Computers in Biology and Medicine 159 (2023) 106957

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Integrative computational modeling to unravel novel potential biomarkers in hepatocellular carcinoma

Andrew Walakira [a,*], Cene Skubic [a], Nejc Nadižar [a], Damjana Rozman [a], Tadeja Režen [a], Miha Mraz [b], Miha Moškon [b,*]

[a] Centre for Functional Genomics and Bio-Chips, Institute for Biochemistry and Molecular Genetics, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia
[b] Faculty of Computer and Information Science, University of Ljubljana, Ljubljana, Slovenia

ARTICLE INFO

Keywords:
Hepatocellular carcinoma
Meta-analysis
Genome-scale metabolic modeling
Network analysis
Enriched pathways
Enriched reactions
Enriched subsystems

ABSTRACT

Hepatocellular carcinoma (HCC) is a major health problem around the world. The management of this disease is complicated by the lack of noninvasive diagnostic tools and the few treatment options available. Better clinical outcomes can be achieved if HCC is detected early, but unfortunately, clinical signs appear when the disease is in its late stages. We aim to identify novel genes that can be targeted for the diagnosis and therapy of HCC. We performed a meta-analysis of transcriptomics data to identify differentially expressed genes and applied network analysis to identify hub genes. Fatty acid metabolism, complement and coagulation cascade, chemical carcinogenesis and retinol metabolism were identified as key pathways in HCC. Furthermore, we integrated transcriptomics data into a reference human genome-scale metabolic model to identify key reactions and subsystems relevant in HCC. We conclude that fatty acid activation, purine metabolism, vitamin D, and E metabolism are key processes in the development of HCC and therefore need to be further explored for the development of new therapies. We provide the first evidence that *GABRP*, *HBG1* and *DAK (TKFC)* genes are important in HCC in humans and warrant further studies.

1. Introduction

Liver cancer is a major human health problem around the world. It can be classified as primary, i.e., originating in the liver, or secondary, i.e., caused by metastasis of another type of cancer to the liver. There were approximately 906,000 new cases of primary liver cancer (4.7% of all new cancer cases) and 830,000 liver cancer deaths (8.3% of all cancer deaths) in 2020 [1]. The 5-year prevalence statistics show that liver cancer is more prevalent in Asia (73.6%), followed by Europe (8.6%), Africa (8.4%), and North America (5%), with an incidence at least two times higher in men compared to women [1]. The two most prevalent forms of primary liver cancer are hepatocellular carcinoma (contributing approximately 80%) and cholangiocarcinoma (approximately 10%–15%) [1]. Chronic infection with the hepatitis B and/or C virus is known to be the major risk factor for hepatocellular carcinoma (HCC). However, the prevalence of hepatitis B and C viruses has decreased in recent years due to the development of effective vaccines and improved management strategies [2] but yet the prevalence of HCC continues to increase. Furthermore, even with improvements in diagnosis and management of liver cancers, especially HCC, major challenges still remain due to the few treatment options available and the lack of easy to use noninvasive diagnostic tools. This has made it necessary to search for novel biomarkers that can be targeted for the diagnosis and/or treatment of HCC.

The diagnosis and treatment of HCC are difficult because clinical symptoms are characteristically present when the disease is in advanced stages. Current diagnostic tools are expensive, require specialized expertise and are often invasive. Imaging techniques such as ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI) provide noninvasive means of diagnosing HCC [3] but are expensive to set up and require specialized training for efficient use. Digital subtraction angiography (DSA) is efficient at detailing the tumor anatomy and its vascular supply [3] but is invasive, expensive and requires specialized personnel. Histological diagnostic techniques are a valuable tool in HCC diagnosis because they provide definitive pathological diagnosis, information on nature of tumor and etiology of the disease but are invasive, require specialized personnel and are not applicable in all scenarios because of the risk of bleeding and needle tract implantation during collection of biopsies [3]. Serum biomarkers such as alpha feto protein (AFP) [4], prothrombin induced by the absence of vitamin K or antagonist II (PIVKA-II) [5], and plasma-free microRNA [6] are gaining momentum, but must be used in conjunction

* Corresponding authors.
E-mail addresses: andrew.walakira@mf.uni-lj.si (A. Walakira), miha.moskon@fri.uni-lj.si (M. Moškon).

https://doi.org/10.1016/j.compbiomed.2023.106957
Received 15 November 2022; Received in revised form 17 March 2023; Accepted 16 April 2023
Available online 20 April 2023

Table 1
Description of datasets.

Accession ID.	Patients(M/F)	HBV/ HCV %	Median age (range)	Year	Tissue	Platform	Country
GSE19665	10 (9/1)	100	67 (51–74)	2009	Tumor /nontumor	GPL570 - Affymetrix Human Genome U133	Tokyo, Japan
GSE39791	72 (58/14)	*83	57.5 (29–77)	2012	Tumor /nontumor	GPL10558 - Illumina HumanHT-12 V4.0	Daegu & Seoul, Korea
GSE41804	20	100	–	2012	Tumor /nontumor	GPL570-Affymetrix Human Genome U133	Kanazawa, Japan
GSE57957	39 (*35/4)	*61	*65 (35–85)	2014	Tumor /nontumor	GPL10558-Illumina HumanHT-12 V4.0	Singapore
GSE64041	60 (*53/7)	*30	+*64	2016	Tumor /nontumor	GPL6244-Affymetrix Human Gene 1.0 ST	Basel, Switzerland
GSE84402	14 (9/5)	–	45 (35–67)	2016	Tumor /nontumor	GPL570-Affymetrix Human Genome U133	Shanghai, China
GSE84598	22 (*17/5)	*47	–	2016	Tumor /nontumor	GPL10558-Illumina HumanHT-12 V4.0	Mainz, German
HCCDB15	49 (28/21)	–	68 (20–81)	~2015	Tumor /nontumor	RNA Seq	USA

* Estimated from reference publication, +mean (not median). "Accession ID" is the identification of the dataset in the public database, "Patients" refers to the number of patients with data from tumor and nontumor tissues, "HBV/HCV" is the percentage of patients with chronic hepatitis B/C infection, "median age" is the median age of the patients in a particular study, "year" refers to publication year, "tissue" is the liver tissue from tumor and surrounding nontumor, "platform" is the type of chip or technical platform used to measure gene expression in the samples, "country" is the country where patients originated.

with more established tools. Early detection of this disease would greatly improve patient management and thus better clinical outcomes. Currently, there are several options for treating HCC, such as surgical resection, local ablation, and liver transplantation, among others. These treatments are expensive, often very aggressive to the patients and do not treat HCC completely as the chances of recurrence are high. Therefore, there is still an urgent need for easy and noninvasive diagnostic tools and novel therapies for the management of HCC.

In all, there is a dire need to find novel biomarkers that can be used as noninvasive diagnostic or drug targets. In this study, we leveraged publicly available transcriptomics data from previous studies that compared tumor and nontumor biopsies from the same patients. We performed a meta-analysis to find differentially expressed genes in all studies and performed a network analysis. Furthermore, we applied genome-scale metabolic modeling to find enriched reactions and subsystems. By comparing transcriptomics data from tumor and surrounding nontumor tissue, the rationale was to find novel genes and biological processes that are important in the development and progression of HCC in humans. The identification of genes associated with HCC in previous studies validates our modeling approach as relevant.

2. Materials and methods

In this section, we describe the data used, their processing, and the analytical methods applied, as summarized in Fig. 1.

2.1. The data

We searched the CancerLivER database [7] and the HCCDB database [8] for transcriptomics datasets curated from public repositories and the literature. The CancerLivER database provided information on microarray datasets, while the HCCDB provided preprocessed RNA seq datasets from The Cancer Genome Atlas (TCGA). We downloaded datasets that contained both tumor and surrounding nontumor samples. We considered only datasets that had 10 or more patients with at least 20,000 genes analyzed. The final analysis included eight datasets described in Table 1. The 8 datasets used for the meta-analysis all together contained 286 patients with approximately 79% males and 21% females. Datasets were from samples collected from Europe, North America, and Asia. Some of the samples were from people who had chronic infection with hepatitis B or C viruses, a major risk factor in HCC.

2.2. Differential gene expression analysis

Transcriptomics data were preprocessed following standard procedures. Briefly, data quality was assessed, expression values were $\log_2$ transformed and normalized using quantile normalization. Gene probes without gene annotation were removed. For multiple probes that matched the same gene, the mean was used as the final expression value. Gene expression values of tumor and nontumor tissue were compared for analysis of differentially expressed (DE) genes using linear models in *limma* library [9]. Log_2 fold change ($\log_2$FC) and p-values were determined. P-values were corrected for multiple testing using the Benjamini–Hochberg (BH) method. All analyses in this section were performed in R software, version 4.0.5 [10].

2.3. Meta-analysis of gene expression

Genes that were present in six or more datasets were considered for meta-analysis. This is because the studies included in the analysis used different gene expression platforms, so we intended to work with genes that were common to most of the platforms. Meta-analysis of differentially expressed genes was done by fitting random effects models for each gene using the *rma.uni* function of the *metafor* library [11,12] in R software. Log_2 fold change ($\log_2$FC) was used as the effect size (y_i) which was assumed to be an unbiased estimate of the true effect size (θ_i).

$$y_i = \theta_i + \epsilon \tag{1}$$

where $i = 1.....\ k$ independent effect size estimates, $\epsilon \sim N(0, \sigma^2)$ with σ^2 denoting the variance in the ith study. Each θ_i consists of the true effect (μ) and the random effect (μ_i) that denotes the difference between θ_i and μ. Hence, the final model was formulated as:

$$y_i = \mu + \mu_i + \epsilon \tag{2}$$

$\theta_i \sim N(\mu, \tau^2)$, the true effect in the study population is normally distributed with μ denoting the average true effect and τ^2 the variance (that is, the amount of heterogeneity) of the true effects in the study population, and $\mu_i \sim N(0, \tau^2)$. The model provides an estimate for μ (mean true effect) and τ^2 (heterogeneity). A gene was considered significantly differentially expressed (SDE) if its mean $|\log_2\text{FC}| >= 1$ and unadjusted p-value was less than 0.05. KEGG enrichment analysis of gene sets was done on SDE genes using the *enrichKEGG* function of the *clusterProfiler* library in R software [13]. Pathways were considered

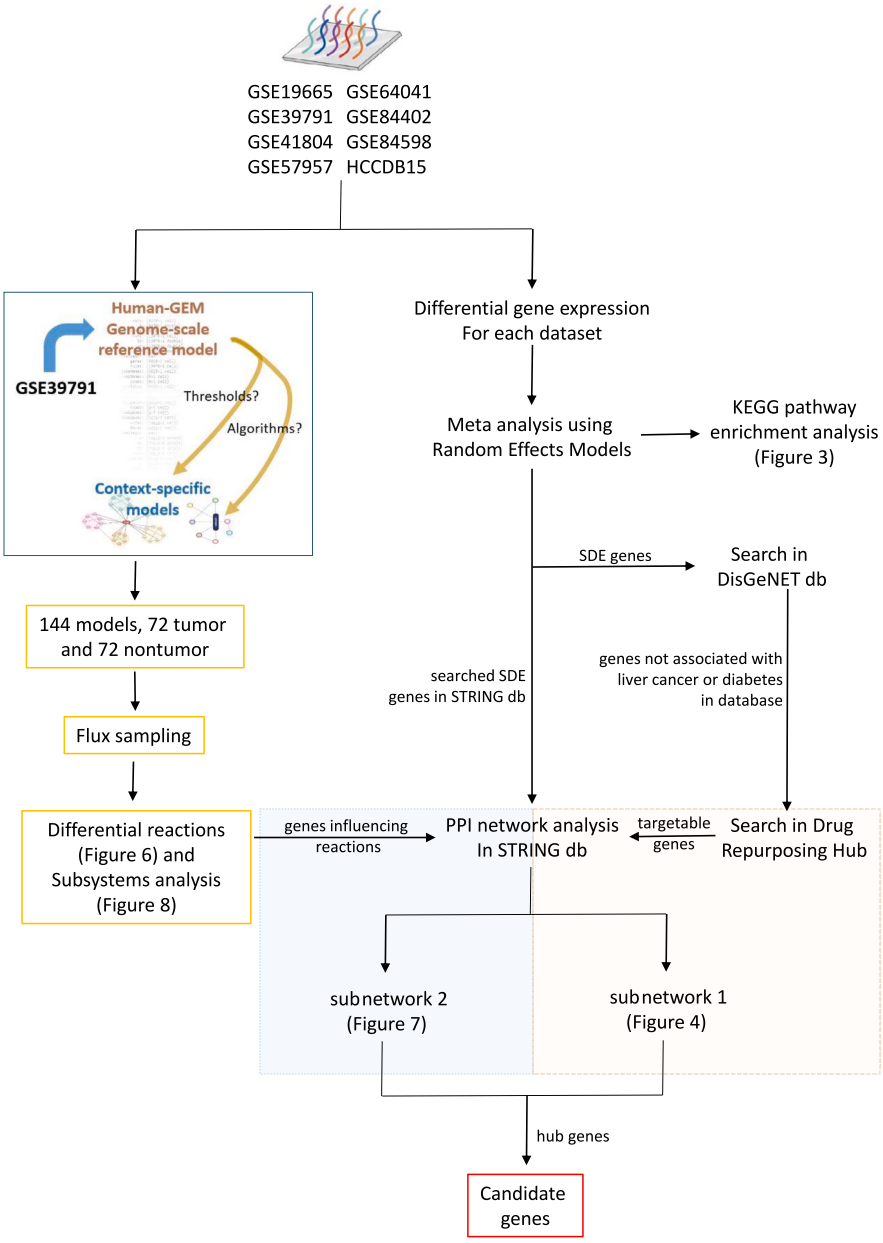

Fig. 1. Analysis workflow. On the right, we did a meta-analysis and generated a network (supp Fig. 4) by searching SDE genes (p value < 0.05) in STRING db. SDE genes were then searched in DisGeNET db to select only genes that are not associated with liver cancer or diabetes. Selected genes were then searched in Drug Repurposing Hub to select a set of genes that is being targeted for treatment of other ailments. This final set of genes was then used to mine the network above to yield a subnetwork which was analyzed for hub genes. On the left, transcriptomics data was integrated into Human-GEM to yield context specific models. This was followed by flux sampling and analysis of differential reactions and subsystems. Genes involved with significantly differentiated reactions and were also in the list of SDE genes from the meta-analysis were selected and used to extract a subnetwork. The subnetwork was analyzed for hub genes. All hub genes were considered as the final set of candidate genes. SDE: Significantly Differentially Expressed, KEGG: Kyoto Encyclopedia of Genes and Genomes, PPI: Protein–Protein Interaction, STRING db: STRING database.

significantly enriched if their Benjamini–Hochberg (BH) adjusted *p*-value was less than 0.1. Furthermore, SDE genes were searched in the DisGeNET database [14] (on 9 and 10 January 2022) to determine if there has been any previous association of a gene with any form of liver cancer or diabetes. We consider the DisGeNET database to be a well curated and up-to-date database for gene–disease associations and therefore used to identify new genes associated with HCC. Diabetes was included in the elimination of genes because it is known to be a precursor of several conditions, including cancers [15]. The remaining genes were then searched in a drug re-purposing database, the Drug Repurposing Hub (DRH) of Broad Institute [16], to identify genes being targeted in other ailments. Genes that were being targeted in other ailments were considered to be viable drug targets.

2.4. Protein–protein interaction (PPI) analysis

To derive Protein–Protein interaction networks, the SDE genes (and corresponding proteins) of the meta-analysis were searched in STRING db [17] with default settings. This yielded what we referred to here as the "parent" network. The genes that were found to be targetable in the DRH were used to extract a subnetwork from the parent network. This was done by querying the parent network in Cytoscape application.

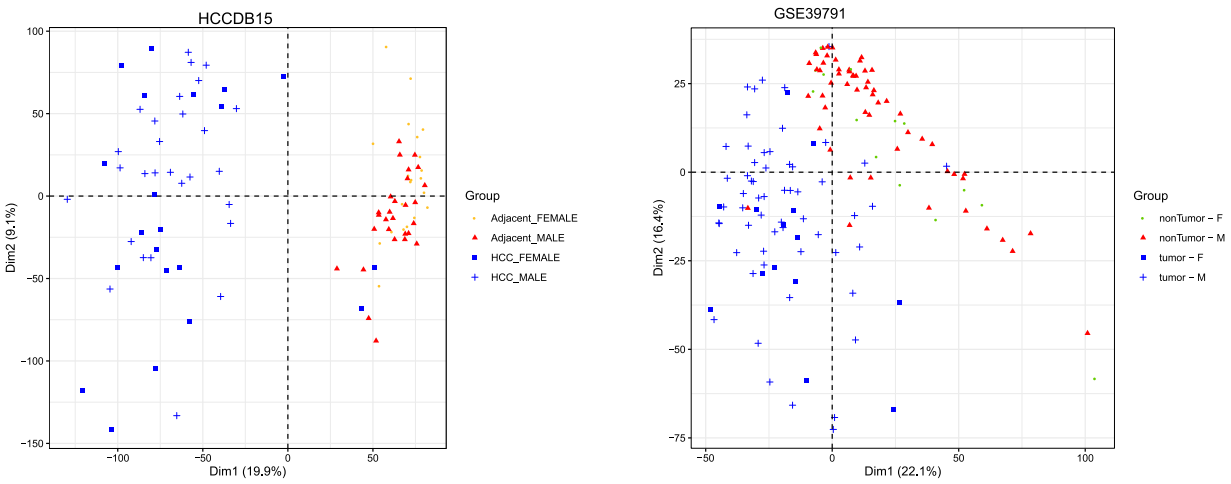

Fig. 2. There was no separation of patients by gender using PCA. Plots are showing two out of eight datasets.

The aim was to identify hub genes among the targetable genes or their immediate neighbors. This analysis was done in Cytoscape version 3.8.2, and STRING db was accessed on 07th January 2022 using the stringApp version 1.7.0.

2.5. Genome-scale metabolic modeling

Extraction of genome-scale metabolic models (GEMs) was performed in Matlab R2019b (MathWorks Inc., Natick, Massachusetts, USA) using the COBRA toolbox [18], RAVEN [19] and gurobi solver [20]. This was done following and extending the protocol described by Walakira et al. [21]. The GSE39791 dataset was chosen because it had a sufficient number of patients, that is, 72 patients with each patient having two samples, thus 72 tumor samples and 72 nontumour tissue samples. Transcriptomics data from the 144 samples were integrated into the human reference genome-scale model called Human-GEM [22] (version 1.9.0). Human-GEM model used has 8370 metabolites, 13,078 reactions, 3625 genes and no dead-end metabolites or blocked reactions. Different model extraction methods (MEMs) were applied to generate personalized models for each patient (a recent review of these methods is available in [23]). The extraction process yielded 144 models from each MEM. The models were extracted using five MEMs namely; Gene Inactivity Moderated by Metabolism and Expression (GIMME) [24,25], Integrative Metabolic Analysis Tool (iMAT) [26], FASTCORE, Integrative Network Inference for Tissues (INIT) [27] and task Integrative Network Inference for Tissues (tINIT) [28].

Different within patient thresholds were set to define low- and highly expressed genes. A lower and upper percentile (L–U) of the gene expression within a patient was used to set the thresholds for lowly and highly expressed genes as shown: 40th–80th, 50th–80th, 60th–80th, 50th–90th, 70th–90th, and 80th–90th. We used within patient thresholds because, for any gene, the degree of expression is different between patients, since individuals are biologically distinct even under similar conditions. For every model, a matrix was generated showing if a reaction was present (1) or absent (0). Reactions that were present or absent in all observed models were removed and the matrix were zero-centered on row means. Then, a principal component analysis (PCA) of the reactions was performed to determine which models separated the best between tumor and nontumor tissue. The t-distributed stochastic neighbor embedding (t-SNE) algorithm was used to confirm the model separation in PCA as used in [29,30]. Models extracted using the GIMME method produced the best separation by tissue type, i.e. tumor vs nontumor tissue (see Fig. 5) and thus these models were used for downstream analysis.

We performed flux sampling using the artificial centering hit-and-run (ACHR) algorithm [31] on the extracted models. This was done to assess their dynamic response. 1000 flux samples were generated for each of the models. Flux samples from models of the same patient i.e., tumor and nontumor were compared for enriched reactions using Mann–Whitney U test. The change in flux (flux change) was assessed as described in [32]. Reactions were considered statistically significantly different if their BH adjusted p-value was less than 0.05 and the flux change was 10-fold or higher, and were used to identify key genes. Reactions with adjusted p-value < 0.05 were then used to perform the enrichment analysis of metabolic subsystems using the hypergeometric test and p-values adjusted using BH. Subsystems with p-value < 0.05 were considered as enriched.

2.6. Gene signatures

We assessed the prognostic potential of selected genes of HCC using the Kaplan–Meier plotter (KM plotter) [33], at default settings. The HCC dataset used was contained in the KM plotter tool and is described in [34]. For each gene, patients were separated into high and low expression groups. A cut-off value to determine the groups of high and low gene expression was calculated by iterating over all values between the upper and lower quantiles of gene expression while computing the Cox regression for each setting. The most significant value was used as the cutoff point to identify the two groups. The two patient cohorts were then compared by a Kaplan–Meier survival plot, the hazard ratio with 95% confidence intervals and log rank p-values were calculated.

Furthermore, we assessed the potential of the selected genes to classify between tumor and nontumor tissue samples from an independent RNA-Seq dataset (the HCCDB18 dataset) which was not previously used in this analysis. This dataset presents transcriptomic data between tumor tissues and nontumor tissues collected from the Japanese population (median age 69 years, age range 31–86 years, 86% had HBV/HBV infection, 75% were men). Seven algorithms were applied using the CMA package in R software [35]; linear discriminant analysis (LDA), diagonal linear discriminant analysis (LDA), random forest (RF), elastic net (EN), L1 penalized logistic regression (Lasso), Partial Least Squares followed by logistic regression (PLS-LR), L2 penalized logistic regression (P-LR). Learning and training sets were generated using Monte-Carlo cross validation with 90% of observations used for learning at each iteration for 5000 iterations, misclassification error (MCE) and area under curve (AUC) were used to assess classification.

3. Results

This study used eight datasets for the meta-analysis, contributing 286 patients in total, 79% males and 21% females. Table 1 summarizes information about the eight datasets used. More than 70% of the

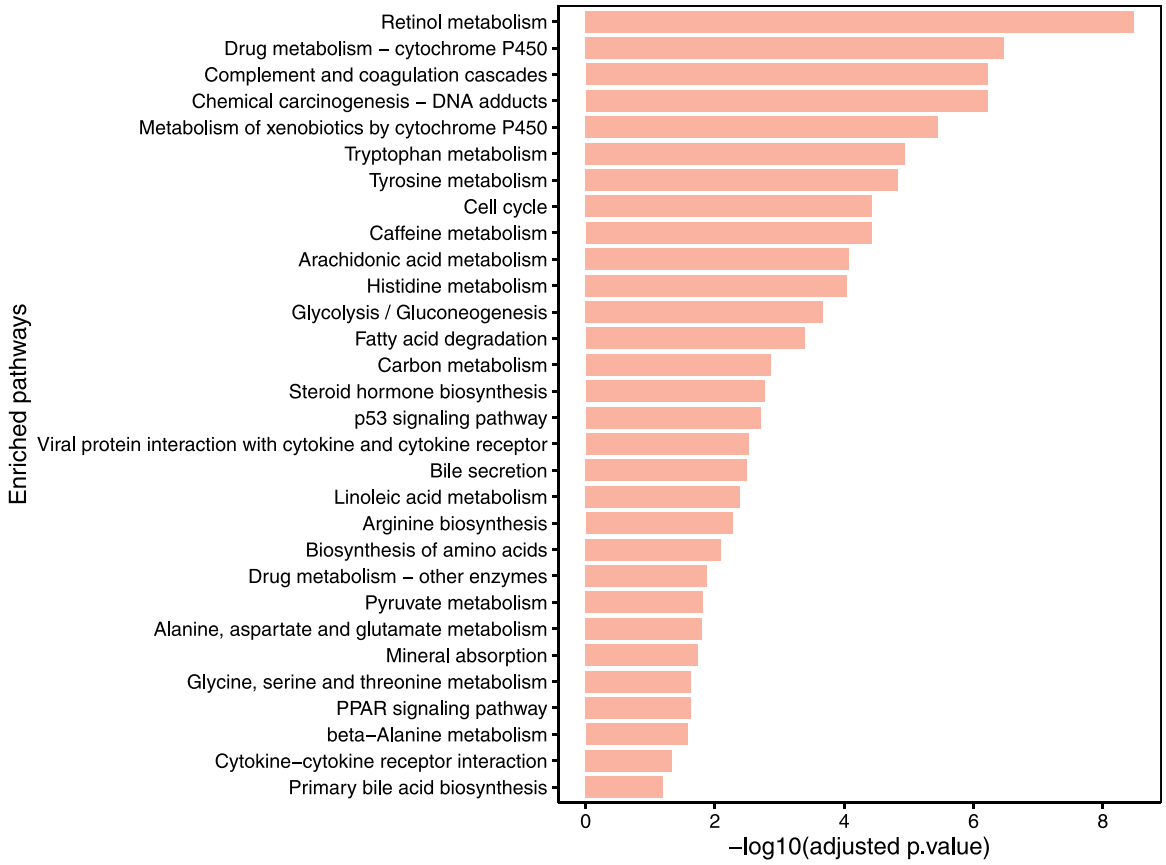

Fig. 3. KEGG enriched pathways.

samples were from patients infected with hepatitis B or C virus. The age distribution of the samples ranged from 20 to 86 years, with the majority of patients aged 50 years or older. For datasets where we had information on age, the age distribution is summarized in Supplementary Figure 1. There was a good separation of the samples by tissue type (tumor versus nontumor) using PCA on the transcriptomics data (Supplementary Figure 2). However, patients did not separate by gender on the PCA plot (Fig. 2).

Differential gene expression analysis was performed for each dataset using limma in R software. This produced a $\log_2$ fold change for each gene, and this was used as the effect size in the meta-analysis. There were 16,625 genes that were common in six or more datasets and these are the ones that were analyzed in the meta-analysis. 690 genes were significantly differentially expressed after the meta-analysis, 500 genes were downregulated while 190 genes were upregulated (Supplementary Figure 3 and Supplementary Table 1). Significantly differentially expressed (SDE) genes for KEGG-enriched pathways were analyzed. 30 pathways were found to be enriched. The majority of enriched pathways were those involved in metabolism while others are involved in signaling, synthesis of biomolecules, mineral absorption and immunity (Fig. 3).

The SDE (690) genes were searched in the DisGeNET database to remove genes associated with any form of liver cancer or diabetes. We remained with 128 genes without any prior association with liver cancer or diabetes. The 128 genes were searched in the DRH database to identify genes that are being targeted for the treatment of any other disease. Genes namely *PIPOX, HBG1, P2RY13, GABRP, CYP3A43, HAO1, IGJ, PDE2 A, PROZ, RDH5, S100P, C7, C8 A, IYD, OXT* were identified to be viable targets. These genes were assigned to the "parent" network (Supplementary Figure 4) to extract a subnetwork showing their connectivity and immediate neighbors. *C8 A, C7, HAO1, PIPOX, RDH5, OXT, IYD, CYP3A43, DAK, GABRP, P2RY13, HBG1* and *PROZ* were identified as key hub genes in the subnetwork (Fig. 4) and were all downregulated in the meta-analysis.

Transcriptomics data, GSE39791, were integrated into the Human-GEM reference human model. Context specific models were extracted using five MEMs as described in methods Section 2.5. Models extracted using GIMME at the 50th and 80th percentiles per patient as thresholds for lowly and highly expressed genes separated best by tissue type (Fig. 5, see also Supplementary Figure 5). These models were analyzed per patient by comparing 1000 flux samples of each reaction of the model from tumor and nontumor tissue. Reactions were considered significantly enriched if their BH adjusted p-value was < 0.05 and the flux change was 10 fold or higher. Reaction enrichment analysis identified 1128 reactions as significantly enriched in at least one patient. Downstream analysis to identify key reactions excluded transport and exchange/demand reactions. Reactions that were fully up or downregulated in all patients were considered as the final set of reactions of interest (Fig. 6). Reactions in red are upregulated while reactions in blue are downregulated. All reactions (19) in fatty acid activation were up regulated except MAR00425, which yields docosapentaenoyl-CoA and AMP but not ATP in the cytosol. 3 of 8 reactions (MAR04020, MAR04452, MAR04603) belonging to purine metabolism producing guanosine monophosphate, guanine and deoxyguanosine, respectively, in cytosol, 1 of 2 reactions (MAR08011) of vitamin D metabolism producing provitamin D_3, and 4 of 7 reactions of vitamin E metabolism (MAR03047, MAR04007, MAR06457, MAR06460) that produced alpha-tocopheryl quinone, alpha-hydroxy-gamma-tocopherone, gamma-tocopheroxyl-radical, and gamma-tocopherol, respectively, were downregulated. All other reactions were upregulated and belonged to aminoacyl-tRNA biosynthesis, nicotinamide metabolism, and Oxidative phosphorylation in addition to

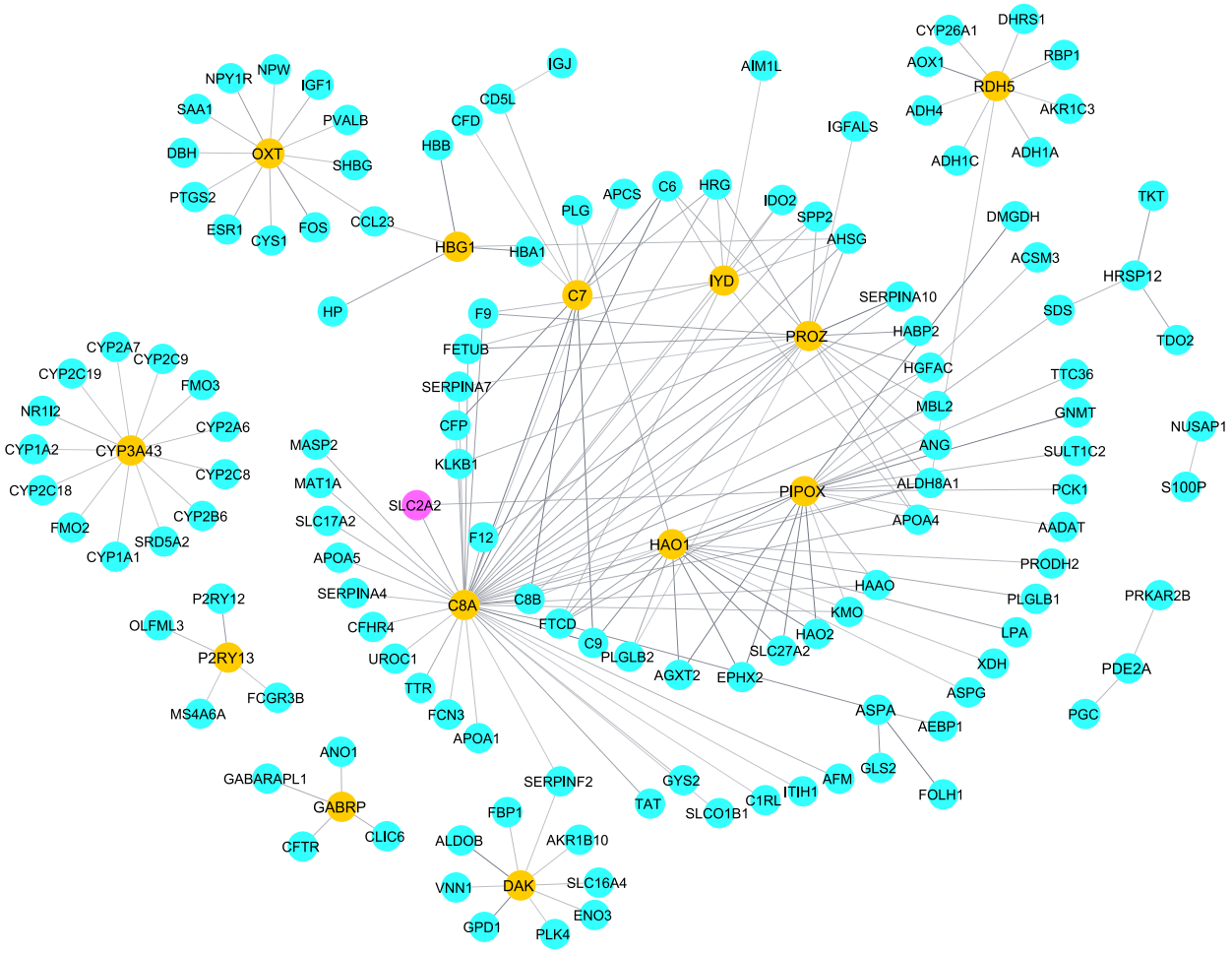

Fig. 4. Subnetwork derived by searching viable targets on the "parent" network. Genes in orange are considered important genes in HCC. *SLC2A2* gene is a major hub gene on the parent network.

the subsystems mentioned above (see Supplementary Table 2). The genes associated with these reactions were identified from the gene rules and searched among the SDE genes from the meta-analysis. 7 genes were identified to be common between the two analyses. Six genes, namely *ACSM3, ACSL1, ACSL4, SLC27A2, ACSM5, ACSM2A*, were involved in fatty acid activation and one gene, *COX7B2*, was involved in oxidative phosphorylation. Of the six genes involved in fatty acid activation, *ACSL4* (Acyl-CoA Synthetase Long Chain Family Member 4) was upregulated, while the others were downregulated. *COX7B2* (Cytochrome C Oxidase Subunit 7B2), involved in oxidative phosphorylation was upregulated. The 7 genes were mapped to the parent network to extract a subnetwork, Fig. 7. *HAO1, PIPOX* genes were common to the two subnetworks.

Subsystem enrichment was done per patient similar to reaction enrichment analysis, but here we considered all subsystems in the human reference model. 18 subsystems were significantly enriched (Fig. 8) and included subsystems involved in oxidation and activation of fatty acids and cholesterol biosynthesis, among others.

We assessed the prognostic potential of genes identified as relevant in HCC (Table 2). All selected genes (except *COX7B2* and *ACSL4*) showed prognostic potential at log rank p-value < 0.05 (see Supplementary Figure 6). Furthermore, we used the 17 selected genes to develop a classifier for tumor vs nontumor HCC samples. Using the random forest algorithm, transcriptomics data for the 17 genes effectively classified tumor and nontumor samples in HCC (median MCE: 3.4%, median AUC: 98%, see Supplementary Figure 7).

4. Discussion

We aimed to unravel the molecular markers and mechanisms that drive HCC in humans by comparing transcriptomics data from tumor and nontumor tissues. A total of 286 patients from eight datasets were included in the meta-analysis. Most of the samples, 79%, were from male patients. HCC disproportionately affects males more than females [1] but the factors responsible for this trajectory are not well understood and were not the focus of this study. Infection with hepatitis B or C viruses is known to be a major risk factor in HCC [59,60]. 72% of the patients included in this analysis had chronic hepatitis B or C infection. The patients were from populations in Europe, Asia, and North America.

We performed a differential gene expression analysis on each of the datasets in the limma in the R software. We did not consider confounders such as age and gender because the public datasets used were not complete in this regard. Furthermore, we did not see a separation of tumor tissue by gender (see Fig. 2) thus calming our concerns. Differential gene expression analysis was followed by meta-analysis of the results using random-effects models (REM). We used REM because it accounts for heterogeneity since the samples were from different populations and gene expression analysis was carried out on different platforms. Furthermore, the conclusions from REM can be extended beyond the samples used, that is, REM gives an unconditional inference about a larger set of studies from which the studies used are

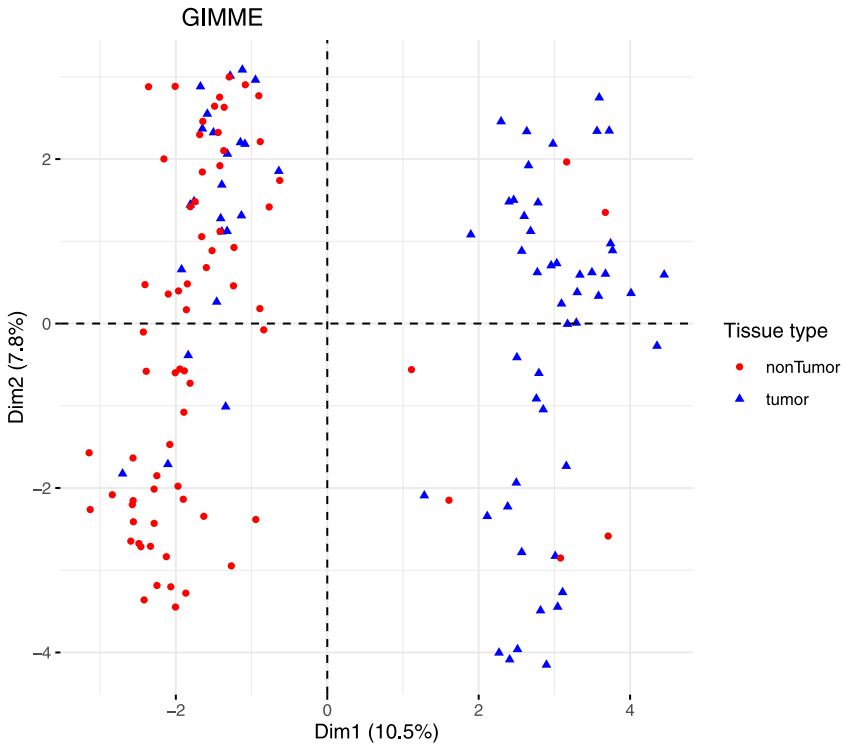

Fig. 5. PCA on models extracted using GIMME MEM. 70% of tumor and 92% of nontumor models separated well according to tissue type.

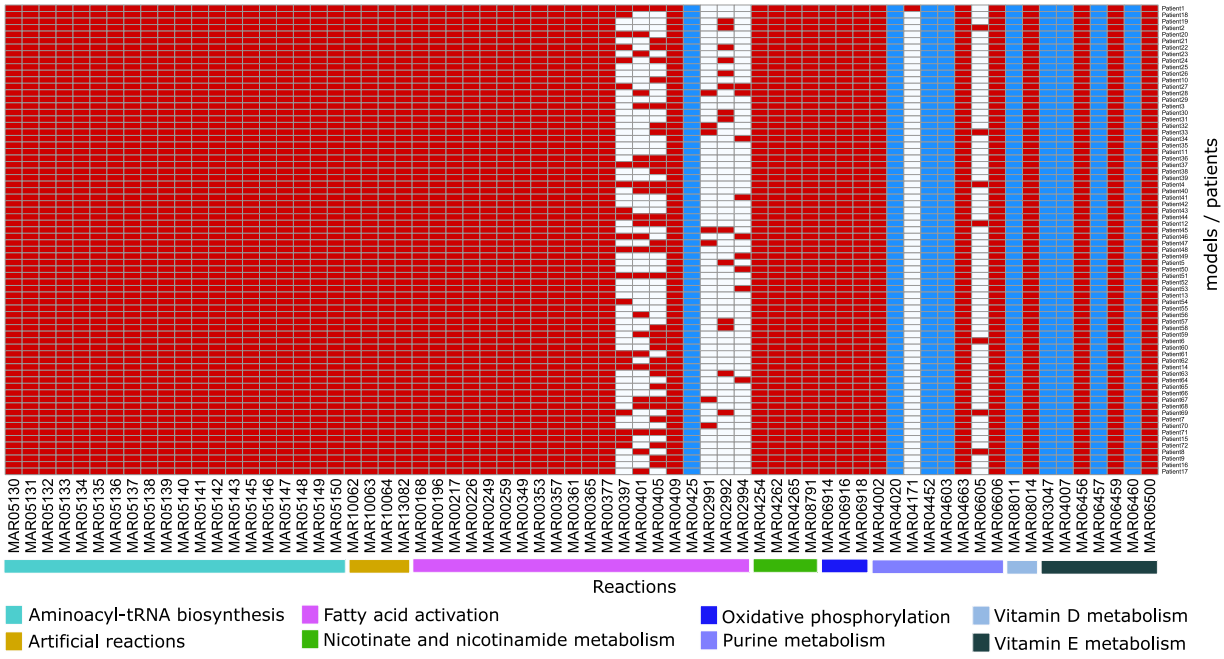

Fig. 6. Significantly enriched reactions at 5% level of significance, and a 10-fold flux change. Reactions in red are up regulated, while those in blue are down regulated.

assumed to be a random sample [12]. Genes were considered significantly differentially expressed (SDE) if their |mean $\log_2$FC| >= 1 and unadjusted *p*-value was less than 0.05. We cautiously used unadjusted *p*-value because of its lower stringency, which allowed us to remain with an appropriate number of SDE genes for downstream analyses. Enrichment analysis exposed key pathways known to be involved in cancer, for example, the cell cycle and p53 signaling pathways. Uncontrolled cell division is a known characteristic of cancers [61]. p53 signaling pathway is known to control the cell cycle, among other functions in the body [62]. Several cancers have been associated with the down regulation of genes in p53 pathway leading to uncontrolled cell division and hence these genes have been considered potential drug targets [62,63]. 16 of the enriched pathways, including fatty acid degradation, are associated with the metabolism of fatty acids, drugs, xenobiotics, glucose, among others. Unraveling the metabolic characteristics of cancer cells remains an unresolved challenge, because tumors dynamically adapt their metabolic characteristics at each step during metastasis, leading to the utilization of unconventional sources of energy [64,65]. Fatty acid degradation is an important pathway that generates acetyl-CoA, the entry molecule for the citric acid cycle that

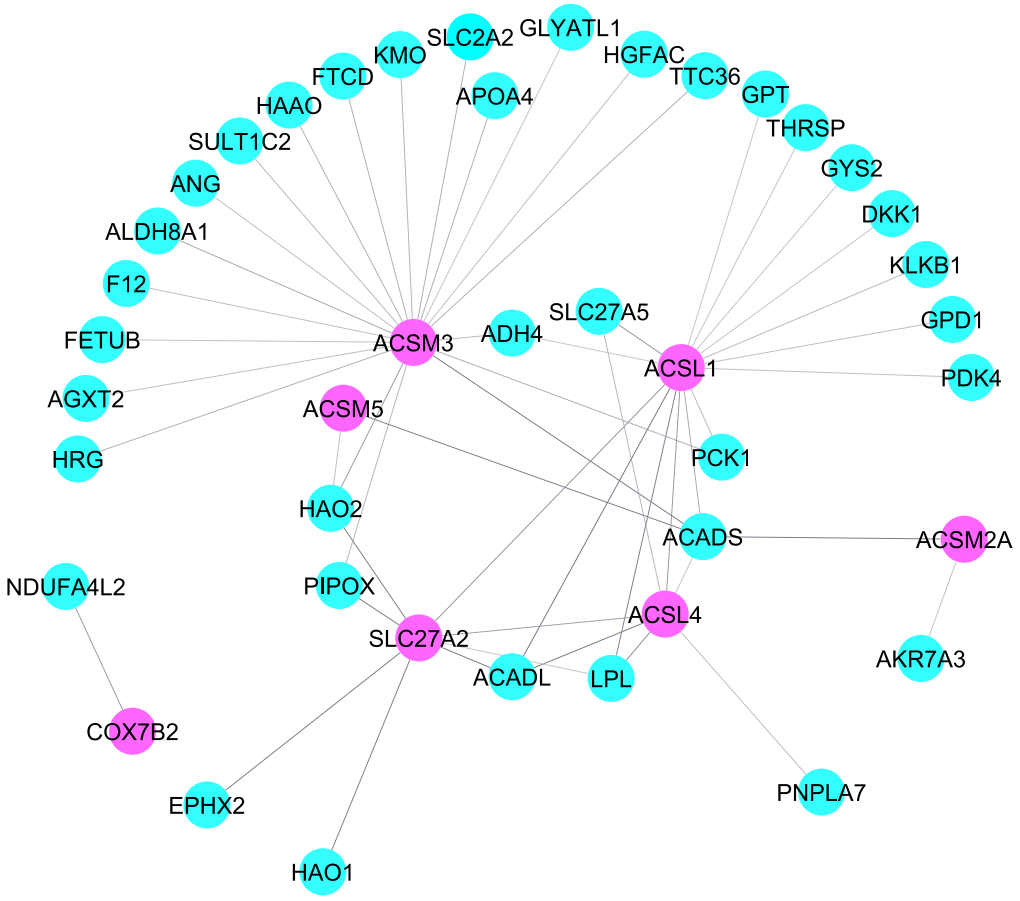

Fig. 7. *ACSM3, ACSL1, ACSL4, SLC27A2, ACSM5, ACSM2A* and *COX7B2* were used to extract a subnetwork from the parent network.

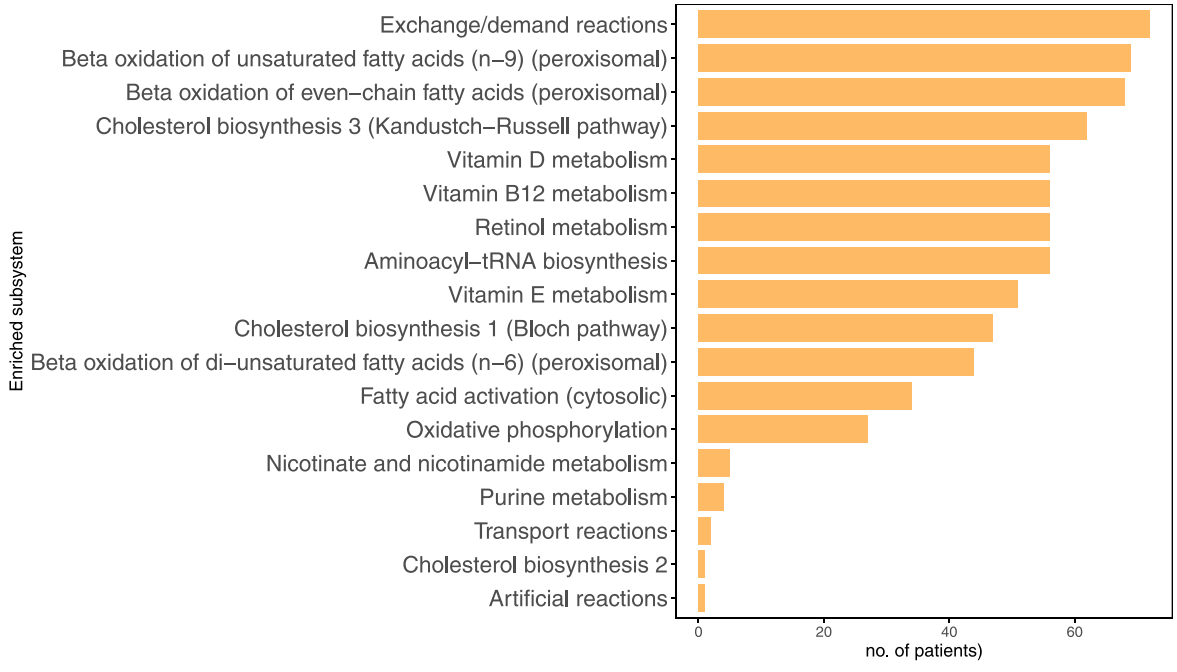

Fig. 8. Significantly enriched subsystems at 5% level of significance.

generates adenosine triphosphate (ATP), the biological form of energy. Enrichment of this pathway indicates a high demand for energy in tumor tissue.

In Table 2, we summarize the genes found in our analysis to be important in HCC. These genes were found to have prognostic potential in HCC (see Supplementary Figure 6) and effectively classified tumor

Table 2
Genes identified as important in HCC and are recommended for further studies.

Gene	Gene full name	Association with HCC	KEGG pathway
ACSL1 & *ACSL4*	Acyl-CoA Synthetase Long Chain Family Member 1 & Acyl-CoA Synthetase Long Chain Family Member 4	These genes encode Long chain acyl-CoA synthetase enzymes and have been associated with HCC. *ACSL1* gene was identified as a member of a prognostic signature of genes in HCC [36] and its disruption with aspirin suppressed abnormal lipid metabolism in liver cancer cells [37]. *ACSL4* gene is known to be upregulated in HCC samples [38].	Fatty acid biosynthesis, Fatty acid metabolism, Fatty acid degradation
ACSM3	Acyl-CoA Synthetase Medium Chain Family Member 3	This gene encodes a subunit of CoA ligases. Downregulation of this gene has been implicated in HCC and is associated with poor prognosis [39]	Butanoate metabolism, Metabolic pathways
C8A & C7	Complement C8 Alpha Chain & Complement C7	These genes are involved in formation of the membrane attack complex that plays a key role in innate and adaptive immunity. Downregulation of these genes impairs the immune system thus enabling tumor growth. One study reported that down regulation of the *C7* gene was associated with advanced cancer stages and high tumor grades in HCC [40].	Complement and coagulation cascades
COX7B2	Cytochrome C Oxidase Subunit 7B2	There is little information about its role in HCC but has been identified as a potential prognostic marker in combination with other genes [41]	Metabolic pathways, Oxidative phosphorylation
CYP3A43	Cytochrome P450 Family 3 Subfamily A Member 43	*CYP3A43* gene encodes enzymes of the cytochrome P450 superfamily which catalyze reactions involved in drug metabolism, and synthesis of steroids, lipids and cholesterol. Low expression of CYP3A43 gene was associated with reduced survival among HCC patients [42].	Chemical carcinogenesis - DNA adducts
DAK/TKFC	Dihydroxyacetone kinase	*DAK/TKFC* gene encodes enzymes which belong to the family of dihydroxyacetone kinases which catalyze formation of riboflavin 4',5'-phosphate. Expression of this gene was reduced in advanced compared to mild stages of fibrosis in hepatitis B patients [43].	RIG-I-like receptor signaling pathway
GABRP	Gamma-Aminobutyric Acid Type A Receptor Subunit Pi	this gene has been associated with ovarian [44], pancreatic [45] and breast cancer [46]. We did not find any studies that robustly associated this gene with HCC. Among genes that encode GABA, the gamma-aminobutyric acid A receptor θ subunit *GABRQ* has been associated with HCC [47]	GABAergic synapse
HAO1	Hydroxyacid oxidase 1	*HAO1* gene is primarily expressed in the liver and pancreas, and it encodes an enzyme that catalyzes the oxidation of glycolate into glycine which is an important step in the detoxification of glyoxylate. Glyoxylate reductase enzyme has previously been reported to be a potent prognostic marker in HCC [48,49]	Peroxisome, Metabolic pathways, Glyoxylate and dicarboxylate metabolism
HBG1	hemoglobin subunit gamma 1	There is no information we came across that robustly associates this gene to HCC. However, increase in fetal hemoglobin has been observed in many cancers [50,50]	–
IYD	Iodotyrosine Deiodinase	*IYD* gene encodes iodotyrosine deiodinase enzyme. Over expression of this gene in HCC cells reduced their utilization of glucose thus exhibiting a tumor suppressive effect [51].	Thyroid hormone synthesis
OXT	Oxytocin	*OXT* gene encodes oxytocin, a hormone that contracts smooth muscles. Alterations in this gene were associated with reduced survival of HCC patients [52]	Oxytocin signaling pathway
P2RY13	purinergic receptor P2Y, G-protein coupled, 13	encodes a receptor which is involved in signaling. Was identified as one of the genes in an HCC prognostic gene signature [53]	Neuroactive ligand–receptor interaction
PIPOX	Pipecolic Acid And Sarcosine Oxidase	*PIPOX* gene encodes an enzyme which is involved in the catabolic process to acetyl-CoA via L-pipecolate. Its downregulation is associated with lung and breast cancer [54] and prostate cancer [55]. It has also been identified as a potential prognostic marker in HCC [48]	Metabolic pathways, Peroxisome, Glycine, serine and threonine metabolism
PROZ	Protein Z, Vitamin K Dependent Plasma Glycoprotein	*PROZ* gene encodes a vitamin -K-dependent glycoprotein produced in the liver, and was found to be a biomarker in HCC [56]. Reduced expression of this gene in HCC has been suggested as a prognostic marker for early HCC [57].	–
RDH5	Retinol Dehydrogenase 5	*RDH5* gene encodes retinol dehydrogenase enzyme which catalyzes the biosynthesis of retinaldehyde used in vision. A study showed that increased expression of RDH5 gene suppressed metastasis in HCC cell lines [58].	Retinol metabolism, Metabolic pathways

Genes identified to be important in HCC. *ACSL1, ACSM3, C7, C8A, CYP3A43, DAK (TKFC), GABRP, HAO1, HBG1, IYD, OXT, P2RY13, PIPOX, PROZ* and *RDH5* were downregulated and *ACSL4, COX7B2* were upregulated in the meta-analysis. *ACSL1, ACSL4, COX7B2, CYP3A43, HAO1* and *RDH5* genes are being targeted in the treatment of other ailments thus are worth studying as viable drug targets in HCC.

and nontumor samples in an independent dataset (see Supplementary Figure 7). Genes in this list are associated with KEGG metabolic pathways linked to HCC such as fatty acid metabolism (including biosynthesis and degradation), complement and coagulation cascade, chemical carcinogenesis and retinol metabolism, which additionally confirms their importance in this disease. Majority of these genes were already linked to HCC pathology either directly or through in vitro assays and this validates our analysis approach. To our knowledge, three genes, *DAK (TKFC)*, *GABRP* and *HBG1* were not previously associated with HCC. *DAK (TKFC)* gene encodes dihydroxyacetone kinase enzyme that catalyzes the formation of riboflavin cyclic 4,5-phosphate and was downregulated in the meta-analysis. This gene has been reported to be associated with liver fibrosis in rats [66] and liver dysfunction in the multi-disease state in humans [67]. *DAK (TKFC)* is involved in the regulation of innate antiviral immunity by suppressing melanoma differentiation-associated gene-5 *(MDA5)* and its deficiency results in impaired innate immunity in response to viruses [68]. It is important to explore the role of *DAK (TKFC)* in HCC and to determine whether differences in its expression are strictly associated with viral-related HCC, as this would determine its application as a biomarker. *HBG1* encodes the gamma subunit 1 of fetal hemoglobin and was downregulated in the meta-analysis. The γ-globin gene has previously been reported to be hypomethylated in colon and breast cancer [69]. Furthermore, older studies (published before 2000) observed an increase in fetal hemoglobin in some cancers [50,70] but it is not clear if this is due to malignancies or treatment. Therefore, it is important to explore the role of fetal hemoglobin in HCC because it is easy to measure in blood and thus a good candidate biomarker. *GABRP* (Gamma-Aminobutyric Acid Type A Receptor Subunit Pi) has previously not been linked to HCC pathology, but is reported to induce growth of pancreatic cancer cell lines through calcium mobilization and ERK1/2 signaling [71]. It has also been associated with ovarian [44], pancreatic [45] and breast cancer [46] in humans. A related gene, the gamma-aminobutyric acid A receptor θ subunit *GABRQ* has been associated with HCC [47]. Therefore, it is worth exploring the role of the *GABRP* gene in HCC.

Transcriptomics data, GSE39791, were integrated into the human reference GEM, Human-GEM. Models were analyzed per patient, i.e., comparing flux samples of models from tumor and nontumor tissue for each patient. We opted for a personalized approach because individuals are biologically distinct and tumors are highly heterogeneous; therefore, some reactions and subsystems may be enriched in one individual but not in the other. Analysis of the enriched reactions revealed that the biosynthesis of aminoacyl-tRNA was upregulated in tumor tissues, indicating increased translation of RNA to peptides (proteins). This is expected in cancer since the biomass accumulation of tumors is higher than that of normal tissues and thus require increased expression of proteins for their growth. Oxidative phosphorylation reaction (MAR06916) yielding ATP and reactions involved in nicotinamide metabolism were upregulated indicating the high energy requirements in tumor compared to nontumor tissue. The gene *COX7B2*, involved in the reaction MAR06914, which is one of the upregulated reactions of oxidative phosphorylation, was upregulated in the meta-analysis. Together, these results confirm the high energy consumption in tumor tissues.

Generally, purine production was upregulated in tumor tissue, and this supports tumor growth. Purines are known to be the building blocks of DNA and RNA and also promote cell proliferation and survival when they function as energy molecules and cofactors [72,73]. Vitamin D metabolism was enriched in tumors and therefore may play an important role in HCC. In fact, vitamin D is known to reduce the risk of many types of cancer [74]. Production of vitamin D starts with conversion of provitamin D_3 (7-dehydrocholesterol) to previtamin D_3 which is finally converted to vitamin D_3. The conversion of provitamin D_3 to previtamin D_3 was downregulated in tumors. This may imply a reduction in vitamin D_3 production in tumors. However, the upregulation of conversion of previtamin to vitamin D_3 is an attempt by the body to produce it since its needed for key functions such as maintaining calcium and phosphate levels and regulation of cell proliferation and differentiation hence whatever little previtamin D_3 available is quickly converted to vitamin D_3. Furthermore, vitamin D is known to be a key regulator of the innate and adaptive immune system hence its deficiency creates a weakened immune system thus a conducive environment for tumor growth. Vitamin E metabolism was generally downregulated in tumor tissues (4 of 7 reactions were downregulated). Vitamin E is a major antioxidant, yet reactive oxygen species have been implicated in the cause of cancers through oxidative stress [75,76] thus downregulating the reactions that process vitamin E favoring tumor growth.

The 20 reactions involved in fatty acid activation (except MAR00425) were upregulated in tumor tissue. Fatty acid activation is the first step in the utilization of fatty acids and produces fatty acyl-CoA which is then entered in the second stage where it is converted into ATP via β-oxidation (catabolic metabolism) or used as building blocks for triacylglycerol, phospholipids and cholesterol esters (anabolic metabolism). There is a growing body of evidence showing an increased role of fatty acids in tumor proliferation [38,77]. 3 of the 18 significantly enriched subsystems are involved in β-oxidation of fatty acids (Fig. 8). Furthermore, all upregulated fatty acid activation reactions yield ATP, the chemical form of energy in cells. *ACSL4*, a gene that encodes an enzyme that catalyzes fatty acid activation reactions, was upregulated in our meta-analysis. This gene is also involved in metabolism of eicosanoids and leukotrienes, both of which are mediators of inflammation, a major characteristic of HCC and its prestages. Dysregulation of the *ACSL4* gene has previously been associated with tumor proliferation, invasion, and evasion of programmed cell death [38]. Another gene, *ACSM3*, which is involved in a few fatty acid activation reactions in the cytosol was down regulated in our meta-analysis. This result is in agreement with a previous experimental study [39] in which downregulation of *ACSM3* was associated with increased metastasis and a poor prognosis in HCC. This gene is mainly active in the mitochondria and thus plays a minor role in the activation of fatty acids in the cytosol, but could be involved in other biological processes relevant to HCC. In all, the results show that fatty acids are important for tumor growth and are mainly used by tumor cells as a source of energy.

The limitation of GEM analysis is that it is still a young field; hence, the reference models are not fully developed, e.g., the Human-GEM used in this analysis contained only 3625 genes, which is a small fraction of more than 20,000 genes known to be active in humans. This limits their application. However, several groups have shown that GEMs are a very useful computational tool to gain insight into systems biology [78,79]. In the present manuscript, we adapted Human-GEM to specific contexts using transcriptomics data and different MEMs. In the future, we plan to extend this analysis by additionally integrating proteomics and metabolomics studies of HCC. Analysis using GEMs can be used to gain insight into potentially useful proteins and metabolites in a disease condition. In this context, GEM-based analyses could be potentially combined with the protocols and tools, such as POSREG [80], NOREVA [81–83], ANPELA [84], MetaFS [85], LargeMetabo [86], and MMEASE [87].

5. Conclusion

We used a robust integrated analysis framework to show that fatty acid activation is a critical process in the development of HCC and that this process is a major source of energy (ATP) for tumors. We conclude that fatty acid activation can be targeted for the development of new therapies. Vitamin D and E supplementation may provide some relief, but it is not clear to what extent and therefore needs to be explored further. Upregulation of *ACSL1, ACSM3, C7, C8 A, CYP3A43, DAK (TKFC), GABRP, HAO1, HBG1, IYD, OXT, P2RY13, PIPOX, PROZ* and *RDH5* and downregulation of *ACSL4, COX7B2* may slow tumor progression. In our study, we provide the first evidence that *GABRP*, *HBG1* and *DAK (TKFC)* genes are important in HCC in humans.

CRediT authorship contribution statement

Andrew Walakira: Methodology, Software, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Cene Skubic:** Investigation, Writing – review & editing. **Nejc Nadižar:** Investigation, Software. **Damjana Rozman:** Conceptualization, Writing – review & editing, Funding acquisition, Supervision. **Tadeja Režen:** Conceptualization, Writing – review & editing. **Miha Mraz:** Writing – review & editing, Funding acquisition. **Miha Moškon:** Conceptualization, Methodology, Formal analysis, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Horizon 2020 TranSYS Marie Curie Initial Training Network Grant Agreement No. 860895, by the scientific-research program Pervasive Computing (P2-0359) founded by the Slovenian Research Agency, by the scientific-research program Functional Genomics and Biotechnology for Health, Slovenia (P1-0390) founded by the Slovenian Research Agency, and by the Network of infrastructure Centres of University of Ljubljana, Slovenia (MRIC-UL-CFGBC, IP-022) founded by the Slovenian Research Agency. We would also like to acknowledge the support by the infrastructure project ELIXIR-SI RI-SI-2 financed by the European Regional Development Fund and by the Ministry of Education, Science and Sport of Republic of Slovenia. The funding sources had no role in the design of the study and collection, analysis, and interpretation of data nor in writing the manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.compbiomed.2023.106957.

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA: Cancer J. Clin. 71 (3) (2021) 209–249.

[2] J.J. Ott, J. Horn, G. Krause, R.T. Mikolajczyk, Time trends of chronic HBV infection over prior decades–a global analysis, J. Hepatol. 66 (1) (2017) 48–54.

[3] J. Zhou, H. Sun, Z. Wang, W. Cong, J. Wang, M. Zeng, W. Zhou, P. Bie, L. Liu, T. Wen, et al., Guidelines for the diagnosis and treatment of hepatocellular carcinoma (2019 Edition), Liver Cancer 9 (6) (2020) 682–720.

[4] J. Zhang, G. Chen, P. Zhang, J. Zhang, X. Li, D. Gan, X. Cao, M. Han, H. Du, Y. Ye, The threshold of alpha-fetoprotein (AFP) for the diagnosis of hepatocellular carcinoma: A systematic review and meta-analysis, PLoS One 15 (2) (2020) e0228857.

[5] E. Degasperi, R. Perbellini, R. D'Ambrosio, S.C. Uceda Renteria, F. Ceriotti, A. Perego, C. Orsini, M. Borghi, M. Iavarone, M. Bruccoleri, et al., Prothrombin induced by vitamin K absence or antagonist-II and alpha foetoprotein to predict development of hepatocellular carcinoma in Caucasian patients with hepatitis C-related cirrhosis treated with direct-acting antiviral agents, Aliment. Pharmacol. Ther. 55 (3) (2022) 350–359.

[6] N.B. Quoc, N.D. Phuong, T.K. Ngan, N.T.M. Linh, P.H. Cuong, N.N.B. Chau, Expression of plasma hsa-miR122 in HBV-related hepatocellular carcinoma (HCC) in Vietnamese patients, MicroRNA 7 (2) (2018) 92–99.

[7] H. Kaur, S. Bhalla, D. Kaur, G.P. Raghava, CancerLivER: a database of liver cancer gene expression resources and biomarkers, Database 2020 (2020).

[8] Q. Lian, S. Wang, G. Zhang, D. Wang, G. Luo, J. Tang, L. Chen, J. Gu, HCCDB: a database of hepatocellular carcinoma expression atlas, Genomics Proteomics Bioinform. 16 (4) (2018) 269–275.

[9] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015) e47.

[10] R Core Team, R: A Language and Environment for Statistical Computing, 2019, URL https://www.R-project.org/.

[11] W. Viechtbauer, M.W. Viechtbauer, Package 'metafor', in: The Comprehensive R Archive Network. Package 'Metafor', 2015, http://cran.r-project.org/web/packages/metafor/metafor.pdf.

[12] W. Viechtbauer, Conducting meta-analyses in R with the metafor package, J. Stat. Softw. 36 (3) (2010) 1–48.

[13] T. Wu, E. Hu, S. Xu, M. Chen, P. Guo, Z. Dai, T. Feng, L. Zhou, W. Tang, L. Zhan, et al., clusterProfiler 4.0: A universal enrichment tool for interpreting omics data, Innovation 2 (3) (2021) 100141.

[14] J. Piñero, J.M. Ramírez-Anguita, J. Saüch-Pitarch, F. Ronzano, E. Centeno, F. Sanz, L.I. Furlong, The DisGeNET knowledge platform for disease genomics: 2019 update, Nucleic Acids Res. 48 (D1) (2020) D845–D855.

[15] T. Ohkuma, S.A. Peters, M. Woodward, Sex differences in the association between diabetes and cancer: a systematic review and meta-analysis of 121 cohorts including 20 million individuals and one million events, Diabetologia 61 (10) (2018) 2140–2154.

[16] S.M. Corsello, J.A. Bittker, Z. Liu, J. Gould, P. McCarren, J.E. Hirschman, S.E. Johnston, A. Vrcic, B. Wong, M. Khan, et al., The Drug Repurposing Hub: a next-generation drug library and information resource, Nat. Med. 23 (4) (2017) 405–408.

[17] D. Szklarczyk, A.L. Gable, K.C. Nastou, D. Lyon, R. Kirsch, S. Pyysalo, N.T. Doncheva, M. Legeay, T. Fang, P. Bork, et al., The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets, Nucleic Acids Res. 49 (D1) (2021) D605–D612.

[18] L. Heirendt, S. Arreckx, T. Pfau, S.N. Mendoza, A. Richelle, A. Heinken, H.S. Haraldsdóttir, J. Wachowiak, S.M. Keating, V. Vlasov, et al., Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v. 3.0, Nat. Protoc. 14 (3) (2019) 639–702.

[19] H. Wang, S. Marcišauskas, B.J. Sánchez, I. Domenzain, D. Hermansson, R. Agren, J. Nielsen, E.J. Kerkhoven, RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on Streptomyces coelicolor, PLoS Comput. Biol. 14 (10) (2018) e1006541.

[20] L. Gurobi Optimization, Gurobi optimizer reference manual, 2021, URL http://www.gurobi.com.

[21] A. Walakira, D. Rozman, T. Režen, M. Mraz, M. Moškon, Guided extraction of genome-scale metabolic models for the integration and analysis of omics data, Comput. Struct. Biotechnol. J. 19 (2021) 3521–3530.

[22] J.L. Robinson, P. Kocabaş, H. Wang, P.-E. Cholley, D. Cook, A. Nilsson, M. Anton, R. Ferreira, I. Domenzain, V. Billa, et al., An atlas of human metabolism, Sci. Signal. 13 (624) (2020) eaaz1482.

[23] M. Moškon, T. Režen, Context-specific genome-scale metabolic modelling and its application to the analysis of COVID-19 metabolic signatures, Metabolites 13 (1) (2023) 126.

[24] S.A. Becker, B.O. Palsson, Context-specific metabolic networks are consistent with experiments, PLoS Comput. Biol. 4 (5) (2008) e1000082.

[25] S. Opdam, A. Richelle, B. Kellman, S. Li, D.C. Zielinski, N.E. Lewis, A systematic evaluation of methods for tailoring genome-scale metabolic models, Cell Syst. 4 (3) (2017) 318–329.

[26] H. Zur, E. Ruppin, T. Shlomi, iMAT: an integrative metabolic analysis tool, Bioinformatics 26 (24) (2010) 3140–3142.

[27] R. Agren, S. Bordel, A. Mardinoglu, N. Pornputtapong, I. Nookaew, J. Nielsen, Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT, PLoS Comput. Biol. 8 (5) (2012) e1002518.

[28] R. Agren, A. Mardinoglu, A. Asplund, C. Kampf, M. Uhlen, J. Nielsen, Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling, Mol. Syst. Biol. 10 (3) (2014) 721.

[29] L. Van Der Maaten, Accelerating t-SNE using tree-based algorithms, J. Mach. Learn. Res. 15 (1) (2014) 3221–3245.

[30] L. Van der Maaten, G. Hinton, Visualizing data using t-SNE, J. Mach. Learn. Res. 9 (11) (2008).

[31] S. Fallahi, H.J. Skaug, G. Alendal, A comparison of Monte Carlo sampling methods for metabolic network models, PLoS One 15 (7) (2020) e0235393.

[32] P. Nanda, A. Ghosh, Genome Scale-Differential Flux Analysis reveals deregulation of lung cell metabolism on SARS-CoV-2 infection, PLoS Comput. Biol. 17 (4) (2021) e1008860.

[33] A. Lánczky, B. Győrffy, et al., Web-based survival analysis tool tailored for medical research (KMplot): development and implementation, J. Med. Internet Res. 23 (7) (2021) e27633.

[34] O. Menyhárt, Á. Nagy, B. Győrffy, Determining consistent prognostic biomarkers of overall survival and vascular invasion in hepatocellular carcinoma, R. Soc. Open Sci. 5 (12) (2018) 181006.

[35] M. Slawski, A. Boulesteix, C. Bernau, CMA: Synthesis of microarray-based classification, 2009, R Package Version. 1.

[36] C. Yue, Y. Ren, H. Ge, C. Liang, Y. Xu, G. Li, J. Wu, Comprehensive analysis of potential prognostic genes for the construction of a competing endogenous RNA regulatory network in hepatocellular carcinoma, OncoTargets Ther. 12 (2019) 561.

[37] G. Yang, Y. Wang, J. Feng, Y. Liu, T. Wang, M. Zhao, L. Ye, X. Zhang, Aspirin suppresses the abnormal lipid metabolism in liver cancer cells via disrupting an NFκB-ACSL1 signaling, Biochem. Biophys. Res. Commun. 486 (3) (2017) 827–832.

[38] Y. Tang, J. Zhou, S.C. Hooi, Y.-M. Jiang, G.-D. Lu, Fatty acid activation in carcinogenesis and cancer development: Essential roles of long-chain acyl-CoA synthetases, Oncol. Lett. 16 (2) (2018) 1390–1396.

[39] H.-Y. Ruan, C. Yang, X.-M. Tao, J. He, T. Wang, H. Wang, C. Wang, G.-Z. Jin, H.-J. Jin, W.-X. Qin, Downregulation of ACSM3 promotes metastasis and predicts poor prognosis in hepatocellular carcinoma, Am. J. Cancer Res. 7 (3) (2017) 543.

[40] X. Qian, Z. Yang, L. Gao, Y. Liu, J. Yan, The role of complement in the clinical course of hepatocellular carcinoma, Immun. Inflamm. Dis. 10 (3) (2022) e569.

[41] Z. Wang, K.S. Embaye, Q. Yang, L. Qin, C. Zhang, L. Liu, X. Zhan, F. Zhang, X. Wang, S. Qin, A novel metabolism-related signature as a candidate prognostic biomarker for hepatocellular carcinoma, J. Hepatocell. Carcinoma 8 (2021) 119.

[42] T. Yu, X. Wang, G. Zhu, C. Han, H. Su, X. Liao, C. Yang, W. Qin, K. Huang, T. Peng, The prognostic value of differentially expressed CYP3A subfamily members for hepatocellular carcinoma, Cancer Manag. Res. 10 (2018) 1713.

[43] M.-Y. Xu, X.-F. Jia, Y. Qu, R.-D. Zheng, Z.-H. Yuan, H.-L. Weng, S. Dooley, X.-P. Wang, L.-J. Zhang, L.-G. Lu, Serum dihydroxyacetone kinase peptide m/z 520.3 as predictor of disease severity in patients with compensated chronic hepatitis B, J. Transl. Med. 11 (1) (2013) 1–14.

[44] H.Y. Sung, S.-D. Yang, W. Ju, J.-H. Ahn, Aberrant epigenetic regulation of GABRP associates with aggressive phenotype of ovarian cancer, Exp. Mol. Med. 49 (5) (2017) e335.

[45] Y. Yang, L. Ren, S. Li, X. Zheng, J. Liu, W. Li, W. Fu, J. Wang, G. Du, GABRP is a potential prognostic biomarker and correlated with immune infiltration and tumor microenvironment in pancreatic cancer, Transl. Cancer Res. 11 (4) (2022) 649.

[46] G. Azizi-Tabesh, Z. Kamaliyan, F. Darbeheshti, R. Omranipour, V. Soleimani, N. Alipour, R. Mirfakhraie, V.R. Yassaee, Overexpression of GABRP gene in triple negative breast cancer: Molecular mechanisms and interpretation, Int. J. Cancer Manag. 14 (11) (2021).

[47] Y.-H. Li, Y. Liu, Y.-D. Li, Y.-H. Liu, F. Li, Q. Ju, P.-L. Xie, G.-C. Li, GABA stimulates human hepatocellular carcinoma growth through overexpressed GABAA receptor theta subunit, World J. Gastroenterol.: WJG 18 (21) (2012) 2704.

[48] M. Desoteux, C. Louis, K. Bévant, D. Glaise, C. Coulouarn, A minimal subset of seven genes associated with tumor hepatocyte differentiation predicts a poor prognosis in human hepatocellular carcinoma, Cancers 13 (22) (2021) 5624.

[49] Y. Pan, R. Ni, Q. Deng, X. Huang, Y. Zhang, C. Lu, F. Li, D. Huang, S. He, B. Chen, Glyoxylate reductase/hydroxypyruvate reductase: a novel prognostic marker for hepatocellular carcinoma patients after curative resection, Pathobiology 80 (3) (2013) 155–162.

[50] M. Wolk, M. Kieselstein, C.G. Ben-Dor, G. Brufman, Fetal hemoglobin screening in whole blood and in plasma of cancer patients, Tumor Biol. 12 (1) (1991) 45–51.

[51] C. Lu, S. Fang, Q. Weng, X. Lv, M. Meng, J. Zhu, L. Zheng, Y. Hu, Y. Gao, X. Wu, et al., Integrated analysis reveals critical glycolytic regulators in hepatocellular carcinoma, Cell Commun. Signal. 18 (1) (2020) 1–14.

[52] T. Harricharran, O.O. Ogunwobi, Oxytocin receptor genetic alterations in hepatocellular carcinoma, SN Compr. Clin. Med. 1 (7) (2019) 523–526.

[53] O. Al-Harazi, I.H. Kaya, M. Al-Eid, L. Alfantoukh, A.S. Al Zahrani, M. Al Sebayel, N. Kaya, D. Colak, Identification of gene signature as diagnostic and prognostic blood biomarker for early hepatocellular carcinoma using integrated cross-species transcriptomic and network analyses, Front. Genet. (2021) 1854.

[54] Y.J. Cha, W.H. Jung, J.S. Koo, et al., Expression of sarcosine metabolism-related proteins according to metastatic site in breast cancer, Int. J. Clin. Exp. Pathol. 7 (11) (2014) 7824.

[55] A.P. Khan, T.M. Rajendiran, A. Bushra, I.A. Asangani, J.N. Athanikar, A.K. Yocum, R. Mehra, J. Siddiqui, G. Palapattu, J.T. Wei, et al., The role of sarcosine metabolism in prostate cancer progression, Neoplasia 15 (5) (2013) 491–IN13.

[56] T.-H. Tsai, E. Song, R. Zhu, C. Di Poto, M. Wang, Y. Luo, R.S. Varghese, M.G. Tadesse, D.H. Ziada, C.S. Desai, et al., LC-MS/MS-based serum proteomics for identification of candidate biomarkers for hepatocellular carcinoma, Proteomics 15 (13) (2015) 2369–2381.

[57] X. Jiang, T. Song, X. Pan, X. Zhang, Y. Lan, L. Bai, PROZ may serve as a prognostic biomarker for early hepatocellular carcinoma, Int. J. Gen. Med. 14 (2021) 4209.

[58] H. Hu, L. Xu, S.-J. Luo, T. Xiang, Y. Chen, Z.-r. Cao, Y.-j. Zhang, Z. Mo, Y. Wang, D.-f. Meng, et al., Retinal dehydrogenase 5 (RHD5) attenuates metastasis via regulating HIPPO/YAP signaling pathway in hepatocellular carcinoma, Int. J. Med. Sci. 17 (13) (2020) 1897.

[59] I. Chemin, F. Zoulim, Hepatitis B virus induced hepatocellular carcinoma, Cancer Lett. 286 (1) (2009) 52–59.

[60] A. Baecker, X. Liu, C. La Vecchia, Z.-F. Zhang, Worldwide incident hepatocellular carcinoma cases attributable to major risk factors, Eur. J. Cancer Prev. Official J. Eur. Cancer Prev. Organ. (ECP) 27 (3) (2018) 205.

[61] H.K. Matthews, C. Bertoli, R.A. de Bruin, Cell cycle control in cancer, Nat. Rev. Mol. Cell Biol. 23 (1) (2022) 74–88.

[62] A.H. Stegh, Targeting the p53 signaling pathway in cancer therapy–the promises, challenges and perils, Expert Opin. Ther. Targets 16 (1) (2012) 67–83.

[63] L. Stramucci, A. Pranteda, G. Bossi, Insights of crosstalk between p53 protein and the MKK3/MKK6/p38 MAPK signaling pathway in cancer, Cancers 10 (5) (2018) 131.

[64] G. Bergers, S.-M. Fendt, The metabolism of cancer cells during metastasis, Nat. Rev. Cancer 21 (3) (2021) 162–180.

[65] L. Sun, C. Suo, S.-t. Li, H. Zhang, P. Gao, Metabolic reprogramming for cancer cells and their microenvironment: Beyond the Warburg Effect, Biochim. Biophys. Acta (BBA)-Rev. Cancer 1870 (1) (2018) 51–66.

[66] M.D.M. AbdulHameed, G.J. Tawa, K. Kumar, D.L. Ippolito, J.A. Lewis, J.D. Stallings, A. Wallqvist, Systems level analysis and identification of pathways and networks associated with liver fibrosis, PLoS One 9 (11) (2014) e112193.

[67] S.B. Wortmann, B. Meunier, L. Mestek-Boukhibar, F. van den Broek, E.M. Maldonado, E. Clement, D. Weghuber, J. Spenger, Z. Jaros, F. Taha, et al., Bi-allelic variants in TKFC encoding triokinase/FMN cyclase are associated with cataracts and multisystem disease, Am. J. Hum. Genet. 106 (2) (2020) 256–263.

[68] F. Diao, S. Li, Y. Tian, M. Zhang, L.-G. Xu, Y. Zhang, R.-P. Wang, D. Chen, Z. Zhai, B. Zhong, et al., Negative regulation of MDA5-but not RIG-I-mediated innate antiviral signaling by the dihydroxyacetone kinase, Proc. Natl. Acad. Sci. 104 (28) (2007) 11706–11711.

[69] S. Ribieras, X.-G. Song-Wang, V. Martin, P. Lointier, L. Frappart, R. Dante, Human breast and colon cancers exhibit alterations of DNA methylation patterns at several DNA segments on chromosomes 11p and 17p, J. Cell. Biochem. 56 (1) (1994) 86–96.

[70] B.I. Carr, S. Rahbar, J.H. Doroshow, D. Blayney, D. Goldberg, L. Leong, Y. Asmeron, Fetal hemoglobin gene activation in a phase II study of 5, 6-dihydro-5-azacytidine for bronchogenic carcinoma, Cancer Res. 47 (15) (1987) 4199–4201.

[71] G.M. Sizemore, S.T. Sizemore, D.D. Seachrist, R.A. Keri, GABA (A) receptor pi (GABRP) stimulates basal-like breast cancer cell migration through activation of extracellular-regulated kinase 1/2 (ERK1/2), J. Biol. Chem. 289 (35) (2014) 24102–24113.

[72] V. Pareek, A.M. Pedley, S.J. Benkovic, Human de novo purine biosynthesis, Crit. Rev. Biochem. Mol. Biol. 56 (1) (2021) 1–16.

[73] J. Yin, W. Ren, X. Huang, J. Deng, T. Li, Y. Yin, Potential mechanisms connecting purine metabolism and cancer therapy, Front. Immunol. 9 (2018) 1697.

[74] M. Moukayed, W.B. Grant, Molecular link between vitamin D and cancer prevention, Nutrients 5 (10) (2013) 3993–4021.

[75] J.E. Klaunig, Oxidative stress and cancer, Curr. Pharm. Des. 24 (40) (2018) 4771–4778.

[76] V. Aggarwal, H.S. Tuli, A. Varol, F. Thakral, M.B. Yerer, K. Sak, M. Varol, A. Jain, M. Khan, G. Sethi, et al., Role of reactive oxygen species in cancer progression: molecular mechanisms and recent advancements, Biomolecules 9 (11) (2019) 735.

[77] M. Wang, J. Han, H. Xing, H. Zhang, Z. Li, L. Liang, C. Li, S. Dai, M. Wu, F. Shen, et al., Dysregulated fatty acid metabolism in hepatocellular carcinoma, Hepat. Oncol. 3 (4) (2016) 241–251.

[78] T. Režen, A. Martins, M. Mraz, N. Zimic, D. Rozman, M. Moškon, Integration of omics data to generate and analyse COVID-19 specific genome-scale metabolic models, Comput. Biol. Med. 145 (2022) 105428.

[79] Y. Seif, E. Kavvas, J.-C. Lachance, J.T. Yurkovich, S.-P. Nuccio, X. Fang, E. Catoiu, M. Raffatellu, B.O. Palsson, J.M. Monk, Genome-scale metabolic reconstructions of multiple Salmonella strains reveal serovar-specific metabolic traits, Nature Commun. 9 (1) (2018) 1–12.

[80] F. Li, Y. Zhou, Y. Zhang, J. Yin, Y. Qiu, J. Gao, F. Zhu, POSREG: proteomic signature discovered by simultaneously optimizing its reproducibility and generalizability, Brief. Bioinform. 23 (2) (2022).

[81] J. Fu, Y. Zhang, Y. Wang, H. Zhang, J. Liu, J. Tang, Q. Yang, H. Sun, W. Qiu, Y. Ma, et al., Optimization of metabolomic data processing using NOREVA, Nat. Protoc. 17 (1) (2022) 129–151.

[82] Q. Yang, Y. Wang, Y. Zhang, F. Li, W. Xia, Y. Zhou, Y. Qiu, H. Li, F. Zhu, NOREVA: enhanced normalization and evaluation of time-course and multi-class metabolomic data, Nucleic Acids Res. 48 (W1) (2020) W436–W448.

[83] B. Li, J. Tang, Q. Yang, S. Li, X. Cui, Y. Li, Y. Chen, W. Xue, X. Li, F. Zhu, NOREVA: normalization and evaluation of MS-based metabolomics data, Nucleic Acids Res. 45 (W1) (2017) W162–W170.

[84] J. Tang, J. Fu, Y. Wang, B. Li, Y. Li, Q. Yang, X. Cui, J. Hong, X. Li, Y. Chen, et al., ANPELA: analysis and performance assessment of the label-free quantification workflow for metaproteomic studies, Brief. Bioinform. 21 (2) (2020) 621–636.

[85] J. Tang, M. Mou, Y. Wang, Y. Luo, F. Zhu, MetaFS: performance assessment of biomarker discovery in metaproteomics, Brief. Bioinform. 22 (3) (2021) bbaa105.

[86] Q. Yang, B. Li, P. Wang, J. Xie, Y. Feng, Z. Liu, F. Zhu, LargeMetabo: an out-of-the-box tool for processing and analyzing large-scale metabolomic data, Brief. Bioinform. 23 (6) (2022).

[87] Q. Yang, B. Li, S. Chen, J. Tang, Y. Li, Y. Li, S. Zhang, C. Shi, Y. Zhang, M. Mou, et al., MMEASE: online meta-analysis of metabolomic data by enhanced metabolite annotation, marker selection and enrichment analysis, J. Proteomics 232 (2021) 104023.