

Integrated regulatory-metabolic network model reveals critical mechanism and potential targets for Hepatocellular Carcinoma

Yi-Zhou Xu

Shanghai Jiao Tong University <https://orcid.org/0000-0003-0158-0345>

Hang Zhang

Shanghai Jiao Tong University

Qiang-Zhen Yang

Shanghai Jiao Tong University

Ren-Liang Sun

Shanghai Jiao Tong University

Ke Wang

Shanghai Jiao Tong University

Yong-Yong Shi

Shanghai Jiao Tong University

Zhuo Wang (✉ zhuowang@sjtu.edu.cn)

<https://orcid.org/0000-0002-6382-1717>

Technical advance

Keywords: Regulatory-metabolic integration, Metabolic model, Stratification, Therapeutic target, Hepatocellular carcinoma, Survival analysis, Metabolic reprogramming, Flux Balance Analysis

Posted Date: April 8th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-21615/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

HCC (Hepatocellular carcinoma), the predominant form of liver cancer, has long been the top three leading cause of death in cancer worldwide. Although researchers have spent lot of effort to identify molecular targets available for treatment, the high tumor heterogeneity makes it difficult to develop effective therapy options and the drug response remains low. Under this circumstance, the precise stratification strategies are more than required. However, previous researches generally focused on single biological level, such as genome, transcriptome, or proteome, and are not able to discover effective therapeutic targets, so the systematic study of both regulation and metabolism of HCC is needed.

Methods

In this paper, we use two different algorithms to reconstruct regulatory networks for both HCC and normal liver cells, then integrate them with corresponding metabolic models in order to discover TFs (transcriptional factors) affecting tumorigenesis. Furthermore, a machine learning algorithm is utilized to classify HCC samples, differentially expressed genes, altered metabolic reactions and biological pathways are identified in lowest overall survival (OS) rate sub-type compared to others.

Results

We classify TCGA-LIHC samples into three sub-types with significantly different OS rate, and this stratification strategy is validated in another independent dataset LIRI-JP. Then, we identify 5 key TFs affecting cancer cell growth and CREB3L3 is believed to be associated with poor prognosis. The comprehensive metabolic analysis on personalized metabolic models highlight 18 metabolic genes essential for tumorigenesis in all three sub-types of patients, besides, ACADSB and CMPK1 are highly possible to be strongly correlated with lower OS.

Conclusions

Among 20 metabolic genes identified through metabolic analysis, 15 of them have already been targeted by approved drugs according to DrugBank. In addition, miRNAs targeting key TFs and genes are also involved in well-known cancer related pathways. The multi-scale regulatory-metabolic model reveals the critical mechanism of HCC cell proliferation and suggests potential targets.

Background

HCC is the most common type of primary liver cancers, it has become the third leading cause of deaths from cancers not only in developing countries but also in developed countries(1). Obesity, diabetes, fatty livers, virus infection and many other diseases can all turn into HCC, and because of the current difficulty in early-diagnosis, the treatment of HCC is unsatisfactory. Although, drugs like Sorafenib and Lenvatinib had been approved by FDA, the drug-response rates were relatively low due to high tumor heterogeneity.

So the precise stratification and effective targets discovery are required for treatment. Recently, several studies reported the subtype classification of HCC based on genetic omics data. For instance, the early-stage Chinese HCC samples have been classified using proteomic data and successfully uncovered the mechanism of early tumor cell development (2). The proteomic of HBV-related HCC samples identified three subgroups with distinct features in metabolic reprogramming, microenvironment dysregulation, and cell proliferation(3).

It has been well recognized that metabolic reprogramming is an important characteristic and driver of cancer. Genome-scale metabolic models (GEMs) have been successfully used to characterize cancer metabolism as well as identify drug targets for cancer treatment, which is a powerful framework to mechanistically represent the relationship between genotype and phenotype by computationally modeling the biochemical constraints that are imposed on the phenotype. They are capable of predicting various biological tasks under given circumstances (4, 5) thus providing the basis for identifying essential genes or reactions for a particular objective function. By comparing the alteration of metabolism between normal and tumor tissue models, many disease-related genes/proteins/metabolites have been identified and then experimentally validated. For example, Folger *et al.*(6) used microarray data to leverage key genes for non-small-cell lung cancer; Agren *et al.*(7) utilized data from HPA Database with INIT algorithm to successfully construct 69 cell-specific models and 16 cancer-specific models. Uhlen *et al.*(8) employed RNA-Seq data from TCGA Database together with tINIT algorithm to reconstruct 6753 patient-specific metabolic models for various cancers. Several anti-cancer drugs designed based on molecular targeting strategy have already been approved by FDA and put into clinical use: Avastin for EGFR-mutation induced non-small cell lung cancer; Pertuzumab for HER2-positive type breast cancer, etc.

In comparison, there has been few breakthroughs in identifying effective therapeutic targets for HCC. The recent study by Bidkhori *et al.* took a new beginning for HCC metabolism study, which utilized metabolic network topology analysis to divide 179 TCGA-LIHC samples into three subtypes and identify potential subtype-specific therapeutic targets (9). However, tumor development is far more complex which requires understanding of the complicated multi-level mechanisms including both regulation and metabolism. It is well known that metabolism can be affected by transcriptional regulation and vice versa, therefore the integration of regulation and metabolism is very important for cross-talking study, and may be useful in precise stratification. So far, there has been few regulatory networks constructed for liver tissue, let alone ones for HCC tissue, more than this, the combination of regulation and metabolism in studying HCC is totally empty. In this study, we leveraged the application of integrated regulatory-metabolic model in disease study to investigate the possible mechanism of HCC using the whole TCGA-LIHC samples. Firstly, we constructed the regulatory networks of HCC and normal liver tissues, then combined with human liver metabolic model using the state-of-art algorithm. We classified HCC patients into different subgroups by expression data of TFs and genes in the integrated model, and evaluated classification results by overall survival outcomes. Using the integrated regulatory-metabolic model, we identified mechanism of HCC tumor cell progression, as well as genes associated with poor prognosis and the potential therapeutic targets, the results showed high consistency with previous *in-silico* and experimental studies. In addition, we analyzed the miRNAs information to further support that the genes identified by the integrated model

are highly important for HCC tumorigenesis and may be the targets for clinical treatment. The main schema of this study was shown in Fig. 1.

Methods

Expression Data

RNA-Seq expression data were obtained from three sources: 315 HCC samples with clinical outcomes from TCGA-LIHC Project; 232 HCC samples with clinical outcomes from ICGC-LIRI Project; 50 HCC paired tumor-normal samples from Liu et al. (10) (GSE77314). GRCh37/hg19 was used as the reference genome and all data have already been pre-processed to quantitate from reads to FPKM through cufflinks.

The GSE77314 dataset was also used to infer tumor and normal liver regulatory networks, and all of the three datasets were employed in model integration.

Metabolic Models

The genome-scale metabolic liver model used for integration was from HMA Database (<https://metabolicatlas.org/gems/repository>). It was built based on the combination of HMR2(7) model with RNA-Seq data of liver tissue to provide approach to explore metabolic and proteomic functions in cancer(11). According to the authors, high-energy compounds cannot be generated from low-energy ones to avoid futile cycle in models and this property is accessed by adding certain constraints.

The patient-specific genome-scale metabolic models used for metabolic analyses are retrieved from Liver Hepatocellular Carcinoma models section in BioModels Database (<https://www.ebi.ac.uk/biomodels/>). Uhlen et al. employed tINIT to do the reconstruction, and the characteristic of metabolic pathways in each model is determined by the protein abundance detected from individual patient RNA-Seq data in TCGA-LIHC Project. Biomass representing cell growth (whose formula is also obtained from Uhlen et al.) is set to be the objective function. We selected 315 out of 338 HCC individual models with clear clinical stage information (exclude 'not reported') for metabolic reprogramming analysis.

Construction of regulatory network

We used two independent algorithms MERLIN (Modular regulatory network learning with per gene information) and CMIP (Conditional mutual information measurement using a parallel computing framework) to construct the tumor/normal regulatory networks from the expression data (GSE77314). MERLIN, which is designed by Roy S. combines both per-gene method and per-module concept based on probabilistic graphical model to infer regulatory network, so it can not only include memberships deduced from individual gene but it can also take the similarity within a group of genes into consideration. The algorithm has been proved to be effective in predicting transcriptional changes in human differentiation neural progenitor cells. In addition, MERLIN outperformed several other state-of-the-art algorithms. We used default settings except modifying folds of cross-validation to five in our study.

CMIP is proposed by Chen *et al.* The interactions between genes are measured by the conditional mutual information measurement to avoid neglecting subtle relations under certain conditions, for example, both A and B are strongly connected to C, then the actual relation between A and B may be confusing because of the interference of C. The performance is evaluated by average AUC and AUPOR of ten benchmark datasets provided by DREAM3, and the results showed that CMIP performs better than other algorithms, at the meantime, parallel computation strategies enable it to handle genome-scale datasets and to complete tasks within a relative short time period compared to other popular methods presented in DREAM Projects(12). We run CMIP in default parameters to let the algorithm automatically decide the threshold dynamically.

The regulatory relations deduced by both of the two algorithms were regarded as 'high confidence' regulations and tagged in the regulatory networks for further integration with metabolic model.

Integration of regulatory network and metabolic model

IDREAM (Integrated Deduced Regulation And Metabolism) algorithm proposed by our group was used to do the integration, which uses bootstrapping regulatory relations delivered by EGRIN (Environment and Gene Regulatory Influence Network) (13) and introduce PROM (Probabilistic Regulation of Metabolism)-like framework to add metabolic constraints. IDREAM has been successfully applied in *Saccharomyces cerevisiae* to predict growth phenotype of gene mutants and shows high consistence with experiment data. In addition, it has uncovered novel synthetic lethal pairs of TFs and metabolic genes with important interaction mechanism.

We collected the liver regulatory network from RegulatoryCircuits published by Marbach *et al.* (14) (<http://regulatorycircuits.org/>) and human regulatory network from RegNetwork published by Liu *et al.* (15)(<http://www.regnetworkweb.org/>). Marbach *et al.* inferred regulatory relations based on the interaction of TFs-promoters, TFs-enhancers, promoters-genes, enhancers-genes and introduced ChIP-Seq, eQTL, RNA-Seq data to do validation. Liu *et al.* constructed regulatory network by considering prior knowledge of transcription factor binding sites and post-transcriptional regulatory by miRNAs, besides, convincing results from literatures were also included. There are 1366 TFs in the union of these two regulatory networks. There are 2456 metabolic genes contained in liver tissue model in HMA Database mentioned above; Then we used IDREAM to build the integrated regulatory-metabolic model to investigate TFs affecting cell growth in tumor and normal respectively.

In total, 3492 genes are determined (1366 + 2456 but exclude the overlapped genes and ones that are unavailable in expression data)

Metabolic Analysis

COBRA Toolbox incorporated in MATLAB was used for metabolism analysis, we replaced default solver with Gurobi to get better results. It focuses on presenting feasible phenotypic states through setting appropriate constraints gained from prior knowledge or assigned conditions. And by identifying the metabolic task to be studied, a reduced set of solutions can be achieved for further analysis.

Our research mainly used SingleGeneDeletion function to find metabolic genes whose knockout will lead to the decrease of cell growth; and OptimizeCbModel function to find the optimal growth rate, and corresponding flux distribution.

DAVID (<https://david.ncifcrf.gov/>) was employed to do the enrichment analysis, we followed the default settings provided, chose proper thresholds and selected the results comprising of KEGG pathways.

Stratification, survival and DEG analysis

The 3492 genes filtered above were used to stratify 315 TCGA-LIHC samples using the NMF(non-negative matrix factorization) consensus cluster method built in R packages called 'NMF'. It is a machine learning algorithm aiming to distinguish different molecular patterns in high-throughput genomic data. We used 200 iterations to determine the best clustering number between 2 and 10, and the result was set to 3 according to the principal of cophenetic score and average silhouette width. Then 300 iterations were performed to identify the group members.

We used clinical outcomes to evaluate the clustering results, Kaplan-Meier survival curve implemented in 'survival' packages in R was employed to assess the overall survival rate. Three classes showed significant differences in survival outcomes

For the analysis of DEG (differently expressed genes), we used a linear model and moderated t-statistics based algorithm implemented in the 'Limma' package, with $\log_2(\text{FC}) \geq 1$ and $P\text{-value} \leq 0.05$. We operated this procedure between three classes, respectively and the intersection were taken to be the final gene set.

Network topology Analysis

Cytoscape software was used to do the topology exploration. The 'Tools'-'Merge'-'networks' function with the optional parameter 'difference' was used to find the difference between tumor and normal liver network. The principle was to remove all identical nodes so to identify TFs/metabolic genes only exist in HCC/normal regulatory network. And then we highlighted the hub genes being responsible for the abnormality on topological structure. For the multi-scale network visualization, we used 'Tools'-'Merge'-'networks'function with optional parameter'union' to integrate miRNA-genes network and TFs-target genes network to present a comprehensive view of the core regulatory-metabolic network.

Results And Discussion

Regulatory network difference between tumor and normal liver cells

Due to the indispensability between metabolic and regulatory mechanisms, it is reasonable to take regulatory network into consideration when studying cancer metabolism. There are many algorithms designed to infer these networks from sequencing profiles and the results have been validated in model

organisms like *S.cerevisiae* and *E.coli*. We employed MERLIN and CMIP algorithms (Methods) together with paired RNA-Seq data obtained from GEO database (GSE77314)(10) to construct the regulatory networks of HCC and paired normal tissues. There are in total 15143 pairs and 29127 pairs of regulation between transcriptional factors and target genes deduced from tumor and normal samples (Supplementary Table 1). Among them, 1654 pairs are same. Cytoscape was used to visualize the topology difference (Methods) of these two networks, after deducting the nodes with little influence, the most cored structure is shown in Fig. 2.

NME2 and *NFKBIA* are clearly the hub TFs playing important roles in normal liver models. In contrast, these two TFs are absolutely absent in HCC tumor model. It is well-known that nuclear factor κ B (NF- κ B) affects multiple biological process by regulating immune response and inflammation, so it has been recognized as hallmark in cancer progression(16). *NFKBIA* is a member of a cellular protein family which can mask the nuclear localization signals of NF- κ B and block its way to bind DNA. Because of this inhibition ability, it has long been considered as a tumor suppressor(17). Also, *NME* is a gene family associated with suppressing cancer metastasis and invasion(18). Especially, *NME2*, which is located on 17q21, its products have been reported to inhibit the metastasis of breast cancer and lung cancer (19, 20). Besides, there are experiments supporting this idea by showing that the RNA expression of *NFKBIA* and *NME2* are significantly decreased in HCC tissues compared to normal tissues. Therefore, the reconstructed regulatory networks effectively revealed the critical differences between liver cancer and normal tissues. It is persuasive to consider *NME2* and *NFKBIA* as tumor suppressor factors for future study.

Integrative regulatory-metabolic network identified abnormality of Hippo signaling as key misregulation in HCC

In our previous work, we developed a framework called Integrated Deduced Regulation And Metabolism (IDREAM) (21), which uses bootstrapping linear regression model on large-scale gene expression dataset to predict transcription factor (TF) regulation on enzyme-encoding genes, and then utilizes a probabilistic regulation of metabolism approach to apply regulatory constraints to the metabolic network. Then the integrated model can predict the influence of each TF knockout on certain objective function, such as cell growth. It has been successfully applied in *Saccharomyces cerevisiae* to effectively predict the influence of transcriptional regulation on metabolic phenotype. Here we used IDREAM algorithm (Methods) to integrate regulatory network with metabolic models to identify potential transcription factors vital to the growth of HCC cells.

The composition of integrated model for HCC and normal liver tissue are listed in Supplementary Table 1. The basic metabolism part is consistent, while the transcriptional factors and target genes are different. There are 1313 TFs in HCC model and 1312 TFs in normal model, among them, 33 and 32 TFs are unique, respectively. To analyze the 32 specific TFs (including *NME2* and *NFKBIA* mentioned above) not involved in HCC regulatory network, the Reactome database show that they were enriched for YAP1- and WWTR1 (TAZ)-stimulated gene expression pathway. They are transcriptional co-activators interacting

with TEAD family genes to promote expression of TFs critical to cell proliferation and apoptosis through Hippo signaling pathway, which means the depletion of these 32 TFs will lead to the abnormality of Hippo signaling and may induce a wide range of cancers. In addition, the 33 specific TFs in HCC integrated model were mainly enriched in pathways in cancer metabolism and transcriptional misregulation. Both conclusions supported that the integrated models we reconstructed are credible.

For each TF knockout, we changed the constraints on corresponding reactions according to activation/inhibition interactions and then simulated the cell growth rate to get the ratio over wild type. We found TFs affecting both tumor and normal cell growth, and TFs only reduce growth of tumor cell particularly (Supplementary Table 2), for example, *ETV7*, whose knockout will lead to nearly 2/3 reduction in tumor cell growth; *CREB3L3*, which seems to be closely correlated with poor prognosis in HCC.

Precise stratification of TCGA-LIHC samples based on metabolic and transcriptional gene expression

Although researchers have long been working on identifying genes or pathways available for treatment, the high heterogeneity of HCC has always been an obstacle in the precise clinical diagnosis, it is partly due to the fact that the current TNM stage classification is not that efficient. Therefore, researchers are devoted to try to stratifying HCC patients on molecular level and hope to find corresponding therapeutic targets. Bidkhor *et al.* utilized metabolic network-based method to divide 179 TCGA-LIHC samples into three subtypes and identified their specific characteristics; Jiang *et al.* used proteomic data to classify HCC patients and explored the mechanism of early-stage HCC tumor cell. Inspired by them, here we used 315 HCC samples attached with actual clinical stage information from TCGA-LIHC dataset and 3492 genes (Methods) obtained from multiple sources to identify altered metabolism among different subgroups and specific characteristics of poor prognosis subgroup.

Using a non-negative matrix factorization consensus-clustering analysis, three major classes are concluded in TCGA-LIHC cohort, 130, 127 and 58 cases are classified into Class1, Class2 and Class3 respectively (Methods). The survival curves are shown in Fig. 3a, from which we can see that Class2 has a significantly lower overall rate of survival. By collecting 159 overlapped samples used in Bidkhor *et al.*'s study and ours, we reached a relatively good agreement in identifying the lowest OS subgroup: 90.57% (48 of 53, their results that are also in ours) and 70% (49 of 70, our results that are also in theirs). Consequently, we focused on digging out the characteristics of poor prognosis subgroup Class2 at both transcriptome and metabolism level.

A supervised analysis through LIMMA(22) revealed Class2-specific gene expression patterns, comprising 287 up-regulated (including three potential therapeutic targets: *ALDOA*, *G6PD*, *ACSS1* specific to lowest OS subgroup identified by Bidkhor *et al.*) and 112 down-regulated genes (Supplementary Table 3) (Methods), which are enriched in 17 and 13 non-overlapped KEGG pathways respectively, as shown in Fig. 3c.

In order to validate the effectiveness of our stratification strategy, we applied the same strategy on LIRI-JP data (ICGC database) to split it into three subgroups, as we can see in Fig. 4a, the P-value of survival curves are still significant. By analyzing the pathways enriched for differentially expressed genes, the result shows high consistency with the ones of TCGA-LIHC data: the up-regulated genes are mainly enriched in well-known cancer related pathways in improving cell proliferation, such as increased glucose uptake as a principal nutrient source in central carbon metabolism of cancer, cell cycle and fructose metabolism; HIF signaling, which consists of master regulators of oxygen homestasis that not only allow tumor cells to adapt to hypoxic environment by enhancing oxygen delivery but also affect important growth factors like *VEGF*; let alone viral carcinogenesis and hepatitis B pathways which can be directly link with HCC development. In the meantime, the down-regulated genes are generally found in pathways contributing to drug metabolism, for example, the PPAR signaling pathway, which has also been identified in less aggressive HCC subtypes through proteomics analysis; drug cytochrome P450 metabolism, which is indeed reduced in advanced cancer patients (23).

Integrated models revealed that PI3K-Akt and mTOR signaling pathways are critical to HCC tumor cell growth

By using IDREAM algorithm, we found 8, 13, 5 TFs being vital for HCC cell growth respectively in Class1, Class2 and Class3 of TCGA-LIHC (after excluding TFs that also affect normal tissue). There are 3 TFs common in all three classes, as listed in Table 1. The knockout of *ETV7* shows the greatest decrease in growth rate in all three classes. *ETV7* is a transcription factor belonging to ETS family, which has been proved to be responsible for the development of different tissues as well as the progression of several cancers, such as HCC(24, 25). Due to its translocation function, the overexpression of *ETV7* has been associated with tumorigenic transformation and anti-apoptosis by blocking Mys-induced apoptosis pathway(26–28). Furthermore, there are growing experimental evidences showing that *ETV7* also plays a significant role in mTOR signaling pathway by assembling mTOR3 complex, which can stimulate cell proliferation and is not sensitive to rapamycin, a common anti-tumor agent, unlike mTOR1/2(29). So the depletion of *ETV7* may cause the inactivation of mTOR3, thus leading to tumor cell death after treatment.

Table 1
Cell-growth ratio by influential TFs knockouts

Ratio after knockout of common TFs in all 3 classes of TCGA-LIHC			
TF	Class1	Class2	Class3
CTBP1	0.926	0.926	0.926
HTATIP2	0.926	0.926	0.926
ETV7	0.234	0.12	0.09
Ratio after knockout of specific TFs in lowest survival class			
TF	TCGA-LIHC	LIRI-JP	
NR1I3	0.978	0.978	
HNF4A	0.969	0.984	
RORC	0.935	0.888	
F2	0.975	0.967	
CREB3L3	0.856	0.876	

In addition, *CTBP1* is a well-known cancer hallmark taking responsibility in pro-tumorigenic process as well as affecting regulatory network(30). It can bind to C-terminus of the adenovirus protein E1A to promote cell-proliferation and invasion(31). Besides, the characteristic of cancer cells (high NADH level) makes it possible for *CTBP1* to bind to NADP with a high affinity thus triggering a conformational change, leading to hyper-activity in both tumorigenesis and tumor progression.

To explore characteristics of TFs leading to low survival rate and poor prognosis, we selected TFs whose knockout only influence samples in Class2 but not in Class1 or Class3 of TCGA-LIHC dataset. There are 5 TFs specific for Class2 (threshold: ratio < 0.98), in addition, 4 of 5 these TFs are also in the result of the lowest survival subgroup (Class3) of LIRI-JP dataset (the ratio of HNF4A exceeded the threshold a little), as shown in Table 1.

The knockout of *CREB3L3*, which also appears in the decreases the growth rate of tumor cells by over 15%, but have no effect on normal tissues. *CREB3L3* is reported to be activated in the lipid metabolism in liver-specific tissue in a mutual manner with PPAR α (32). They both play important roles in the utilization of fatty acid for energy under fasted state, just like that in cell proliferation, so it is not surprising to see its absence will result in the decreased growth rate in tumor cells. The expression of *CREB3L3* is link with anti-apoptosis, cell survival and HBV-associated HCC development by regulating hepatic genes in PI3K-Akt and AMPK signaling pathways. With the combination of consistent *in-silico* analyses and biological knowledge, we suggested *CREB3L3* as a therapeutic target, especially for those HCC patients in advanced stage.

Metabolic genes in cholesterol biosynthesis are druggable targets in HCC treatment

We incorporated patient-specific models established by Uhlen *et al.* to do metabolic analyses, including identification of metabolic genes essential for tumor cell growth and annotation of the specific reactions altered during tumor development. All 315 metabolic models were built at the aim of representing tumor growth. On the basis of genetic human metabolic model HMR2 and RNA-Seq expression data from TCGA-LIHC, a task-driven model reconstruction algorithm called tINIT was employed to construct all models.

We conducted the single gene deletion simulation by a function provided in COBRA Toolbox (33). The total gene number of each model ranges from 1106 to 2169. We first differentiate the essential genes respectively in the three subtypes of TCGA-LIHC samples by the NMF stratification strategy mentioned above. And then collected genes that existed in at least half of the samples in each class. After wiping out those having no influence on tumor cell growth, there are 19, 20, and 18 genes left, respectively. The 18 genes found in Class3 (relatively high overall survival rate) are same with other two classes, and one other gene, *ACADSB* was shared by Class1 and Class2, while the other gene *CMPK1* was identified in Class2 alone. We checkout the total 20 influential genes to see whether there is prior knowledge about their therapeutic potentiality in Drugbank, as shown in Table 2.

Table 2

Lethal metabolic genes as potential targets and corresponding drugs in DrugBank

Lethal Gene	Target Drug	Drug state	Brief Description of drug
IDI1	Dimethylallyl Diphosphate	Experimental	
SQLE	Ellagic Acid	investigational	Antioxidant and anti-proliferative/anti-cancer effects
FDFT1	TAK-475	Investigational	Target rate-limiting enzyme in the hepatic biosynthesis of cholesterol
CRAT	Levocarnitine	Approved	Treatment of primary systemic carnitine deficiency
EBP	Tamoxifen	Approved	Treatment of metastatic breast cancer
ACADSB	Isoleucine Valproic Acid	Approved Approved	Anti-proliferative effects useful in cancer therapy
SLC22A5	Levocarnitine	Approved	Treatment of primary systemic carnitine deficiency, affect fatty-acid synthesis
HMGCR	Lovastatin Cerivastatin Simvastatin Atorvastatin Rosuvastatin Meglutol	Approved Approved Approved Approved Approved Experimental	Lowering LDL cholesterol and triglycerides, hypercholesterolemia; Target rate-limiting enzyme in the hepatic biosynthesis of cholesterol
CMPK1	Gemcitabine Lamivudine Sofosbuvir	Approved Approved Approved	Various advanced cancers Treatment of HBV Treatment of HCV Reduce incidence of HCC
MVK	Farnesyl thiopyrophosphate	Experimental	
HSD17B7	NADH	Approved	Treating Parkinson's disease, chronic fatigue syndrome, Alzheimer's disease and cardiovascular disease
NSDHL	NADH	Approved	Treating Parkinson's disease, chronic fatigue syndrome, Alzheimer's disease and cardiovascular disease
DHCR7	NADH	Approved	Treating Parkinson's disease, chronic fatigue syndrome, Alzheimer's disease and cardiovascular disease
ACACB	Soraphen A	Experimental	Anti HCV viral activity
LSS	R048-8071 Lanosterol	Experimental Experimental	

Lethal Gene	Target Drug	Drug state	Brief Description of drug
FDPS			
PMVK			
MVD			
CYP51A1			
SC5D			

According to Drugbank, 9 genes (*CRAT, EBP, ACADSB, CMPK1, SLC22A5, HMGCR, HSD17B7, NSDHL, DHCR7*) have already been targeted by approved drugs in the treatment of cancer or relative diseases, other 6 genes have corresponding drugs under experimental or investigational state. Both *CMPK1* and *ACADSB* seem to be vital to tumor cell growth in HCC models with lower survival rate, and they actually have long been regarded as prognosis biomarkers associated with worse survival in multiple tumor(34–39). In fact, *CMPK1* has also been targeted by three FDA approved cancer drugs in clinical usage, especially in the treatment of diseases induced by virus infection, such as HCC caused by HBV/HCV: Gemcitabine, Lamivudine and Sofosbuvir. Li *et al.*(2) recently reported that in Kaposi’s sarcoma, a common AIDS-related malignancy caused by infection of KSHV, cells can increase invasiveness and motility by over-expression of *CMPK*, meanwhile, this theory has been validated by the knockout experiments carried out in cell line.

Among 18 genes lethal in all three classes, there are 15 genes participating in cholesterol biosynthesis via desmosterol (DESMOL pathway), which is the dominant form of liver cholesterol biosynthesis. (40): Firstly, *HMGCR, MVK, PMVK, MVD* and *ID1* are genes involving in mevalonate pathway to convert acetyl-CoA into dimethylallyl pyrophosphate (DMAPP), and enzyme encoded by *FDPS* helps DMAPP synthesize farnesyl pyrophosphate (FAPP), then *FDFT1* catalyzes the dimerization of two FAPP into squalene (SQNE). In the next step, *SQLE* and *LSS* play the important rate-limiting roles in cholesterol biosynthesis by catalyzing SQNE to lanosterol (LNSOL). After that, LNSOL goes through demethylation, oxidation, reduction catalyzed by *CYP51A1, NSDHL, HSD17B7* to become zymosterol (ZYMOL), the precursor in DESMOL pathway, in which *EBP, SC5D, DHCR7* are taking responsibilities by order to convert ZYMOL into DESMOL. Finally, DESMOL is reduced by *DHCR24* to produce cholesterol. So the knockout of any gene will cause the break of cholesterol biosynthesis and lead to the depletion of cholesterol, which is disastrous for tumor cell growth.

There are only five of the predicted essential genes having not been recorded in Drugbank, but the high confidence of existing drug targets hitting suggested that those five metabolic genes are possibly be potential targets and worthy to be explored in depth for future researches.

Enhancement of glutathione and fatty-acid biosynthesis is important metabolic reprogramming associated with poor

prognosis

It is widely accepted that tumor cells reprogram certain metabolic pathways consisting of numerous reactions to meet the demand of unlimited cell proliferation, aggressive invasiveness and anti-apoptosis. We investigated 1329 reactions existing in all 315 patient-specific models to see the flux patterns and enzymes differed between the poor-prognosis subgroup (Class2) and the other two classes. We conducted the flux balance analysis with cell growth as objective function to get flux distribution for each patient, and then selected candidate reactions having similar flux change in over half samples of each subgroup. There are four types of flux alteration patterns: 1) from negative flux value in Class1 and Class3 to positive flux in Class2; 2) from positive flux value in Class1 and Class3 to negative flux in Class2; 3) from non-zero flux in Class1 and Class3 to zero flux in Class 2; 4) from zero flux in Class1 and Class3 to non-zero flux in Class 2. The altered reactions, formula, enzymes and corresponding types of flux patterns are shown in Supplementary Table 4.

There are two reactions having flux change of type1 and type2, respectively. According to these four reactions, GSH (glutathione) production is suspected to increase in Class2 samples due to the enhancement of AKG biosynthesis and cysteine accumulation in cytosol. GSH is a key member of cell immune response system, lacking it can easily lead to cell death, several labs have confirmed its common occurrence in all cancers(41) and considered it as potential therapeutic target. Besides, the loss of *SLC25A11*, the enzyme catalyzing these reactions, has been validated to inhibit tumor cell growth in non-small cell lung cancer (42). Baulies *et al.* (43) suggested that the over-expression of *SLC25A11* works as an adaptive mechanism of HCC to provide enough GSH for vast cell growth, meanwhile, *SLC25A11* induced exportation of AKG to cytosol also activates mTOR pathway to promote cell growth and anabolism through EGLNs (egl-9 family hypoxia-inducible factors)(44).

There are eleven reactions carried no flux in Class 2 but having positive flux in Class 1/3 (type 3), four of them belongs to porphyrin metabolism. The enzyme *UROD* involved in this pathway is a recently identified potential anti-cancer target due to its ability to convert uroporphyrinogen to coproporphyrinogen(45). Another enzyme *ALAD* has been found over-expressed in breast cancer patients with a favorable clinical outcome for its up-regulation can suppress cell proliferation and invasion(46). In addition, a set of enzymes responsible for carnitine shuttle including *SLC22A1*, *SLC25A20*, *SLC25A29* and *CPT2* are also down-regulated in HCC tumor cells, they mainly play rate-limiting roles in controlling fatty acid oxidation(47). Their low-expression are proved to be significantly associated with worse patient survival(48) and differentiation state by impairing NO production and mTOR signaling pathway mediated by arginine, in some situation, even leading to severe autophagy(49, 50).

There are three reactions having non-zero flux in Class 2 but zero flux in Class 1/3 (type 4), involving fatty acid activation responsible for providing adequate ATP and CoA for tumor cell growth; glycine, serine, threonine metabolism helping reduce ROS pressure through SGOC metabolic network during tumor metastasis(51), and arginine/proline metabolism, which can regulate response to nutrient and oxygen deprivation in oncogenesis thus avoiding cell apoptosis(52). Furthermore, the exploration of enzymes

revealed that *ACADSB* (which is also indicated by previous analyses), *ACSL3* and *ACSL4* regulate distinct proteins like p-AKT, LSD1 and β -catenin to stimulate tumor cell proliferation(53).

To conclude, the altered reactions specific to Class2 samples generally aim at promoting tumor cell growth and decreasing their sensitivity towards normal apoptosis signals. Several key enzymes have already been regarded as biomarkers in cancers, with deeper exploration, it is convincing to figure out their detailed mechanism and find more available therapeutic targets. Meanwhile, the metabolic reprogramming accounting for poor prognosis also supports that our stratification of the HCC patients are meaningful.

miRNAs regulating the influential genes for HCC cell proliferation

To disclose a more comprehensive interplay between regulation and metabolism of HCC, we further retrieved microRNAs regulating the influential genes highlighted in our previous analyses: including 3 common TFs showing influence in all three Classes; 5 overlapped TFs specifically affecting the lowest survival subgroup of TCGA-LIHC(Class2) and LIRI-JP(Class3); and 20 metabolic genes revealed by single gene deletion result. From MIRNET, we found six highly-connected microRNAs, three of which have been validated to be directly linked with HCC, including miR-124-3p, miR-1-3p, and miR-24-3p (Supplementary Table 5).

MiR-124-3p and miR-1-3p have been reported to be down-regulated in HCC patients compared to normal(54, 55). While miR-24-3p has been considered to be involved in HCC diagnosis panel for its abnormal over-expression.

The specific mechanism of how the loss-function or gain-function of these miRNAs contribute to tumorigenesis remains unclear, however there are some experiment-based theories available. Especially, miRNA-124-3p, which also correlates with a wide range of other cancers including breast cancer, lung cancer, colorectal cancer, etc. appears to be the key microRNA during oncogenesis(56–58). Zheng *et al.* (59) believes that miR-124-3p participates in reducing tumor cell motility and invasion by controlling epithelial–mesenchymal cell transition as well as cytoskeletal events through cpG-island methylation(60). As for miR-1-3p, Zhang *et al.*(61) suggests that its over-expression can inhibit cell proliferation and induce apoptosis by targeting PI3K-Akt and mTOR pathways through ETV7. While the down-expression of miR-24-3p can assist this process by deactivating Fas-receptor in NOTCH pathway and inhibiting *HNF4A* to drive a feedback loop to lead to cancer-related inflammatory reaction (62). Plus, studies by Wang *et al.*(63) and Chen *et al.*(64) indicates the regulatory impact of miR-24-3p not only on altering cell-cycle by inducing p53 mutation but also on avoiding cell death by targeting Fas-receptor in NOTCH pathway(65).

In addition, miR-26b-5p, which has the maximal connection (9) in regulating 28 highlighted genes, has been experimentally validated to under-express in HCC patients with worse prognosis. Because it can suppress tumor invasion as well as induce apoptosis by targeting *SMAD1*(66), which is consistent with

our conclusion about *SMAD* gene. Nevertheless, three genes obtained by our integrated regulatory-metabolic analysis in previous section, *CMPK1*, *ACADSB*, *RORC* are directly regulated by miR-26b-5p, and the fact that they are all specific genes for Class2 (the worst overall survival rate) proved our conclusion is credible. By combining the *in-silico* results in our work together with prior knowledge obtained from literatures, Fig. 5a shows the core network comprising of microRNAs, TFs and genes involved in HCC, which can provide basis for further studies in HCC development.

To take 28 candidate genes and 6 top-connected miRNAs into whole consideration, we conducted IPA (Ingenuity Pathway Analysis) to explore the biological connection among them. As shown in Fig. 5b, *EGFR* is inferred and densely linked with our core gene set. *EGFR* is one of the most crucial genes responsible for cancer cell growth, its over-expression can lead to unlimited cell proliferation, just like that in tumor cells, so it has long been focused and treated as therapeutic targets in cancer therapy. Multiple FDA-approved drugs such as Gefitinib, Lapatinib have been proved to be effective in EGFR-related non-small-cell lung cancer and several other cancers. So the IPA result indicated that our core-gene set concluded from integrated regulation-metabolism analyses are convincing and worthy further researches.

Conclusions

Integrated regulatory-metabolic network uncovered differences between HCC and normal liver cells

Due to the comprehensive information linking genes-proteins-reactions contained in genome-scale metabolic models, researchers have successfully identified numerous potential disease-related biomarkers by metabolic analyses. Meanwhile, because of the inseparability between metabolic and regulatory mechanism, the integration of transcriptional regulation with metabolism would allow us to better describe the impact of mutations and environmental perturbations on functional metabolism. This kind of integration has been proved to be effective in model organisms such as *S. cerevisiae* and *E. coli*, but has not yet been applied to study malignant disease in human, such as cancers.

Although there are a few regulatory networks of normal liver tissue, the regulatory networks of HCC is still in depletion, not to mention the integration of regulatory network with metabolic model. In this study, we used two different approaches CMIP and MERLIN to construct regulatory networks for HCC and normal liver tissue samples. Through topology analysis, *NME2* and *NFB1KA* are highlighted as tumor suppressor TFs because of their absence in tumor regulatory network and high connectivity in non-tumor one. Then we integrated the regulatory networks with human liver metabolic model, and compared the effects of TFs on cell growth in tumor and normal models, respectively. We found TFs only reduce growth of tumor cell particularly as potential targets. Further, by comparing the influential TFs in early stage and advanced stage samples, we found five TFs, including *SMAD2*, *HEY2*, *ELK1*, *CREB3L3*, *CCDC6* are vital to advanced stage HCC patients.

Three subtypes of HCC samples demonstrated significantly different prognosis

By splitting TCGA-LIHC samples using pre-filtered 3492 genes, we defined three patient subgroups which can be distinguished by overall survival rate and Class2 showed the worst survival. We identified 3 essential TFs for HCC tumor cell growth common in all three groups, among which *ETV7* showed the greatest impact, especially in Class2 by decreasing cell growth rate by about 88%. *ETV7* is a transcription factor belonging to ETS family, which has been long proved to be responsible for the progression of several cancers, including HCC. Because of its translocation function, the overexpression of *ETV7* has been associated with tumorigenic transformation and anti-apoptosis by blocking Mys-induced apoptosis pathway. There are growing experimental evidences showing that *ETV7* also plays a significant role in mTOR signaling pathway by assembling mTOR3 complex to stimulate cell proliferation and prevent cell from rapamycin, a common anti-tumor agent. In addition, the knockout of 5 TFs specifically affecting samples in Class2, we identified potential TFs related to poor prognosis. Among them, *CREB3L3*, which has also been identified as influential for advanced stage HCC samples by integrated model in previous analysis. The same stratification strategy has been applied on LIRI-JP dataset to validate the effectiveness: the survival outcomes shows significantly differences between subgroups and the result of TFs affecting the lowest survival subgroup shows high consistency with that of TCGA-LIHC dataset.

In addition, the poor prognosis group Class2 also exhibited specific pattern of altered metabolism. Flux alterations in Class2 samples showed the accumulation of both AKG and cysteine, which indicated the over production of GSH, a key member of cell immune response system in improving cell proliferation and avoiding apoptosis. Besides, the biosynthesis of fatty acids, mTOR signaling were also hyper-activated, and pathways like glycine, serine, threonine metabolism were employed to reduce ROS pressure during tumor homeostasis.

Key metabolic genes in cholesterol biosynthesis identified by patient-specific models are potential targets

The metabolic analyses based on patient-specific models found 20 metabolic genes playing important roles in HCC tumor cell growth by participating in cholesterol biosynthesis pathway, 15 of them have already been targeted in various cancers or cancer-related diseases according DRUGBANK, such as *ACADSB* and *SQLE*, which are targeted to suppress cancer cell-proliferation; *EBP*, which is one of the targets of Tamoxifen in breast cancer therapy; *CMPK1*, *ACACB*, *FDFT1*, *CRAT*, *SLC22A5*, *HMGCR* which play important roles in treatment of hepatic related disease to prevent HCC, providing strong proof for metabolic network-based target discovery in cancer. Therefore the other 5 candidate genes might also be potential targets. What's more, we further found *ACADSB* and *CMPK1* appeared to be specifically essential in Class2, indicating these two genes are associated with poor prognosis and may be the targets for treatment of more serious HCC patients.

The multi-scale regulatory-metabolic network revealed the critical mechanism of HCC cell proliferation

In addition to the integration of transcriptional regulation with metabolism, we further expanded our analysis to the epigenetic regulation carried by miRNAs. Based on the highlighted genes (total 28 key genes), we dig out miRNAs regulating these candidates from MIRNET. There are three miRNAs, miR-124-3p, miR-1-3p, and miR-24-3p reported as very important factors associated with HCC tumorigenesis and functioned in well-known cancer-related pathways like NOTCH, PI3K-Akt and mTOR. We illustrated the core network of HCC cell proliferation involving regulations between miRNAs-TFs, miRNAs-Targets and TFs-Targets across multi-scale (Fig. 5a), and emphasized the targets that were highlighted in the combined analyses. In general, the inhibition of miRNAs on over-expressed genes in HCC are consistent with their validated function such as suppressing tumorigenesis, which do provide supports to the key genes predicted as affecting cancer cell growth through cross-talking.

Especially, the direct regulation of miR-26b-3p on *ACADSB* and *CMPK1* provide more experimentally evidence to support the idea that these two metabolic genes are link with lower OS in HCC. Moreover, the biological connection inferred by IPA indicated these highlighted genes are closely connected to *EGFR*, which plays a significant role in cancer cell proliferation, providing evidence for our analyses.

Abbreviations

HCC

Hepatocellular carcinoma

TF

Transcriptional factor

OS

Overall survival

MERLIN

Modular regulatory network learning with per gene information

CMIP

Conditional mutual information measurement using a parallel computing framework

IDREAM

Integrated Deduced Regulation And Metabolism

EGRIN

Environment and Gene Regulatory Influence Network

NMF

Non-negative matrix factorization

IPA

Ingenuity Pathway Analysis

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The expression data that support the findings of this study are available in GEO Datasets with the identifier [doi:10.18632/oncotarget.8927](10)

The expression and clinical data that support the findings of this study are available in TCGA Database, LHC section [<https://portal.gdc.cancer.gov/>].

The expression and clinical data that support the findings of this study are available in ICGC Database, JIRI-JP section [<https://icgc.org/>].

The genome-scale metabolic liver model that support the findings of this study is available in HMA database [<https://metabolicatlas.org/gems/repository>].

The patient specific metabolic models that support the findings of this study are available in BioModels Database [<https://www.ebi.ac.uk/biomodels/>] with the identifier [doi:10.1126/science.aan2507](8)

Competing interests

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be constructed as influencing the position presented in, or the review of the manuscript entitled.

Funding

This work was supported by National Key Research and Development Program (2019YFA0905401, 2017YFC0908105), Shanghai Municipal Science and Technology Major Project (2017SHZDZX01), the Natural Science Foundation of China (U1804284, 81421061, 81701321, 31571012, 81501154), the Shanghai Natural Science Funding (16ZR1449700), Shanghai Hospital Development Center (SHDC12016115), Shanghai Science and Technology Committee (17JC1402900 and 17490712200),

Shanghai municipal health commission (ZK2015B01 and 201540114). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

YX was responsible for the data gathering, data preprocessing, statistical analysis, regulatory network inferring, model integration, metabolic analysis and manuscript writing; HZ was responsible for the TCGA-LIHC RNA-Seq data gathering and data preprocessing; QY was responsible for the IPA analysis; RS was responsible for the previous literature sharing; KW was responsible for the data preprocessing; YS managed the project.; ZW was responsible for the design of integrated model, explanation of results, and manuscript writing.

All authors have read and approved the manuscript.

Acknowledgements

Not applicable

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer*. 2015;136(5):E359-86.
2. Jiang Y, Sun A, Zhao Y, Ying W, Sun H, Yang X, et al. Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma. *Nature*. 2019;567(7747):257-61.
3. Gao Q, Zhu H, Dong L, Shi W, Chen R, Song Z, et al. Integrated Proteogenomic Characterization of HBV-Related Hepatocellular Carcinoma. *Cell*. 2019;179(2):561-77.e22.
4. Mardinoglu A, Nielsen J. Systems medicine and metabolic modelling. *Journal of internal medicine*. 2012;271(2):142-54.
5. Mardinoglu A, Nielsen J. New paradigms for metabolic modeling of human cells. *Current opinion in biotechnology*. 2015;34:91-7.
6. Folger O, Jerby L, Frezza C, Gottlieb E, Ruppin E, Shlomi T. Predicting selective drug targets in cancer through metabolic networks. *Mol Syst Biol*. 2011;7:501.
7. Mardinoglu A, Agren R, Kampf C, Asplund A, Uhlen M, Nielsen J. Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. *Nature communications*. 2014;5:3083.
8. Uhlen M, Zhang C, Lee S, Sjostedt E, Fagerberg L, Bidkhori G, et al. A pathology atlas of the human cancer transcriptome. *Science (New York, NY)*. 2017;357(6352).

9. Bidkhorji G, Benfeitas R, Klevstig M, Zhang C, Nielsen J, Uhlen M, et al. Metabolic network-based stratification of hepatocellular carcinoma reveals three distinct tumor subtypes. *Proceedings of the National Academy of Sciences of the United States of America*. 2018;115(50):E11874-e83.
10. Liu G, Hou G, Li L, Li Y, Zhou W, Liu L. Potential diagnostic and prognostic marker dimethylglycine dehydrogenase (DMGDH) suppresses hepatocellular carcinoma metastasis in vitro and in vivo. *Oncotarget*. 2016;7(22):32607-16.
11. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. *Science (New York, NY)*. 2015;347(6220):1260419.
12. Marbach D, Schaffter T, Mattiussi C, Floreano D. Generating realistic in silico gene networks for performance assessment of reverse engineering methods. *Journal of computational biology : a journal of computational molecular cell biology*. 2009;16(2):229-39.
13. Bonneau R, Facciotti MT, Reiss DJ, Schmid AK, Pan M, Kaur A, et al. A predictive model for transcriptional control of physiology in a free living cell. *Cell*. 2007;131(7):1354-65.
14. Marbach D, Lamparter D, Quon G, Kellis M, Kutalik Z, Bergmann S. Tissue-specific regulatory circuits reveal variable modular perturbations across complex diseases. *Nature methods*. 2016;13(4):366-70.
15. Liu ZP, Wu C, Miao H, Wu H. RegNetwork: an integrated database of transcriptional and post-transcriptional regulatory networks in human and mouse. *Database : the journal of biological databases and curation*. 2015;2015.
16. Huang F, Tang J, Zhuang X, Zhuang Y, Cheng W, Chen W, et al. MiR-196a promotes pancreatic cancer progression by targeting nuclear factor kappa-B-inhibitor alpha. *PloS one*. 2014;9(2):e87897.
17. Laos S, Baeckstrom D, Hansson GC. Inhibition of NF-kappaB activation and chemokine expression by the leukocyte glycoprotein, CD43, in colon cancer cells. *International journal of oncology*. 2006;28(3):695-704.
18. Steeg PS, Bevilacqua G, Kopper L, Thorgeirsson UP, Talmadge JE, Liotta LA, et al. Evidence for a novel gene associated with low tumor metastatic potential. *J Natl Cancer Inst*. 1988;80(3):200-4.
19. Hennessy C, Henry JA, May FE, Westley BR, Angus B, Lennard TW. Expression of the antimetastatic gene nm23 in human breast cancer: an association with good prognosis. *J Natl Cancer Inst*. 1991;83(4):281-5.
20. Thakur RK, Yadav VK, Kumar A, Singh A, Pal K, Hoepfner L, et al. Non-metastatic 2 (NME2)-mediated suppression of lung cancer metastasis involves transcriptional regulation of key cell adhesion factor vinculin. *Nucleic acids research*. 2014;42(18):11589-600.
21. Wang Z, Danziger SA, Heavner BD, Ma S, Smith JJ, Li S, et al. Combining inferred regulatory and reconstructed metabolic networks enhances phenotype prediction in yeast. *PLoS computational biology*. 2017;13(5):e1005489.
22. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research*. 2015;43(7):e47.
23. Rivory LP, Slaviero KA, Clarke SJ. Hepatic cytochrome P450 3A drug metabolism is reduced in cancer patients who have an acute-phase response. *British journal of cancer*. 2002;87(3):277-80.

24. Matos JM, Witzmann FA, Cummings OW, Schmidt CM. A pilot study of proteomic profiles of human hepatocellular carcinoma in the United States. *The Journal of surgical research*. 2009;155(2):237-43.
25. Peeters P, Raynaud SD, Cools J, Wlodarska I, Grosgeorge J, Philip P, et al. Fusion of TEL, the ETS-variant gene 6 (ETV6), to the receptor-associated kinase JAK2 as a result of t(9;12) in a lymphoid and t(9;15;12) in a myeloid leukemia. *Blood*. 1997;90(7):2535-40.
26. Carella C, Potter M, Bonten J, Rehg JE, Neale G, Grosveld GC. The ETS factor TEL2 is a hematopoietic oncoprotein. *Blood*. 2006;107(3):1124-32.
27. Cardone M, Kandilci A, Carella C, Nilsson JA, Brennan JA, Sirma S, et al. The novel ETS factor TEL2 cooperates with Myc in B lymphomagenesis. *Molecular and cellular biology*. 2005;25(6):2395-405.
28. Alessandrini F, Pezze L, Menendez D, Resnick MA, Ciribilli Y. ETV7-Mediated DNAJC15 Repression Leads to Doxorubicin Resistance in Breast Cancer Cells. *Neoplasia (New York, NY)*. 2018;20(8):857-70.
29. Harwood FC, Klein Geltink RI, O'Hara BP, Cardone M, Janke L, Finkelstein D, et al. ETV7 is an essential component of a rapamycin-insensitive mTOR complex in cancer. *Science advances*. 2018;4(9):eaar3938.
30. Blevins MA, Huang M, Zhao R. The Role of CtBP1 in Oncogenic Processes and Its Potential as a Therapeutic Target. *Molecular cancer therapeutics*. 2017;16(6):981-90.
31. Hildebrand JD, Soriano P. Overlapping and unique roles for C-terminal binding protein 1 (CtBP1) and CtBP2 during mouse development. *Molecular and cellular biology*. 2002;22(15):5296-307.
32. Vecchi C, Montosi G, Garuti C, Corradini E, Sabelli M, Canali S, et al. Gluconeogenic signals regulate iron homeostasis via hepcidin in mice. *Gastroenterology*. 2014;146(4):1060-9.
33. Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nature protocols*. 2019;14(3):639-702.
34. Liu NQ, De Marchi T, Timmermans A, Trapman-Jansen AM, Foekens R, Look MP, et al. Prognostic significance of nuclear expression of UMP-CMP kinase in triple negative breast cancer patients. *Scientific reports*. 2016;6:32027.
35. Zhou D, Zhang L, Sun W, Guan W, Lin Q, Ren W, et al. Cytidine monophosphate kinase is inhibited by the TGF-beta signalling pathway through the upregulation of miR-130b-3p in human epithelial ovarian cancer. *Cellular signalling*. 2017;35:197-207.
36. Zhou D, Zhang L, Lin Q, Ren W, Xu G. Data on the association of CMPK1 with clinicopathological features and biological effect in human epithelial ovarian cancer. *Data in brief*. 2017;13:77-84.
37. Ryu JS, Shin ES, Nam HS, Yi HG, Cho JH, Kim CS, et al. Differential effect of polymorphisms of CMPK1 and RRM1 on survival in advanced non-small cell lung cancer patients treated with gemcitabine or taxane/cisplatin. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2011;6(8):1320-9.
38. Ohmine K, Kawaguchi K, Ohtsuki S, Motoi F, Ohtsuka H, Kamiie J, et al. Quantitative Targeted Proteomics of Pancreatic Cancer: Deoxycytidine Kinase Protein Level Correlates to Progression-Free

- Survival of Patients Receiving Gemcitabine Treatment. *Molecular pharmaceutics*. 2015;12(9):3282-91.
39. Zhang B, Wu Q, Wang Z, Xu R, Hu X, Sun Y, et al. The promising novel biomarkers and candidate small molecule drugs in kidney renal clear cell carcinoma: Evidence from bioinformatics analysis of high-throughput data. *Molecular genetics & genomic medicine*. 2019;7(5):e607.
40. Song BL, Javitt NB, DeBose-Boyd RA. Insig-mediated degradation of HMG CoA reductase stimulated by lanosterol, an intermediate in the synthesis of cholesterol. *Cell metabolism*. 2005;1(3):179-89.
41. Mehrmohamadi M, Liu X, Shestov AA, Locasale JW. Characterization of the usage of the serine metabolic network in human cancer. *Cell reports*. 2014;9(4):1507-19.
42. Lee JS, Lee H, Lee S, Kang JH, Lee SH, Kim SG, et al. Loss of SLC25A11 causes suppression of NSCLC and melanoma tumor formation. *EBioMedicine*. 2019;40:184-97.
43. Baulies A, Montero J, Matias N, Insausti N, Terrones O, Basanez G, et al. The 2-oxoglutarate carrier promotes liver cancer by sustaining mitochondrial GSH despite cholesterol loading. *Redox biology*. 2018;14:164-77.
44. Villar VH, Merhi F, Djavaheiri-Mergny M, Duran RV. Glutaminolysis and autophagy in cancer. *Autophagy*. 2015;11(8):1198-208.
45. Yip KW, Zhang Z, Sakemura-Nakatsugawa N, Huang JW, Vu NM, Chiang YK, et al. A porphodimethene chemical inhibitor of uroporphyrinogen decarboxylase. *PloS one*. 2014;9(2):e89889.
46. Ge J, Yu Y, Xin F, Yang ZJ, Zhao HM, Wang X, et al. Downregulation of delta-aminolevulinic acid dehydratase is associated with poor prognosis in patients with breast cancer. *Cancer science*. 2017;108(4):604-11.
47. Lin M, Lv D, Zheng Y, Wu M, Xu C, Zhang Q, et al. Downregulation of CPT2 promotes tumorigenesis and chemoresistance to cisplatin in hepatocellular carcinoma. *OncoTargets and therapy*. 2018;11:3101-10.
48. Heise M, Lautem A, Knapstein J, Schattenberg JM, Hoppe-Lotichius M, Foltys D, et al. Downregulation of organic cation transporters OCT1 (SLC22A1) and OCT3 (SLC22A3) in human hepatocellular carcinoma and their prognostic significance. *BMC cancer*. 2012;12:109.
49. Keshet R, Erez A. Arginine and the metabolic regulation of nitric oxide synthesis in cancer. *Disease models & mechanisms*. 2018;11(8).
50. Xiong L, Teng JL, Botelho MG, Lo RC, Lau SK, Woo PC. Arginine Metabolism in Bacterial Pathogenesis and Cancer Therapy. *International journal of molecular sciences*. 2016;17(3):363.
51. Amelio I, Cutruzzola F, Antonov A, Agostini M, Melino G. Serine and glycine metabolism in cancer. *Trends in biochemical sciences*. 2014;39(4):191-8.
52. Phang JM, Liu W, Hancock CN, Fischer JW. Proline metabolism and cancer: emerging links to glutamine and collagen. *Current opinion in clinical nutrition and metabolic care*. 2015;18(1):71-7.

53. Wu X, Deng F, Li Y, Daniels G, Du X, Ren Q, et al. ACSL4 promotes prostate cancer growth, invasion and hormonal resistance. *Oncotarget*. 2015;6(42):44849-63.
54. Koberle V, Kronenberger B, Pleli T, Trojan J, Imelmann E, Peveling-Oberhag J, et al. Serum microRNA-1 and microRNA-122 are prognostic markers in patients with hepatocellular carcinoma. *European journal of cancer (Oxford, England : 1990)*. 2013;49(16):3442-9.
55. Lang Q, Ling C. MiR-124 suppresses cell proliferation in hepatocellular carcinoma by targeting PIK3CA. *Biochemical and biophysical research communications*. 2012;426(2):247-52.
56. Murakami Y, Yasuda T, Saigo K, Urashima T, Toyoda H, Okanoue T, et al. Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. *Oncogene*. 2006;25(17):2537-45.
57. Dai Y, Sui W, Lan H, Yan Q, Huang H, Huang Y. Comprehensive analysis of microRNA expression patterns in renal biopsies of lupus nephritis patients. *Rheumatology international*. 2009;29(7):749-54.
58. Van Vlierberghe P, De Weer A, Mestdagh P, Feys T, De Preter K, De Paepe P, et al. Comparison of miRNA profiles of microdissected Hodgkin/Reed-Sternberg cells and Hodgkin cell lines versus CD77+ B-cells reveals a distinct subset of differentially expressed miRNAs. *British journal of haematology*. 2009;147(5):686-90.
59. Zheng F, Liao YJ, Cai MY, Liu YH, Liu TH, Chen SP, et al. The putative tumour suppressor microRNA-124 modulates hepatocellular carcinoma cell aggressiveness by repressing ROCK2 and EZH2. *Gut*. 2012;61(2):278-89.
60. Furuta M, Kozaki KI, Tanaka S, Aii S, Imoto I, Inazawa J. miR-124 and miR-203 are epigenetically silenced tumor-suppressive microRNAs in hepatocellular carcinoma. *Carcinogenesis*. 2010;31(5):766-76.
61. Zhang H, Zhang Z, Gao L, Qiao Z, Yu M, Yu B, et al. miR-1-3p suppresses proliferation of hepatocellular carcinoma through targeting SOX9. *OncoTargets and therapy*. 2019;12:2149-57.
62. Salloum-Asfar S, Arroyo AB, Teruel-Montoya R, Garcia-Barbera N, Roldan V, Vicente V, et al. MiRNA-Based Regulation of Hemostatic Factors through Hepatic Nuclear Factor-4 Alpha. *PLoS one*. 2016;11(5):e0154751.
63. Wang G, Dong F, Xu Z, Sharma S, Hu X, Chen D, et al. MicroRNA profile in HBV-induced infection and hepatocellular carcinoma. *BMC cancer*. 2017;17(1):805.
64. Chen L, Luo L, Chen W, Xu HX, Chen F, Chen LZ, et al. MicroRNA-24 increases hepatocellular carcinoma cell metastasis and invasion by targeting p53: miR-24 targeted p53. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie*. 2016;84:1113-8.
65. Nicolas M, Wolfer A, Raj K, Kummer JA, Mill P, van Noort M, et al. Notch1 functions as a tumor suppressor in mouse skin. *Nature genetics*. 2003;33(3):416-21.
66. Wang Y, Sun B, Zhao X, Zhao N, Sun R, Zhu D, et al. Twist1-related miR-26b-5p suppresses epithelial-mesenchymal transition, migration and invasion by targeting SMAD1 in hepatocellular carcinoma. *Oncotarget*. 2016;7(17):24383-401.

Additional Files

Supplementary Table1: SupplementaryTable1.xlsx

Composition of normal/hcc liver model: The regulatory network of normal/hcc liver and the composition of integrated models

Supplementary Table2: SupplementaryTable2.xlsx

TFs affecting cell growth: TFs affecting normal/hcc liver cell growth and corresponding growth ratio after TFs knocking down

Supplementary Table3: SupplementaryTable3.csv

Differentially expressed genes: Differentially expressed genes in Class2 compared to Class1 and Class3 in TCGA-LIHC samples

Supplementary Table4: SupplementaryTable4.xlsx

Altered metabolic reactions: Altered metabolic reactions in Class2 compared to Class1 and Class3 in TCGA-LIHC samples

Supplementary Table5: SupplementaryTable5.csv

Genes for core network: 28 core genes filtered according to our research

Figures

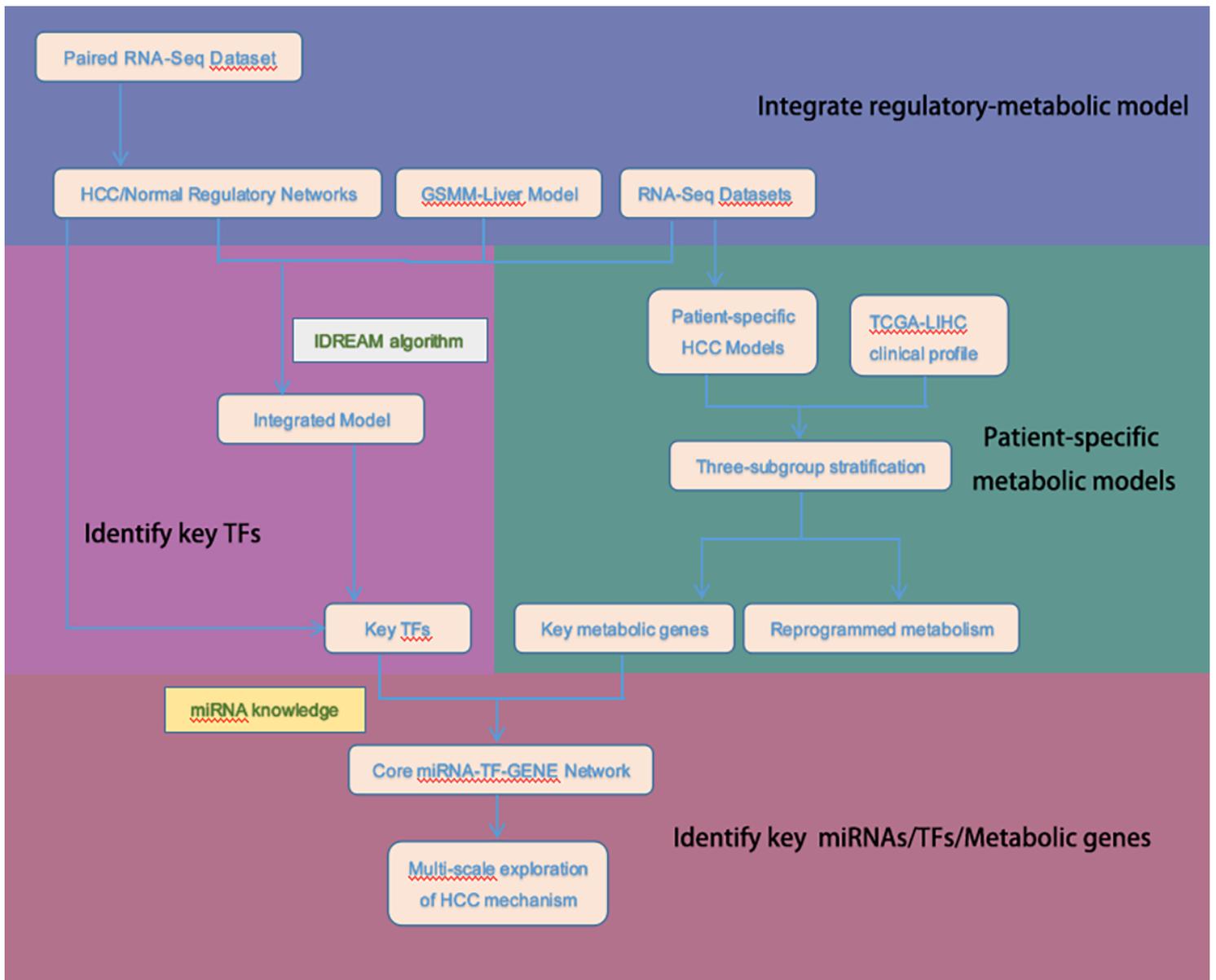


Figure 1

The schema of the integrated model for stratification and key targets discovery.

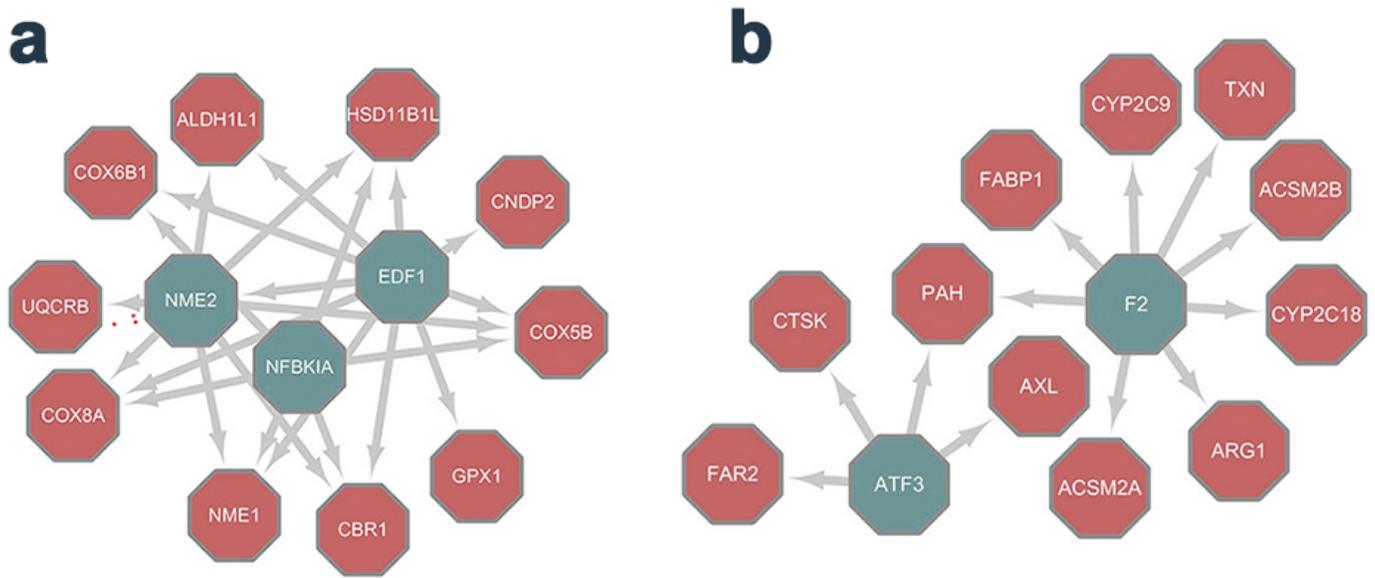


Figure 2

Core structure of different nodes between Normal/Tumor Regulatory topology networks. a. Difference between normal and tumor networks. b. Difference between tumor and normal networks; Nodes filled with green are TFs while nodes filled with red are metabolic genes.

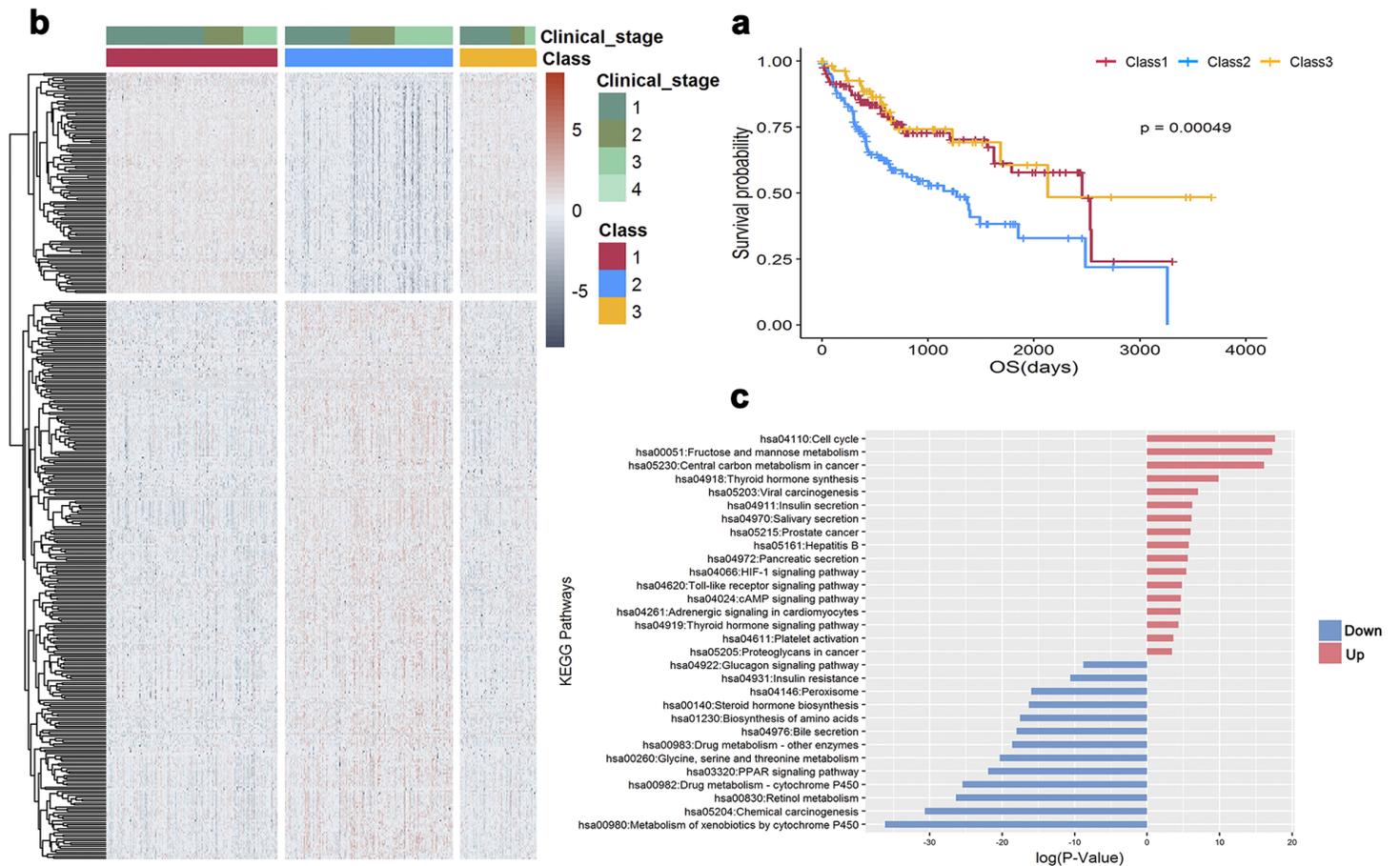


Figure 3

Stratification of 315 TCGA-LIHC samples. a. Kaplan-Meier curve shows the survival outcomes of Class1 (red), Class2 (blue) and Class3 (yellow), the P-value is 0.00049 which is significant. b. Heatmap of 399 differentially expressed genes revealed a distinguishable pattern between Class2 and Class1&3. Red color represents up-regulated while blue represents down-regulated. c. Enrichment analysis of KEGG pathways of 287 up-regulated and 112 down-regulated genes, respectively.

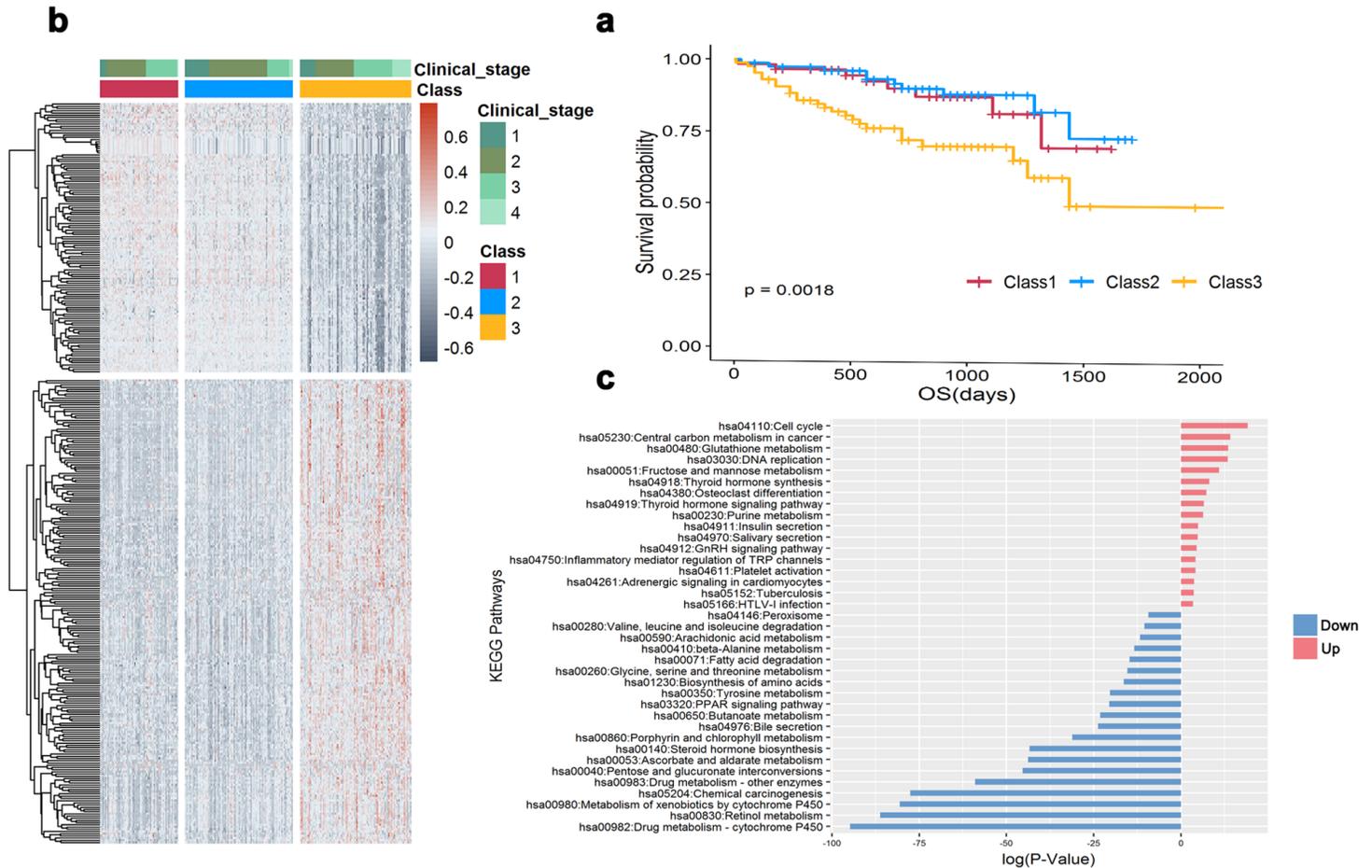


Figure 4

Stratification of 232 LIRI-JP samples. a. Kaplan-Meier curve shows the survival outcomes of Class1 (red), Class2 (blue) and Class3 (yellow), the P-value is 0.0018 which is significant. b. Heatmap of 332 differentially expressed genes revealed a distinguishable pattern between Class2 and Class1&3. Red color represents up-regulated while blue represents down-regulated. c. Enrichment analysis of KEGG pathways of 210 up-regulated and 122 down-regulated genes, respectively.

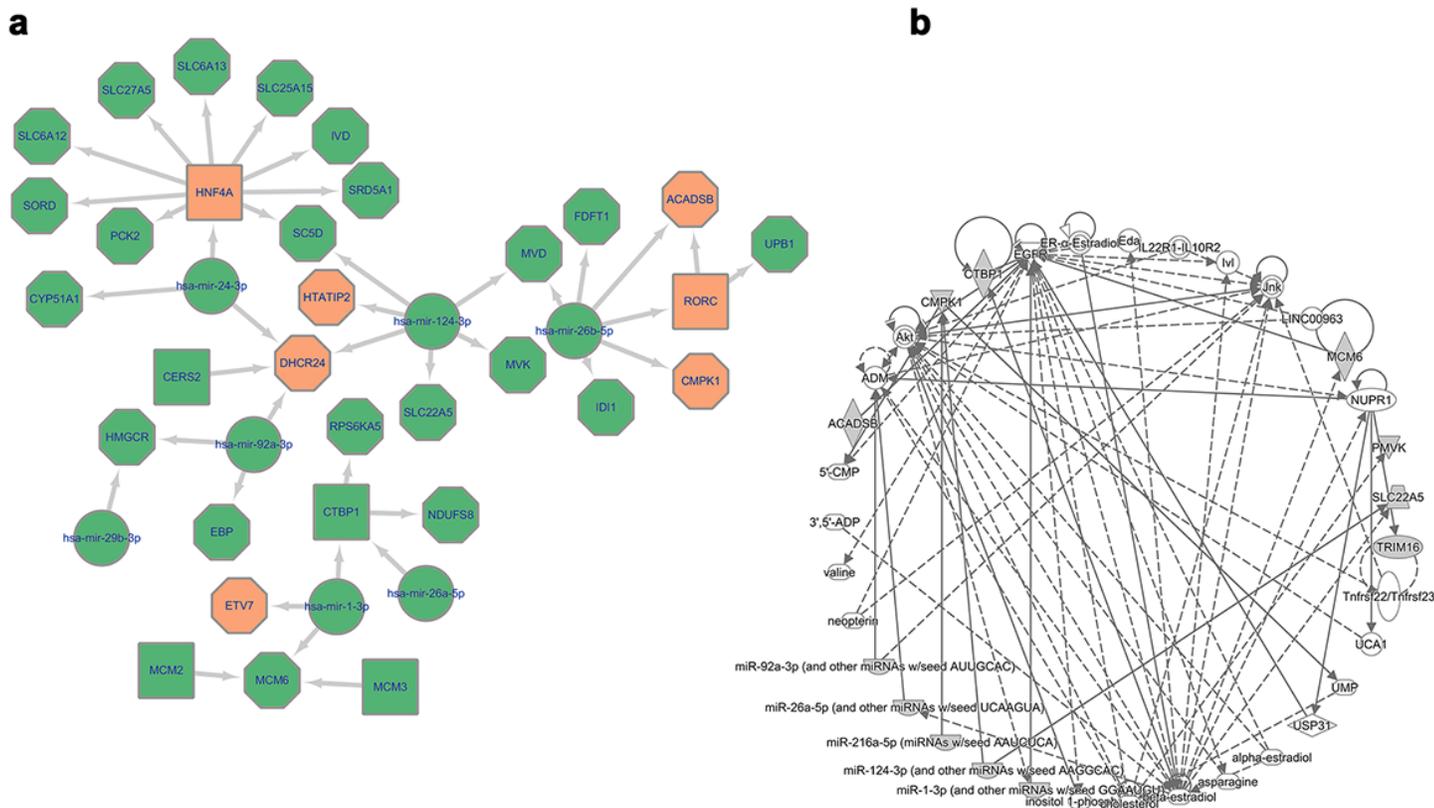


Figure 5

Multi-scale network exploration of HCC mechanism. a. Core miRNA relation network of potential therapeutic targets. The octagon represents metabolic genes while the square represents TFs and the circle represents miRNAs; The orange color represents genes/TFs that have been experimentally validated as crucial genes in HCC tumor cells; b. Biological connection of 34 genes (28 candidate genes and top 6 miRNAs)

Supplementary Files

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

- [SupplementaryTable1.xlsx](#)
- [SupplementaryTable4.xlsx](#)
- [SupplementaryTable5.csv](#)
- [SupplementaryTable3.csv](#)
- [SupplementaryTable2.xlsx](#)