA-383 is a tumor suppressor in human lung cancer by targeting endothelial PAS domain-containing protein 1. *Cell biochemistry and function* **34**, 613–619, <https://doi.org/10.1002/cbf.3237> (2016).
- Azarbarzin, S., Feizi, M. A. H., Safaralizadeh, R., Kazemzadeh, M. & Fateh, A. The Value of MiR-383, an Intronic MiRNA, as a Diagnostic and Prognostic Biomarker in Intestinal-Type Gastric Cancer. *Biochemical genetics* **55**, 244–252, <https://doi.org/10.1007/s10528-017-9793-x> (2017).
- Chen, L. *et al.* miR-383 inhibits hepatocellular carcinoma cell proliferation via targeting APRIL. *Tumour biology: the journal of the International Society for Oncodevelopmental Biology and Medicine* **37**, 2497–2507, <https://doi.org/10.1007/s13277-015-4071-1> (2016).
- Fang, Z. *et al.* The miR-383-LDHA axis regulates cell proliferation, invasion and glycolysis in hepatocellular cancer. *Iranian journal of basic medical sciences* **20**, 187–192, <https://doi.org/10.22038/ijbms.2017.8246> (2017).

32. Shi, Y. & Huang, A. Effects of sorafenib on lung metastasis in rats with hepatocellular carcinoma: the role of microRNAs. *Tumour biology: the journal of the International Society for Oncodevelopmental Biology and Medicine* **36**, 8455–8463, <https://doi.org/10.1007/s13277-015-3565-1> (2015).
33. Le, T. X. A. T. miRBaseConverter: A comprehensive and high-efficiency tool for converting and retrieving the information of miRNAs in different miRBase. *R package version 1.0.4* (2017).
34. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).
35. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences* **98**, 5116–5121 (2001).
36. Huang, W. C. *et al.* BCRP/ABCG2 Inhibition Sensitizes Hepatocellular Carcinoma Cells to Sorafenib. *PLoS One* **8**, e83627, <https://doi.org/10.1371/journal.pone.0083627> (2013).
37. Feng, R. *et al.* miR-126 functions as a tumour suppressor in human gastric cancer. *Cancer Lett* **298**, 50–63, <https://doi.org/10.1016/j.canlet.2010.06.004> (2010).
38. Wang, J. *et al.* CRKL promotes cell proliferation in gastric cancer and is negatively regulated by miR-126. *Chemico-biological interactions* **206**, 230–238, <https://doi.org/10.1016/j.cbi.2013.09.003> (2013).

Acknowledgements

The authors appreciate Shen Chen, Di Ma, Ye Lu, Xiaoyong Gong, Jiajun Ren and Yuchen Yang for providing valuable technical supports and assistance. This study was supported by grants from the following: National Natural Science Foundation of China (Nos 81602544; 81372187).

Author Contributions

J.W. wrote the main manuscript text. Y.Z. conducted the statistics analysis and participated in the manuscript writing. X.F. directed pathological detection and prepared Figures. X.C. and Y.C. directed and supported the study. All authors have reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-29271-3>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018