Research Article

Network Analysis in the Identification of Genes Conferring Metastatic Potential in Hepatocellular Carcinoma

Hao Dong Tan,[1] Hazel Jing Yi Leong,[1] Wei-Hsum Yap,[1,2] Adeline Yoke Yin Chia,[1] Yin-Quan Tang[1,2]

[1]School of Biosciences, Faculty of Health and Medical Sciences Taylor's University, Subang Jaya, Malaysia
[2]Medical Advancement for Better Quality of Life Impact Lab, Taylor's University, Subang Jaya, Selangor Darul Ehsan, Malaysia

Abstract

Objectives: Hepatocellular carcinoma (HCC) is a common liver cancer accounting with high mortality rate owing to metastasis. Anti-metastatic treatment is scant while proposed mechanisms are in excess, yet specific molecular drivers of HCC remain at large. Therefore, our study aims to identify drivers of HCC metastasis using protein-protein interaction (PPI) networks to identify key driver genes associated with HCC metastasis.

Methods: From differential expression genes (DEGs) analysis using GSE45114 microarray dataset, four main hub genes that correlated with patient survival were identified. The first hub gene, SERPINC1 had the highest centrality parameter in impeding HCC metastasis, implicating thrombin mediation through thrombin-induced tumor growth and angiogenesis.

Results: Our study reveals that thrombin was not differentially expressed, hence, suggesting the involvement of other, less-well studied pathways in impeding metastasis, such as KNG1, PAH, AMBP, and TTR. Findings for CD44 were consistent with existing literature. Meanwhile, FGG and APOA5, both less studied genes in the context cancer metastasis studies, were found to be crucial in impeding HCC metastasis.

Conclusion: This study identified four potential proteins (SERPINC1, CD44, FGG and APOA5) to be therapeutic targets or biomarkers and demonstrates the use of PPI networks for understanding HCC metastasis at a more profound level.

Keywords: Biomarkers, hepatocellular carcinoma, metastasis, Protein-Protein Interaction Network, therapeutic targets

Hepatocellular carcinoma (HCC) arises from malignant hepatocytes and has been estimated to have an incidence of >1 million cases by 2025. HCC is the most common type of liver cancers accounting for approximately 90% of liver cancer cases.[1] According to the Ministry of Health of Malaysia, liver cancer is ranked sixth for the most common cause of cancer accounting for both genders, while it is ranked third for the cancer-related death in the year 2020.[2] Others have also reported that the annual mortality rate per 100, 000 people is trended upwards, as the rate in the year 2013 was 6.1%, which is a drastic increase of 42.8% since 1990.[3]

Furthermore, HCC has been a subject of poor prognosis, accounting for the second leading cause of global cancer-related mortality and 9.1% of total death in the year 2012, despite its relative lower occurrence when compared with other, more common cancers. Numerous treatment methods including liver transplantation, radiofrequency ablation, and hepatic resection are available, but only modestly improves 5 years relative survival (of 3.1%). This is largely due to the limitations of the treatment methods, being limited to early stages of the disease. In the late stages of HCC, the cancer preferentially metastasizes to various sites including the portal vein, lungs, bones, brain, lymph nodes

and adrenal glands, and leading to the observed poor over-all survival (OS).[4] Therefore, investigating the metastatic mechanisms and progression of HCC is crucial to improve the OS of HCC patients.

Gene expression studies with the use of microarray technologies can quantify the expression of up to tens of thousands of genes simultaneously. From which, differentially expressed genes (DEGs) can be identified between groups of samples.[5] Although previous studies have performed said analysis to uncover the mechanisms of HCC metastasis, these studies often focus on gene-centric approaches, or at most, on enriching the DEGs to corresponding pathways.[6-8] With the advent of graph theory in the context of protein-protein interaction (PPI) networks, the scale-free nature of these biological networks of which can be taken advantage, allowing for the investigation and ranking of genes and pathways. One such application, is the identification of hub proteins. Hub proteins, that may play a gateway role and serve as therapeutic targets or biomarkers, along with the dynamic interactions between genes and pathways can be captured and be investigated at depth.[9]

In the present study, bioinformatics analyses were performed to identify the DEGs associated with HCC metastasis. This was performed, by analyzing previously submitted microarray data sets on Gene Expression Omnibus (GEO), whereby DEGs between metastatic HCC and non-metastatic HCC. Enrichment analysis was also performed to identify the pathways associated with the DEGs. After which, this study would analyze the network characteristics of the PPI network, along with the investigation of the network of enriched pathways, with various network parameters to better understand the underlying mechanisms of HCC metastasis and to identify potential hub genes for therapeutic efforts. Survival analysis would be performed on the identified hub genes to evaluate their significance to the OS of patients.

Methods

The flow chart of the methodology used in this study is shown graphically in Figure S1.

Microarray Data Retrieval and Processing

First, genes expression profiles must first be selected, to identify genes that are differentially expressed in metastatic HCC. Microarray data set search was conducted on GEO using the keywords "HCC." One gene expression profiling by array dataset was selected (accession no., GSE45114), of which, follows the following criteria: sample groups contain intratumoral or peritumoral HCC samples and that HCCs could be further classified into metastasis-present
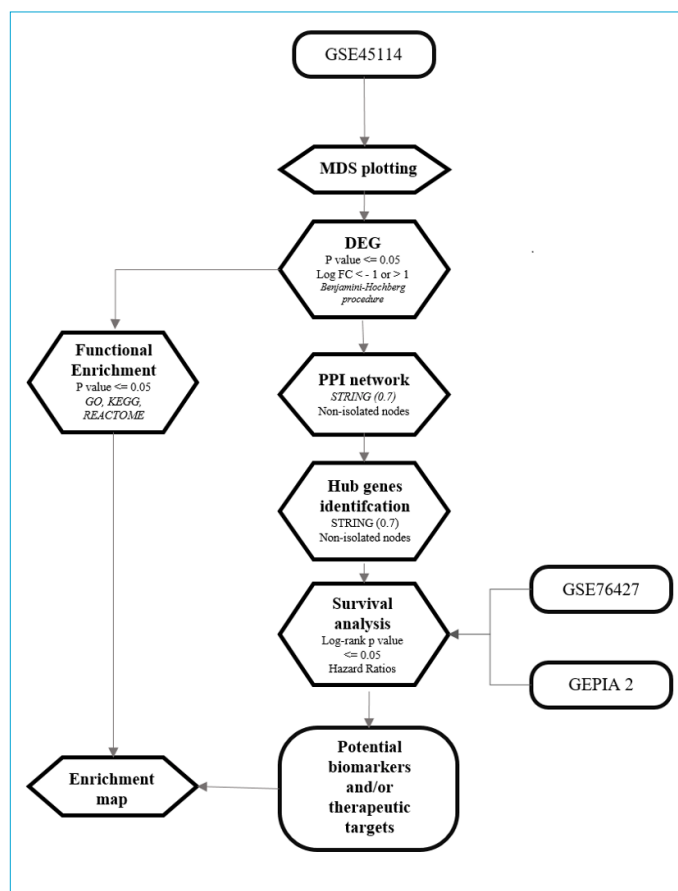
Figure S1. Flow chart depicting the overall design of the study at present. MDS, multidimensional scaling; DEG, differentially expressed genes; GO, go ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

or metastasis-absent. GSE45114 deposited by Wei et al.[8] provided 49 sets of samples, of which contained 24 sets of homogenous primary HCC (C) and 25 sets of pericancer liver tissues (P), in which, could be further categorized by the history of distant metastasis, of which, 27 sets had history of metastasis (M) while the remaining 22 sets did not (N). This resulted in a ratio of samples, CM:PM:CN:PN of 13:11:14:11. Besides, the dataset GSE76427 deposited by Grinchuk et al.[10] offered 167 sets of samples, 115 of which derived from primary HCC tumors, with the remaining 52 derived from the surrounding liver tissue of HCC tumors was used. This dataset has not recorded any information regarding metastasis, rather, it would be used for Kaplan Meier survival analysis, as it offered a longer follow-up period on overall patient survival.[10] The elaborate information regarding the datasets is provided in Table S1.

Differentially Expressed Genes (DEG) Analysis

Subsequently, data preparation and cleaning, gene annotation, and unsupervised clustering were conducted in R using the appropriate packages, namely, GEOquery,[11] Ti-

Table S1. Details of datasets used in present study.

Datasets	Sample	Array	Platform	HCC samples	Peritumor samples	Detail
GSE45114	HCC	mRNA	GPL5918	14 +13(M)[a]	11 +11(M)[a]	CapitalBio Human 22k oligonucleotide microarray
GSE76427	HCC	mRNA	GPL10558	115	52	Illumina HumanHT-12 V4.0 expression beadchip

[a]M denotes samples with patient history of distant metastasis. [b]H denotes endothelial tissues from Hemangioma of patients without HCC.

dyverse,[12] AnnotationDBI[13] with Ensembl gene database, and Limma,[14] before identification of DEGs. Unsupervised clustering, using multidimensional scaling (MDS) method, allows for the verification of correct group labeling for each sample (e.g., normal liver tissues should cluster together and not cluster separately), while also allowing a preliminary inspection, on the extent of DEGs one would be able to find.[15] To identify DEGs, the multiple linear model fitting and contrast fitting functions within the R package, Limma, was then used to apply linear model fitting for genes between groups of samples. The results were further adjusted for false discovery and filtering, by applying Benjamini and Hochberg's method of false discovery rate control, and by setting the criteria: $p<=0.05$ and absolute log fold change ($|logFC|$) $>=1$. Results are then visualized with the Volcano plot. Ultimately, the DEGs that were differentially expressed in metastatic HCC were subjected to enrichment analysis and PPI network construction.

Functional Enrichment Analysis

Following which, the biological functions and pathways associated with the DEGs in metastatic HCC were identified using functional enrichment analysis. This process encompasses the identification of relevant pathways or biological functions by associating the occurrence, or "hits," of genes in a gene set to pathways or biological functions curated in well-established, public pathway databases. In the present study, functional enrichment was performed using the DEGs along with their corresponding logFC values, using the msigdbr[16] and fgsea[17] packages in RAll three Gene Ontology (GO) categories, biological processes, cellular component (CC) and molecular function (MF) were selected and downloaded using the msigdbr package, and then enrichment was run using fgseasimple function fgsea had to offer. The analysis was repeated using Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and Reactome.

PPI Construction and Validation

Following the identification of genes was differentially expressed in metastatic HCC samples, they were used to construct a PPI network, to investigate the interactions between the DEGs. Cytoscape was the desktop application of

choice to visualize the PPI networks,[18] while the STRING app, that is an extension available in cytoscape, was the application of choice to import the network of DEGs.[19] The STRING app is able to predict interactions between proteins on the basis of omics data, including homology, genome features, co-expression and text mining, on top of experimental data, which allows for the discovery of novel functions for the proteins.[20] The cumulative score of all these data that predict interactions between genes, is quantified as STRING confidence scores, or full STRING interaction, and ranges between 0.0 and 1.0, with 0.0 indicating very low possibility of interaction and 1.0 indicating very high possibility of interaction. Interaction between the nodes was full STRING interactions set at the confidence score of 0.7. Network centrality values for the nodes in the PPI network were then calculated using the Network Analyzer app.[21]

Hub Genes Identification

Hub genes which have high centrality parameters and play important roles in regulating the information flow of the network were then identified. From the HCC metastasis PPI network, the top 10% of genes that scored high values in the centrality parameter, degree, were selected, and were identified as the hub genes. However, the betweenness centrality of the potential hub genes would also be investigated, to ascertain the importance of the gene in the context of information flow throughout the network.[21]

OS

To validate the importance of the hub genes in driving HCC metastasis, survival analysis was performed on the hub genes. Namely, the dataset GSE76427 is used to perform Kaplan Meier survival analysis. The median expression of hub genes was used as the cutoffs defining the low and high expression groups of the hub genes across the HCC samples in the GSE76427. Univariate Cox proportional hazard regression was performed in R using the survival[22] and survminer,[23] whereby the log rank p value and hazard ratios were calculated and compared. To mitigate biases and to investigate HCC subtype-specific survival analysis, GEPIA2-powered survival analysis was performed on the hub genes that correlated with patient survival in GSE76427. The survival

analyses using GEPIA2 online service offered a total of 4 available HCC expression datasets: The general dataset for HCC samples (164 samples), the dataset for iCluster 1 subtype of HCC samples (60 samples), the dataset for iCluster 2 subtype of HCC samples (54 samples) and the dataset for iCluster 3 subtype of HCC samples (62 samples).[24]

Pathway Networks Visualization

Next, to investigate the interactions or crosstalk between pathways, a network of pathways was then constructed and analyzed. This network comprised nodes depicting pathways and edges depicting the interactions between the pathways. This network is coined as the enrichment map. To construct an enrichment map, STRING functional enrichment of all nodes was first performed to identify the pathways associated with the PPI network. Then, an enrichment map was then created using the enrichment results through EnrichmentMap app.[25] Only the significant pathways generated from the prior enrichment analysis from fgsea were selected, with their NES values added into the nodes table. The AutoAnnotate app was then used to cluster and summarize the pathway networks with semantic annotations.[26]

Results

DEGs Identified in Metastatic HCC

The results for MDS clustering are shown in Figure S2. The plot generated was in agreement with the expected, whereby tumor and non-tumor samples were well differentiated against one another, while metastatic and non-metastatic samples were less well-differentiated with one. When subjecting the microarray data set to the Limma microarray analysis pipeline, a total number of 157 DEGs were identified, 130 of which were downregulated while the remaining 27 were upregulated. The results of the analysis can be visualized in the volcano plot Figure 1a.
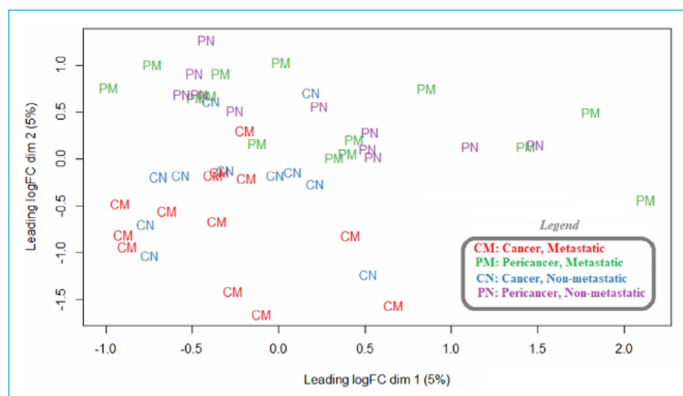
Figure S2. Unsupervised clustering of samples using Multidimensional Scaling.
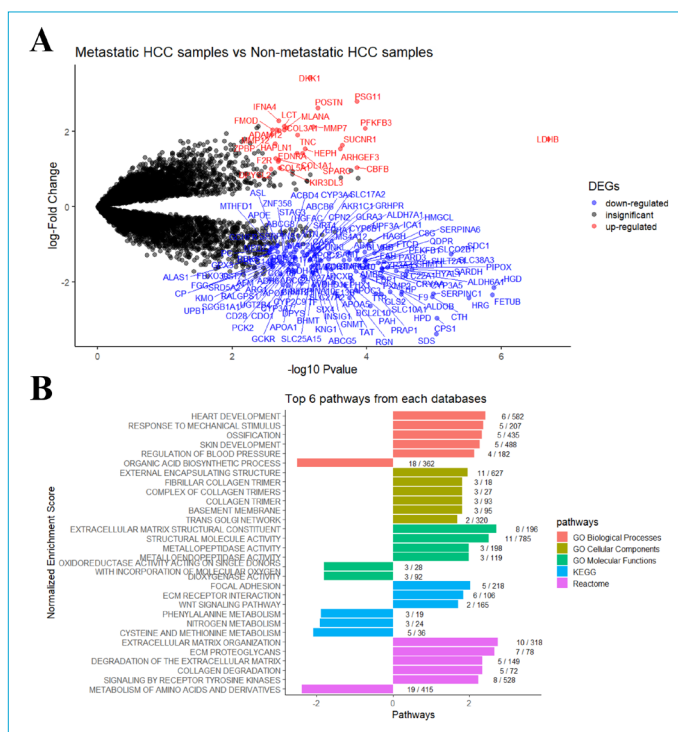
Figure 1. (a) Volcano plotting. Volcano plots were generated manually using the ggplot2 package. Significant DEGs were annotated using the DecideTests function, with "global" setting. **(b)** Functional enrichment. GO Biological processes, GO cellular components and GO molecular functions, KEGG and Reactome functional were performed on the DEGs. The results were filtered with p value <= 0.05.

## Functional Enrichment Analysis on DEGs

The identified DEGs were then subjected to functional enrichment analysis with the three GO pathways, KEGG and Reactome. The results were filtered with p<=0.05, resulting in total of 339 pathways, 252 of which from GO:BP, 17 of which from GO:CC, 33 of which from GO:MF, 9 of which from KEGG, and 28 of which from Reactome. The top 6 pathways from each database are visualized in Figure 1b.

PPI Construction Using STRING

Significantly expressed genes were queried on cytoscape using the STRING protein query (confidence score of >=0.7). To ensure the relevancy of the network with HCC metastasis, nodes that were isolated from the network which forms small clusters with 2 or less nodes were discarded from the main network. The network, was then adjusted to have node sizes to correspond to degree (increase of node size indicating increase of degree) the color of the nodes to correspond to the calculated logFC values (with blue indicating negative logFC, white indicating 0 logFC, and red indicating positive logFC). This resulted in a network with 99 nodes that are shown in Figure 2.

Hub Genes Identification

Hub genes selection was then carried out by selecting nodes with centrality values that were consistently high across the centrality parameters. The centrality parameter considered for hub gene determination was the degree parameter. A total of 10 hub genes were selected based on the above criteria and are visualized in Table 1. Of the 10 hub genes identified with the degree centrality parameter, their closeness centrality and between ness centrality were also investigated. The 10 hub genes were then subjected to survival analysis using microarray expression data provided by the GSE76427.
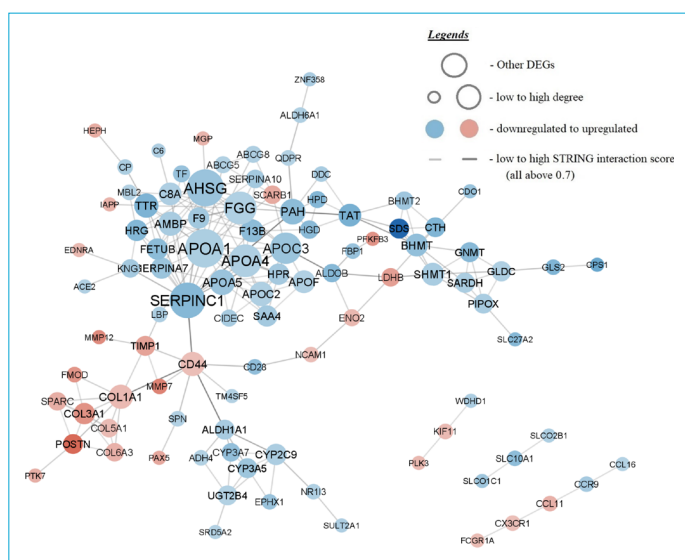
Figure 2. STRING (0.7) Network. The figure shows the network comprising of nodes of genes that were found to be differentially expressed in tumor samples with varying metastatic potential. Small clusters of nodes that were disconnected from the biggest cluster, with two or less number of nodes were removed.

## Overall Survival Analysis on Hub Genes

Univariate Cox proportional hazard regression performed on the 10 hub genes revealed that four of the 10 hub genes to significantly correlate with patients' OS. Namely, SERPINC1, CD44, FGG, and APOA5 have log rank p value of below the 0.05 threshold. The results are visualized in Table S2. The four hub genes were then subjected to a more thorough survival analysis using GEPIA2 services. Using the Univariate Cox proportional hazard regression analysis once more, the hazard ratio was estimated, the results of which are shown in Figure 3a-d. SERPINC1, FGG, and APOA5 demonstrated hazard ratios of below 1.000, indicating that patients that have higher than median expression levels of the genes, were more likely to succumb to death. On the other hand, patients with higher than median levels of CD44 are 1.589 times as likely to die than those who have lower than median levels of expression of CD44. In a nutshell, patients with higher expressions of SERPINC1, FGG, and APOA5 or patients who have lower expressions of CD44 are more likely to survive. Further information on these 4 hub genes, retrieved from the STRING database and the information regarding the DEG analysis is displayed in Table 2.

Table S2. Univariate Cox proportional hazard regression. Hub genes were subjected to Univariate Cox proportional hazard regression, in which the log-rank p values were calculated. The entries that are bolded in the table are hub genes with their log-rank p-values that do exceed the 0.05 threshold

Genes	Log-rank p	Genes	Log-rank p
1 APOA1	0.882	6 APOC3	0.727
2 AHSG	0.728	7 PAH	0.262
3 SERPINC1	0.017	8 APOA5	0.005
4 FGG	0.006	9 CD44	0.034
5 APOA4	0.294	10 COL1A1	0.466

Table 1. Hub genes selection. Hub genes are selected by selecting nodes with consistently high centrality values across degree and betweenness centrality

	ENSEMBL	SYMBOL	Degree	Closeness	Betweenness
1	ENSG00000118137	APOA1	18	0.328	0.109
2	ENSG00000145192	AHSG	17	0.343	0.093
3	ENSG00000117601	SERPINC1	16	0.384	0.468
4	ENSG00000171557	FGG	14	0.321	0.054
5	ENSG00000110244	APOA4	14	0.339	0.049
6	ENSG00000110245	APOC3	13	0.335	0.162
7	ENSG00000171759	PAH	9	0.269	0.264
8	ENSG00000110243	APOA5	9	0.326	0.004
9	ENSG00000026508	CD44	8	0.336	0.450
10	ENSG00000108821	COL1A1	8	0.265	0.151
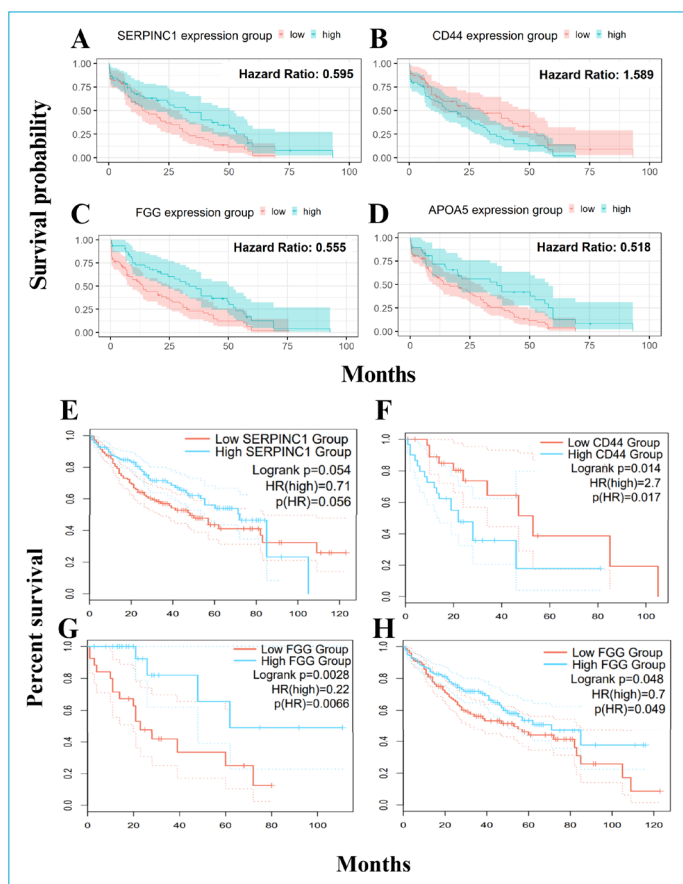
Figure 3. De novo Survival Analysis. The expression levels of SERPINC1 **(a)**, FGG **(b)**, CD44 **(c)** and APOA5 **(d)** were extracted from GSE76427 and was stratified into two groups: higher than median expression group in blue, and lower than median expression group in red. Survival Analysis with GEPIA 2. **(e)** Survival analysis of SERPINC1 with the general dataset. **(f)** Survival analysis of CD44 with the iCluster3 dataset. **(g)** Survival analysis of FGG with the general dataset. **(h)** Survival analysis of FGG with the iCluster2 dataset.

The hub genes are also subjected to survival analysis using the available GEPIA2 database and online services. The entire results are shown in Figure S3, while the results with significant log rank p values are shown in Figure 3e-h.

SERPINC1 survival analysis using the general HCC dataset, FGG survival analyses using the general HCC and iCluster 2 datasets, and CD44 survival analysis using the iCluster 3 dataset demonstrated significance. The hazard ratios of the four analyses were relatively consistent with the ones generated in the present study, whereby SERPINC1 and FGG had hazard ratios of below 1.000, while CD44 had a hazard ratio of above 1.000. APOA5 expression levels, however, did not demonstrate correlation with patient survival using any of the GEPIA 2 datasets.

Pathway Networks Visualization

The constructed PPI network was then subjected to STRING functional enrichment and enrichment map generation, in which the results generated prior with fgsea were imported into cytoscape to outfit the pathway nodes with NES values and to filter pathways. This resulted in a network of 59 pathways. Following the use of Autoannotate app to cluster the pathways, numerous large clusters can be seen. The results are visualized in Figure 4. The biggest cluster, "Fibrillar Collagen Trimers" contains 19 upregulated pathways and is associated with the hub genes, FGG and CD44. This is followed up by the second biggest cluster, "Small Process Catabolic," which contains 10 downregulated pathways and is directly associated with APOA5 and CD44 hub genes. There were also two small clusters, each with two upregulated pathways, "Response wounding wound" and "cell migration locomotion," that are associated with all hub genes except for APOA5 and were linked to the "Fibrillar Collagen Trimers" cluster.

Discussion

HCC is a relatively common cancer with a high mortality rate due to the limitations of contemporary treatment to only be effective in early stages of the disease.[4] In the late stages of HCC, the tumor or tumors progress, accumulating dysregulation of genes and pathways that ultimately confers the disease the ability to metastasize. Therefore,

Table 2. Potential therapeutic targets. Hub genes that correlated with patient overall survival are displayed here, along with the information retrieved from the STRING database and the DEG analysis

STRING ID	SYMBOL	ENSEMBL ID	Gene name	logFC	p
9606.ENSP00000356671	SERPINC1	ENSG00000117601	serpin family C member 1	-1.699	0.000
9606.ENSP00000398632	CD44	ENSG00000026508	CD44 molecule (Indian blood group)	1.014	0.048
9606.ENSP00000336829	FGG	ENSG00000171557	fibrinogen gamma chain	-1.064	0.016
9606.ENSP00000445002	APOA5	ENSG00000110243	apolipoprotein A5	-1.396	0.013
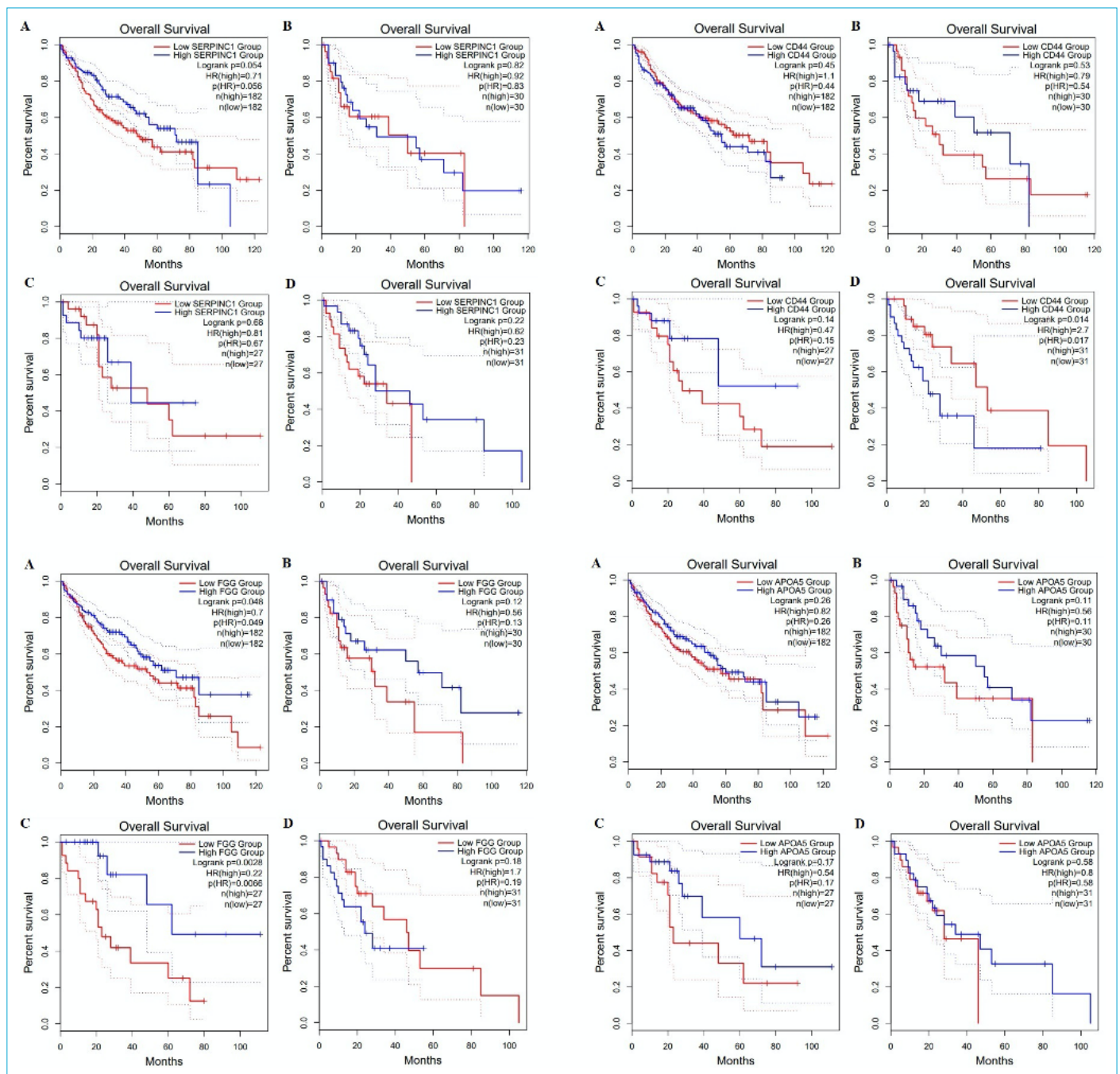
Figure S3. GEPIA 2 survival analysis on hub genes.

this study aims to capture the genes and pathways that are dysregulated in the advanced stage of the disease, the very same genes and pathways that drive metastasis in HCC. By the identification of the relevant genes and pathways which are important in the context of network centralness, the mechanism behind HCC metastasis could be better understood, while potentially novel biomarkers and therapeutic targets may be derived. These potential biomarkers or therapeutic targets present themselves as hub genes, which were identified as SERPINC1, CD44, FGG, and APOA5.

The full name of SERPINC1 is serpin family C member 1. It is also known as AT3, AT3D, THPH7, ATIII, ATIII-R2, ATIII-T1, or ATIII-T2. SERPINC1 encodes the protein, antithrombin III, which functions to inhibit thrombin in the coagulation system.[27] SERPINC1 has also been reported to suppress the invasion and metastasis of several cancers. Some studies have highlighted the role of SERPINC1 preventing hepatic ischemia/reperfusion-induced metastasis, namely, through interaction with thrombin, macrophage inhibitory factor (MIF) and the tumor necrosis factor alpha (TNF-α) produc-
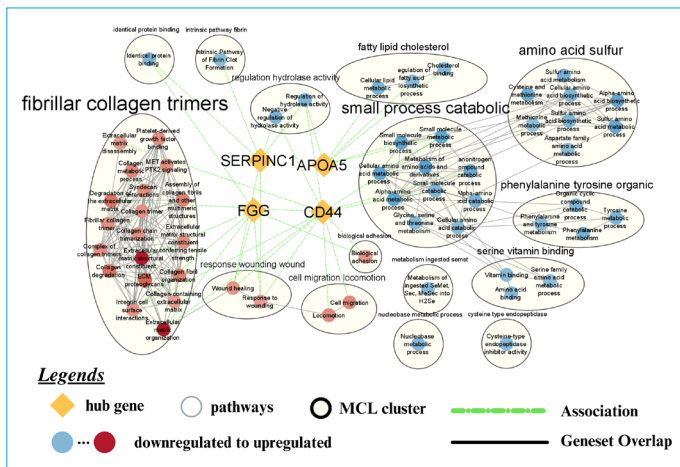
Figure 4. Pathway Network Visualization. STRING functional enrichment was performed to generate an enrichment map. The enrichment results table, which compiled all 3 categories of GO terms, along with KEGG and Reactome pathways, were filtered by p value <= 0.05 and were then imported into the nodes table. The diamond nodes are the identified hub genes. The translucent yellow bubbles are the results of auto annotate clustering using MCL algorithm. The grey edges depict interactions between pathways, via gene set overlap, while the green edges depict the involvement of hub genes in the pathways.

tion pathways.[28]

Thrombin has been demonstrated to be associated with more malignant phenotype in vivo through cancer.[28] First, protease activated receptor 1 which functions as a thrombin receptor, has been shown to be expressed in HCC cell lines. The activation the receptor through binding with thrombin, leads to tumor growth and angiogenesis.[29] Furthermore, thrombin cleaves osteopontin. The osteopontin fragments in turn, have demonstrated to increase likelihood of vascular invasion, most likely through means of osteopontin-CD44 interactions.[30,31] Finally, thrombin has demonstrated to increase the expression and secretion of MIF. SERPINC1, however, is able to cleave, and consequently inactivate thrombin, leading to diminished metastasis-promoting activity of thrombin.[27] These may explain the role of SERPINC1 in conferring protective effects against HCC metastasis.

MIF behaves as a regulatory element for the innate immune system, and increased expression of which normally results in increased systemic or local inflammation.[32] In the context of cancer metastasis, MIF exerts its metastatic effects through its tautomerase activity.[33] MIF's metastasis-promoting activities are thwarted by binding with SERPINC1, as ATIII–MIF complexes, the result of SERPINC1-MIF interaction, exert less biological effects due to the blocking of the reactive, and N-terminal proline residue in MIF that is required for tautomerase activity.[34]

In the context of colon cancer cells in a rat models, TNF-α

is a pro-inflammatory cytokine and has been demonstrated to increase the expression levels of leukocyte adhesion molecules, which are able to enhance the lodgment of HCC cells on sinusoidal surfaces, hence, promoting the formation of metastatic modules. This occurs in a not-well studied mechanism, which does not involve thrombin.[35]

However, these mechanisms did not appear significant in this study, as all three mediators, TNF-α, MIF, and thrombin were not differentially expressed when comparing HCC tissues with varying metastatic potential, as demonstrated in the PPI network in Figure 5a. Further investigation also revealed that although thrombin did not qualify as a hub gene through Benjamini-Hochberg procedure, it did have a p value of 0.006 and a logFC of −1.515, indicating likelihood that it was downregulated in metastatic HCC instead. Likewise, pathways involved in TNF-α production or MIF-mediated functions were not enriched in the present study. This could be partially explained by the nature of the study design employed in the GSE45114 dataset, which is strictly mRNA-focused and does not take into consideration of protein activity. As thrombin and MIF were not differentially expressed in HCC samples with varying metastatic potential, it could be inferred that there would be
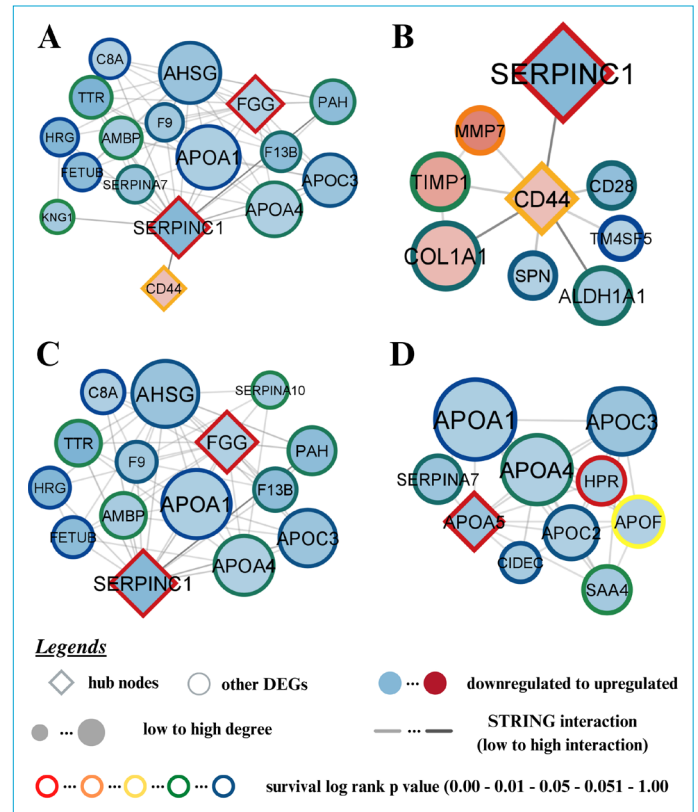
Figure 5. The PPI networks was sub-networked by selecting only hub genes and its immediate neighbors. Survival analysis-derived log rank p value was also merged with the network, shown by the colored borders. **(a)** SERPINC1. **(b)** CD44. **(c)** FGG. **(d)** APOA5.

similar levels of the two proteins in both HCC types. However, due to the varying levels of SERPINC1, the extent of which thrombin and MIF are active in either HCC types may vary. Hence, it may be inferred that in HCC samples with lower metastatic potential, SERPINC1 is expressed at higher levels, causing the binding and subsequent inactivation of both thrombin and MIF, leading to the observed lower metastatic potential. TNF-α related pathways on the other hand, can be more confidently inferred as less relevant due to the absence of TNF-α differential expression.

As mentioned by prior studies, wound healing processes are inextricably linked with tumor progression and metastasis.[36] In particular, the pathways associated with wound healing are co-opted or corrupted in cancers, causing aberrations in the three wound healing processes: Inflammation, regeneration and remodeling, and leading to metastasis.[37] In the present study, expression levels of SERPINC1 appears to correlate with the wound healing and wound healing response processes, as shown in Figure 4 and Table S3. The three processes of wound healing were also apparent, in the involvement of inflammatory processes, as demonstrated by the dysregulation of HRG, KNG1, and CXCR1, in the involvement of regeneration processes, as demonstrated by the dysregulation of SPARC, SCARB1 and PTK7, and even in remodeling processes, as demonstrate by the dysregulation of TIMP1 and MMP12.[37-45] At the same time, all nine of these aforementioned genes have been demonstrated to be involved in cancer metastasis.

Furthermore, the importance of wound healing pathways in driving metastasis in HCC, can be further demonstrat-ed in the results of functional enrichment of all DEGs, as demonstrated in the enrichment map, whereby the largest cluster of pathways is "fibrillar collagen trimmers," also in GO biological processes such as skin development, tissue morphogenesis, morphogenesis of an epithelium processes et cetera, in GO molecular functions such as extracellular matrix (ECM) structural activity, growth factor activity et cetera, in KEGG pathways such as ECM interaction, WNT signal pathway and JAK STAT signaling pathway, and in reactome pathways such as degradation of ECM, collagen degradation, and developmental biology.

Finally, the present study also identified other potential mediators and pathways that may play a crucial role in HCC metastasis. In particular, SERPINC1 has also been predicted using STRING algorithm to interact with kininogen 1 (KNG1), phenylalanine hydroxylase (PAH), adrenomedullin binding protein (AMBP), and transthyretin (TTR). KNG1 has been demonstrated in inhibiting proliferation and inducing apoptosis in glioma cells, potentially via P13/AKT signaling pathway.[45] PAH, a less-well studied biomarker for HCC, has been reported as a prognostic marker for poor outcome, albeit with undiscovered means of involvement. [46] AMBP, on the other hand, which binds to adrenomedullin, an angiogenic peptide that functions under putative hypoxic conditions, reduces the availability, hence bioactivity of adrenomedullin, and reducing the risks of metastasis.[47] And finally, TTR which functions in oxidative inducing and detecting pathways, has also been demonstrated to be inversely correlated with metastatic potential in HCC, although with less-well studied mechanisms.[48,49] These

Table S3. Pathways involving SERPINC1 are extracted from the node table from the enrichment map network. Corresponding NES are also depicted. Other genes that are present in the PPI network and are involved in the particular pathways are also depicted, whereby the hub genes are also bolded.

Pathways	NES	Genes
1 Identical protein binding	-1.753	NCAM1|IAPP|COL1A1|SHMT1|APOA1|FGG|APOA4|TTR|TAT|SDS|SERPINC1|CTH|SCARB1|GNMT|APOC2|ALDOB|GLDC|AMBP|LDHB|CD28|FBP1|QDPR|HGD
2 Intrinsic Pathway of Fibrin Clot Formation	-1.619	F9|KNG1|SERPINC1
3 Negative regulation of hydrolase activity	-1.849	COL6A3|TIMP1|APOC3|HRG|SERPINA7|APOA1|SERPINC1|SERPINA10|KNG1|FETUB|AMBP|AHSG|CD44
4 Wound healing	1.743	F9|TIMP1|COL3A1|SPARC|COL1A1|HRG|FGG|F13B|SERPINC1|MMP12|SCARB1|SERPINA10|KNG1|PTK7|COL5A1|CD44
5 Response to wounding	1.588	F9|TIMP1|COL3A1|SPARC|COL1A1|HRG|APOA1|FGG|F13B|SERPINC1|MMP12|SCARB1|SERPINA10|KNG1|PTK7|COL5A1|CD44|CX3CR1
6 Regulation of hydrolase activity	-1.753	COL6A3|CCL16|ALDH1A1|TIMP1|APOC3|CCL11|HRG|SERPINA7|APOA1|APOA4|SERPINC1|SERPINA10|APOC2|EDNRA|ALDOB|KNG1|FETUB|AMBP|AHSG|CD44|APOA5

mediators and pathways may present as potential HCC metastasis mechanisms, which are thwarted by the effects of SERPINC1.

Due to the high closeness and between ness centrality parameters of SERPINC1, it is highly possible that SERPINC1 assumes high importance at the regional and global levels in the HCC metastasis PPI network.[50] In other words, SERPINC1 may serve as a gateway gene that mediates the interactions between pathways, both at the regional level, for example, in the mediation among wound healing pathways, as well as at the global level, for example, in the mediation of interactions between wound healing pathways and pathways at the periphery of the network. Likewise, wound healing may be the central pathway that mediates and interacts with the pathways in the periphery, such as AMBP-mediated angiogenic pathways or TTR-mediated oxidative pathways. Therefore, SERPINC1 downregulation may be indispensably linked with the occurrence's metastasis, especially in a non-HCC-subtype related manner, as indicated in the OS analysis.

CD44 is a transmembrane glycoprotein with no kinase activity. CD44 is expressed in a wide range of tissues, with expression in the skin ranking the highest. On binding with its ligand hyaluronan, cascades of signaling pathways involved in cell survival, cytoskeletal changes and cell motility, as well as cell proliferation are all induced. It can also interact with collagens, osteopontin, and matrix metalloproteinases.[51,52] Furthermore, it is overexpressed in cancer stem cells and has shown implication in cancer progression, in particularly its isoform, CD44s, may have a role in regulating epithelial to mesenchymal transition (EMT).[51] It has been demonstrated that CD44 promotes proliferation and migration of HCC cells through YAP, CASP1, and ZEB1.

YAP is an important downstream regulator of the Hippo pathway, which is a pathway that plays a key role in driving metastasis in a variety of cancers.[53,54] The Hippo pathway enhances the translocation of transcription activators from the cytoplasm into the nucleus, subsequently activating a number of oncogenic genes.[53] Shin et al.[54] has demonstrated that CD44 upregulates YAP, and further demonstrated that following the knockout of CD44, the migration and proliferation ability of HCC cell lines were restored through overexpression of YAP.

CASP1 is activated in macrophages following inflammation-promoting stimuli, such as hypoxic stimuli, and is involved in the activation of various inflammatory cytokines, namely, IL1B.[55] Following CASP1 activation, the secreted IL1B promotes HCC progression and EMT.[56] Li et al.[57] demonstrated that CD44 promotes CASP1/IL1B pathway through inhibiting autophagic degradation of CASP1.

ZEB1 belongs to a group of transcription factors that represses epithelial gene expression. The activation of these transcription factors leads to EMT.[56] CD44, has been demonstrated by Li et al.[57] to be engaged in a cyclic feedback loop, causing the upregulation of both genes in HCC. CD44, in particular the CD44s isoform was shown to be able to activate ZEB1 expression, which in turn downregulates ESRP1, further promoting the synthesis of, CD44s. These cyclical mechanisms allow HCC cells to maintain stemness of tumor, independent of external stimuli.[57]

This study, however, did not find these three major mechanisms to be significant. This is because the two of the mediators, YAP and CASP1 and did not appear to be significantly differentially expressed as shown in Figure 5b, while ZEB1 was not detectable, due to the lack of corresponding probes to capture its mRNA in the microarray dataset GSE45114. Likewise, ESRP1 was not detectable due to the lack of ESRP1 mRNA capturing probes, leading to inability to confirm or deny the involvement of CD44-ZEB1 interactions in driving metastasis. As such, the mechanism behind metastasis driven by CD44 may be mediated through other less-well studied pathways and genes.

In the present study, numerous other pathways are enriched along with the differential expression of CD44 as shown in Table S4. These included pathways concerning ECM organization, ECM degradation, ECM disassembly, locomotion, cell migration, biological adhesion, and integrin cell surface interactions. With the exception of FGG, none of the other 2 hub genes are involved in these pathways.

ECM remodeling is required in the surrounding non-tumor tissues to elicit further tumor growth, increased migration, and ultimately distant metastasis.[58] In the present study, ECM remodeling could be activated from CD44, through mediators such as MMP7, which is able to degrade the basal membrane and to cleave E-cadherin ectodomain leading to inhibition of cell aggregation and induction of cell invasion,[59] and TIMP-1, which has been demonstrated to trigger formation of pre-metastatic niche in distant sites.[60] MMP7-mediated mechanism appeared to have high involvement in the present study, due to its significant log rank p value as shown in Figure 5b. Furthermore, glycoproteins on tumor cells, including CD44, are often upregulated, resulting in a bulkier and thicker glycocalyx layer.[58] These would also result in increased tension applied from the glycocalyx layer onto the ECM-bound integrins, independent of myosin involvement, resulting in integrin priming and activation of integrin signaling which would result in pro-tumorigenic signals, such as growth and survival signals.[61]

CD44 also appeared to interact with COL1A1 as well, as

Table S4. Pathways involving CD44 are extracted from the node table from the enrichmentmap network. Corresponding NES are also depicted. Other genes that are present in the PPI network and are involved in the particular pathways are also depicted, whereby the hub genes are also bolded.

Pathways	NES	Genes
1 Locomotion	1.949	LBP|CCR9|CCL16|NCAM1|COL1A1|CCL11|HRG| APOA1|SPN|MMP12|SCARB1|PTK7|COL5A1|CD44|SAA4|CX3CR1
2 Extracellular matrix organization	2.72	COL6A3|TIMP1|COL3A1|NCAM1|SPARC|COL1A1|MMP7|FMOD|FGG|TTR|MMP12|COL5A1|CD44
3 Biological adhesion	1.675	COL6A3|COL3A1|NCAM1|CCL11|FGG|APOA4| SPN|ACE2|SCARB1|POSTN|PTK7|AMBP|COL5A1| CD44|CX3CR1
4 Integrin cell surface interactions	1.734	COL6A3|COL3A1|COL1A1|COL5A1|FGG|CD44
5 Extracellular matrix disassembly	1.653	MMP12|TIMP1|MMP7|CD44
6 Negative regulation of hydrolase activity	-1.849	COL6A3|TIMP1|APOC3|HRG|SERPINA7|APOA1| SERPINC1|SERPINA10|KNG1|FETUB|AMBP|AHSG|CD44
7 Wound healing	1.743	F9|TIMP1|COL3A1|SPARC|COL1A1|HRG|FGG| F13B|SERPINC1|MMP12|SCARB1|SERPINA10| KNG1|PTK7|COL5A1|CD44
8 Organonitrogen compound catabolic process	-2.072	HPD|PAH|ALDH6A1|CYP3A5|GLS2|PIPOX|SHMT1|FMOD|TAT|CDO1|SDS|CTH|MMP12|SARDH| GLDC|AMBP|CD44|BHMT|QDPR|HGD
9 Response to wounding	1.588	F9|TIMP1|COL3A1|SPARC|COL1A1|HRG|APOA1| FGG|F13B|SERPINC1|MMP12|SCARB1| SERPINA10|KNG1|PTK7|COL5A1|CD44|CX3CR1
10 Cell migration	1.943	LBP|CCR9|CCL16|COL1A1|CCL11|APOA1|SPN| MMP12|SCARB1|PTK7|COL5A1|CD44|SAA4| CX3CR1
11 Degradation of the extracellular matrix	2.326	MMP12|COL6A3|TIMP1|COL3A1|COL1A1|MMP7|COL5A1|CD44
12 Regulation of hydrolase activity	-1.753	COL6A3|CCL16|ALDH1A1|TIMP1|APOC3|CCL11| HRG|SERPINA7|APOA1|APOA4|SERPINC1| SERPINA10|APOC2|EDNRA|ALDOB|KNG1|FETUB| AMBP|AHSG|CD44|APOA5
13 Small molecule catabolic process	-1.780	HPD|PAH|ALDH6A1|ALDH1A1|SULT2A1|GLS2| PIPOX|SHMT1|TAT|CDO1|SDS|CTH|SCARB1| SARDH|ALDOB|GLDC|ADH4|SLC27A2|CD44| QDPR|ENO2|HGD
14 Small molecule metabolic process	-2.197	PAH|ALDH6A1|EPHX1|SULT2A1|CYP3A5|FMOD| APOA1|APOA4|TTR|CDO1|BHMT2|SDS|DDC| CYP2C9|SCARB1|SRD5A2|PFKFB3|ADH4|SLC27A2|CYP3A7|BHMT|QDPR|HGD|HPD|ALDH1A1| UGT2B4|GLS2|PIPOX|SHMT1|TAT|CTH|SARDH| GNMT|ALDOB|GLDC|LDHB|APOF|CD44|CPS1| FBP1|ENO2|APOA5

shown in Figure 5b. COL1A1 has been associated with numerous cancers including HCC.[62] COL1A1 expression level had been shown to be directly correlated with CCND1 and BCL2, which have both been established to be implicated in EMT and cell survival, respectively.[63] In the present study, it is noteworthy that although CCND1 and BCL2 were not included in the PPI network, they were identified as DEGs, having p values of lower than 0.05 and log-fold change of above +1.000, indicating potential involvement. Besides, COL1A1 has been shown to promote Wnt signaling, a signaling pathway that was enriched in the present

study, had been demonstrated in other studies to increase cell migration in colorectal carcinoma.[63,64]

Furthermore, CD44 appeared to interact with ALDH1A1. ALDH1A1 has been demonstrated to be implicated in modification of metabolism and maintenance of cancer stem cell properties, as well as DNA repair in the context of cancer involvement.[65] Seubert et al.[62] demonstrated that ALDH1A1 is able to increase CD44 expression through AURKA. This could partially explain as to why concurrent expression of CD44 and ALDH1A1 increases the aggressiveness of tumors, and increased likelihood for metastasis occurrence.
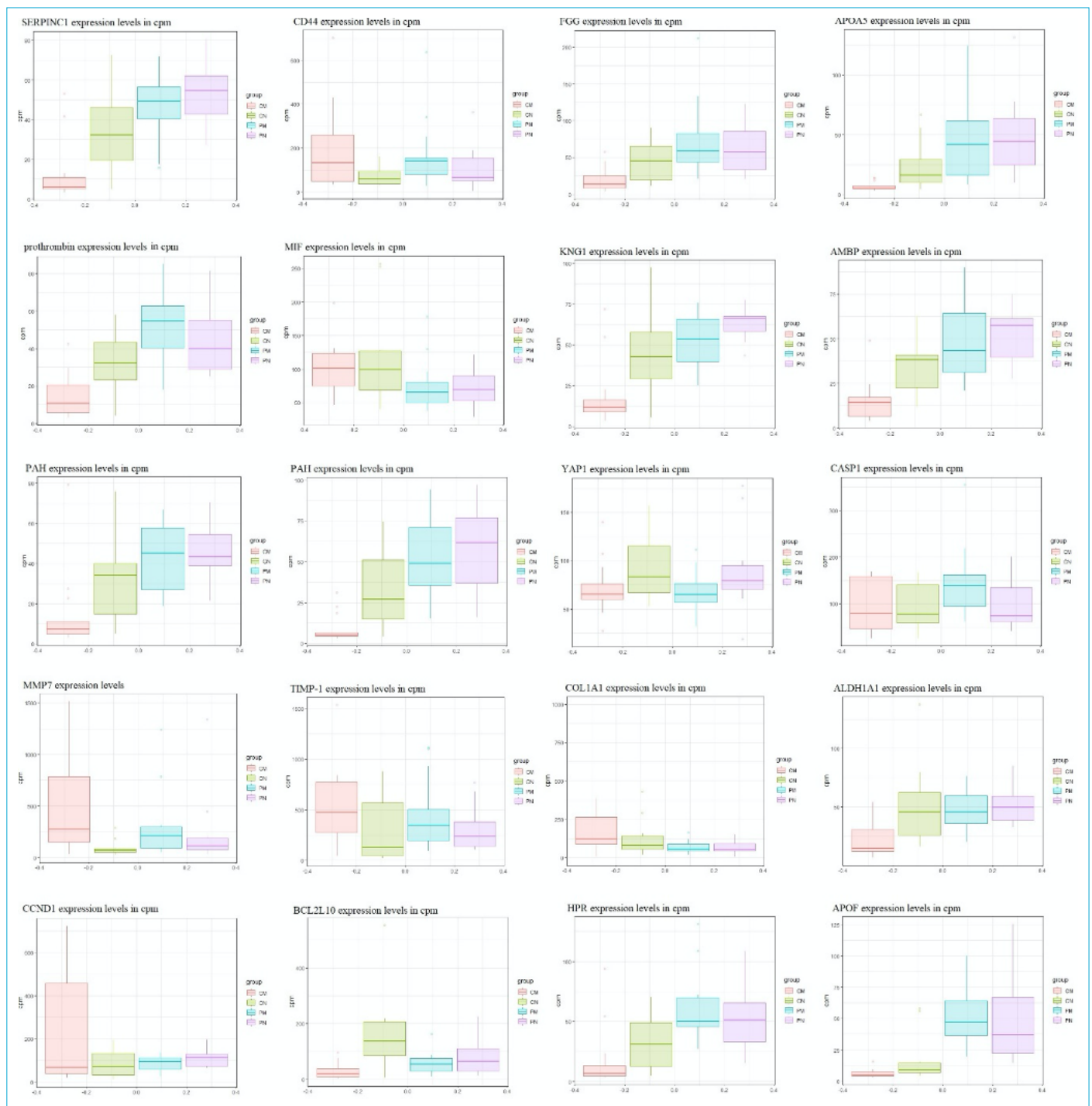
Figure S4. Boxplot of gene expression levels in counts per million (cpm) from GSE45114 of the hub genes and other mentioned genes.

Although CD44 had a relatively low degree of 8 when compared with the other potential hub genes, CD44 had the third highest closeness centrality and the second highest betweenness centrality, indicating its importance at regional and global levels.[66] These regional and global levels of significance of CD44 in the HCC metastasis PPI network may be better understood through investigating its interactions with other genes and pathways. First, its regional importance may lie in the facts that it plays a crucial role in regulating ECM remodeling, locomotion and migration related pathways. On the other hand, its global level of importance may be seen in its importance in connecting the metastasis-promoting influences of COL1A1 and ALDH1A1, both genes that lie in the periphery of the network, to the more central pathways associated with SERPINC1, respectively.

The FGG gene, which encodes the gamma chain of fibrinogen, is a blood-borne glycoprotein that is uniquely expressed by liver tissues. Under physiologic conditions, the cleavage products of fibrinogen, including FGG, play important roles in regulating cell adhesion and spread, along roles in vasoconstriction and chemotactic activities.[67] FGG overexpression has been reported to be associated with increasing metastatic potential of HCC, whereby FGG overexpression has been linked to HCC tumors exhibiting more advanced TNM stage and increased microvascular invasion, leading to overall higher recurrence rate and shorter OS time of patients.[68] The same study further demonstrated that FGG promotes the migration of SK-HEP-1 cell lines through regulation of Slug and ZEB1, leading to increased EMT.

The finding in this study, however, contradicted the notion that FGG overexpression's association with more aggressive phenotypes. As demonstrated in Figure 5c, FGG appeared to be under-expressed in metastatic HCC from the microarray dataset GSE45114, and its lowered expression in GSE76427 samples to be associated with lower survival. GEPIA2 analysis of FGG, however, offered a slightly different perspective. GEPIA2 survival analysis using the general and iCluster 2 datasets revealed that FGG under expression is associated with lower survival with significant log rank p value, while using the iCluster 3 dataset revealed that high expression of FGG to associate with lower survival instead, albeit with an insignificant log rank p value as seen in Figure S3. This could be indicative that the role of FGG in HCC metastasis is context dependent, whereby its upregulation drives metastasis in iCluster 3 HCC tumor types, while its

downregulation in general and iCluster 2 HCC types drive metastasis instead.

FGG expression had been associated with numerous pathways in the present study, as shown in Table S5. Most noticeably, FGG shared its involvement in the wound healing pathways with SERPINC1, at the same time, shared its involvement in the ECM-related pathways with CD44. As tempting as it would be to associate FGG as the gateway between the two pathways, STRING analysis revealed that FGG only interacted with SERPINC1, but not CD44. Furthermore, FGG also shared with SERPINC1, interactions with AMBP, TTR, and PAH, which further demonstrates the likelihood of AMBP-, TTR-, and PAH- related mechanisms.

In the context of network importance, FGG ranked 4th for degree with 14 edges, but did not have superior closeness and betweenness centrality values when compared to the other hub genes. This is indicative that FGG is only important in the local context, and not in the regional and global context of the PPI network.[66] Regardless of which, FGG may still have potential to qualify as a potential biomarker and therapeutic target in advanced HCC. Further research is warranted to unveil the role of FGG in impeding or driving HCC metastasis, in more detailed contexts such as HCC subtypes.

Apolipoprotein A5 or APOA5 is a component of high-density lipoprotein and is essential in the regulation of plasma levels of triglycerides. It is found in high density lipoproteins and its dysregulation have been associated with hypertriglyceridemia and hyperlipoproteinemia.[69] Lipoproteins have also been demonstrated to have a role in the progression of cancers. Namely, APOA1 has been demonstrated by

Table S5. Pathways involving FGG are extracted from the node table from the enrichment map network. Corresponding NES are also depicted. Other genes that are present in the PPI network and are involved in the particular pathways are also depicted, whereby the hub genes are also bolded.

Pathways	NES	Genes
1 Identical protein binding	-1.753	NCAM1|IAPP|COL1A1|SHMT1|APOA1|**FGG**|APOA4| TTR|TAT|SDS|SERPINC1|CTH|SCARB1|GNMT|APOC2|ALDOB|GLDC|AMBP|LDHB|CD28|FBP1|QDPR|HGD
2 Extracellular matrix structural constituent	2.733	COL6A3|COL3A1|COL1A1|MGP|COL5A1|**FGG**
3 Extracellular matrix organization	2.720	COL6A3|TIMP1|COL3A1|NCAM1|SPARC|COL1A1| MMP7|FMOD|**FGG**|TTR|MMP12|COL5A1|**CD44**
4 Biological adhesion	1.675	COL6A3|COL3A1|NCAM1|CCL11|**FGG**|APOA4|SPN|ACE2|SCARB1|POSTN|PTK7|AMBP|COL5A1|**CD44**| CX3CR1
5 Integrin cell surface interactions	1.734	COL6A3|COL3A1|COL1A1|COL5A1|**FGG**|**CD44**
6 Wound healing	1.743	F9|TIMP1|COL3A1|SPARC|COL1A1|HRG|**FGG**|F13B| SERPINC1|MMP12|SCARB1|SERPINA10|KNG1|PTK7| COL5A1|**CD44**
7 Response to wounding	1.588	F9|TIMP1|COL3A1|SPARC|COL1A1|HRG|APOA1| **FGG**|F13B|SERPINC1|MMP12|SCARB1|SERPINA10| KNG1|PTK7|COL5A1|**CD44**|CX3CR1

Ren et al.[70] to potentially downregulate mitogen-activated protein kinase (MAPK) pathway, leading to subsequent cell cycle arrest promoting apoptosis and inhibiting HCC proliferation, while APOA2 and APOA4 have shown promise as biomarkers, due to their downregulation in HCC.[71] APOA5, on the other hand, has not been reported in HCC development. Such findings were relatively consistent with the findings in the present study, whereby all APOAs were found to be downregulated, as demonstrated in Fig. 5d. However, the previously reported MAPK involvement did not appear significant in the present study. Therefore, the role of APOA1 was unclear in this study.

Another study has demonstrated that APOA5 expression in hepatocytes can be reduced following treatment with inflammatory signals, TNF-α and IL-1.[72] The study further demonstrated that the promoter sequence of APOA5 contains PPARα response element, which on signaling instigated by TNFα, leads to the downregulation of the APOA5 gene expression. This is indicative that APOA5 expression can be regulated through inflammation, which also plays crucial processes in cancer progression.

In the STRING network constructed, it was also revealed that APOA5 may interact with HPR and APOF as well. These two genes also had correlation with OS of patients with HCC as shown in Figure 5d. HPR has also shown promise as a prognostic marker in HCC, whereby its higher expression of HPR was correlated with better differentiation of HCC cells and subsequent better OS in patients.[73] Such was consistent with the present study, as HPR expression did appear to correlate with patient survival as well, as indicated in Figure 5d. Besides such, HPR have yet been associated with potential HCC metastasis-driving pathways in existing literature,

nor in the present study. Meanwhile, APOF has also been investigated and showed promise to be a predictor for HCC progression.[74,75] The same study further demonstrated that APOF may have tumor suppressive functions, whereby SMMC-7721 and Huh-7 cell lines displayed lowered proliferative abilities following APOF upregulation, and in vivo analysis revealed tumor growth inhibition.

In the context of network centrality, APOA5 did not have superior degree, closeness centrality, and betweenness centrality values when compared against other hub genes, indicating its relatively lowered importance for the information flow in the HCC metastasis PPI network.[66] But once again, by virtue of having degree values in the upper 10% of the network and of having significant correlation with survival of HCC patients, APOA5 may be a crucial hub gene linking other genes and pathways, hence, may subsequently impede metastasis in HCC (Table S6). Although with less weight of importance compared with the three other hub genes, APOA5 therefore may still be a potential biomarker or therapeutic target in advanced stages of HCC.

Figure 6 encapsulates all of the pathways and its mediators, in which the hub genes may drive or impede metastasis. The figure contains all of the mechanisms that were suggested by the results of the present study and had been proven by prior studies in regards to their relevance with cancer metastasis.

Conclusion

The present study provided a comprehensive bioinformatics analysis to identify potential predictive biomarkers for advanced HCC. Four identified candidate therapeutic targets, namely, SERPINC1, CD44, FGG, and APOA5 were found

Table S6. Pathways involving APOA5 are extracted from the node table from the enrichment map network. Corresponding NES are also depicted. Other genes that are present in the PPI network and are involved in the particular pathways are also depicted, whereby the hub genes are also bolded.

Pathways	NES	Genes
1 Cholesterol binding	-1.695	APOC3|APOF|APOA1|APOA4|APOA5
2 Regulation of fatty acid biosynthetic process	-1.645	APOC2|APOC3|APOA1|APOA4|APOA5
3 Regulation of hydrolase activity	-1.753	COL6A3|CCL16|ALDH1A1|TIMP1|APOC3|CCL11|HRG|SERPINA7|APOA1|APOA4|SERPINC1|SERPINA10|APOC2|EDNRA|ALDOB|KNG1|FETUB|AMBP|AHSG|CD44|APOA5
4 Cellular lipid metabolic process	-1.841	ALDH1A1|EPHX1|CYP3A5|APOC3|APOA1|APOA4|TTR|CYP2C9|APOC2|ADH4|SLC27A2|CYP3A7|CPS1|APOA5
5 Small molecule biosynthetic process	-2.237	PAH|GLS2|SHMT1|APOA1|APOA4|CDO1|BHMT2|SDS|CYP2C9|CTH|SRD5A2|ALDOB|SLC27A2|BHMT|CPS1|FBP1|QDPR|ENO2|APOA5
6 Small molecule metabolic process	-2.197	PAH|ALDH6A1|EPHX1|SULT2A1|CYP3A5|FMOD|APOA1|APOA4|TTR|CDO1|BHMT2|SDS|DDC|CYP2C9|SCARB1|SRD5A2|PFKFB3|ADH4|SLC27A2|CYP3A7|BHMT|QDPR|HGD|HPD|ALDH1A1|UGT2B4|GLS2|PIPOX|SHMT1|TAT|CTH|SARDH|GNMT|ALDOB|GLDC|LDHB|APOF|CD44|CPS1|FBP1|ENO2|APOA5
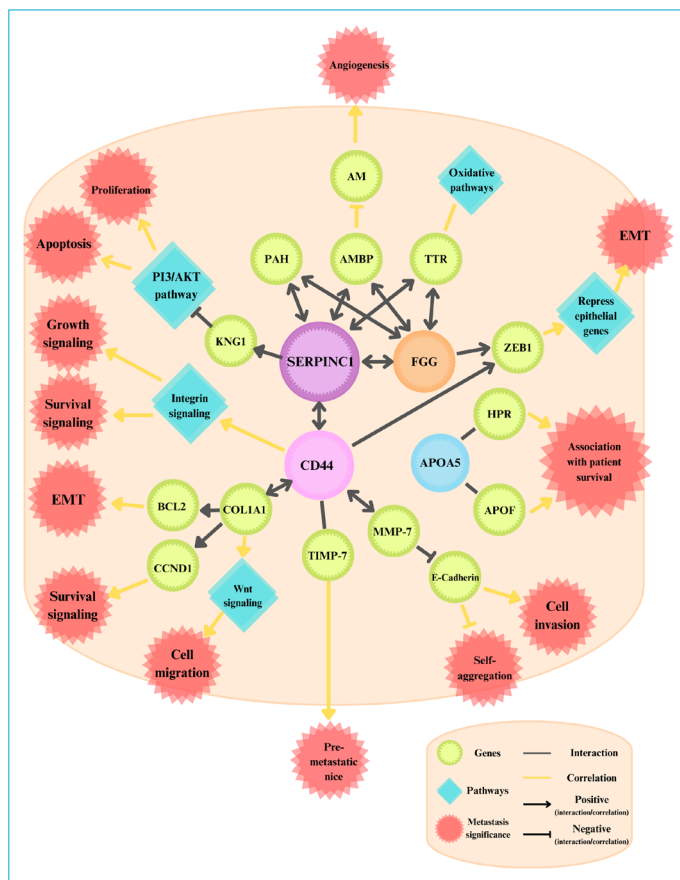
Figure 6. Schematic diagram presenting the interaction of four potential predictive biomarkers for advanced HCC (created using Canva).

significantly associated with the OS rate of HCC patients and may exert critical function in HCC progression. Enrichment denotes the possible mechanisms such as proliferation, survival signals, and EMT play essential roles in the development and progress of HCC. These findings offered novel insights into the current understanding of the pathways and interactions using PPI networks and suggests that these findings may have great clinical significance.

References

1. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. Nat Rev Dis Primers 2021;7:6. [CrossRef]

2. WHO. International Agency for Research on Cancer. Malaysia Globocan 2020. Available at: https://gco.iarc.fr/today/data/factsheets/populations/458-malaysia-fact-sheets.pdf. Accessed Jan 3, 2023.

3. Raihan R, Azzeri A, H Shabaruddin F, Mohamed R. Hepatocellular carcinoma in Malaysia and its changing trend. Euroasian J Hepatogastroenterol 2018;8:54–6. [CrossRef]

4. Wu W, He X, Andayani D, Yang L, Ye J, Li Y, et al. Pattern of distant extrahepatic metastases in primary liver cancer: a SEER based study. J Cancer 2017;8:2312–8.

5. Zhang M, Lv X, Jiang Y, Li G, Qiao Q. Identification of aberrantly methylated differentially expressed genes in glioblastoma multiforme and their association with patient survival. Exp Ther Med 2019;18:2140–52. [CrossRef]

6. Xiang ZL, Zeng ZC, Tang ZY, Fan J, He J, Zeng HY, et al. Potential prognostic biomarkers for bone metastasis from hepatocellular carcinoma. Oncologist 2011;16:1028–39.

7. Xiang ZL, Zeng ZC, Fan J, Tang ZY, Zeng HY, Gao DM. Gene expression profiling of fixed tissues identified hypoxia-inducible factor-1α, VEGF, and matrix metalloproteinase-2 as biomarkers of lymph node metastasis in hepatocellular carcinoma. Clin Cancer Res 2011;17:5463–72. [CrossRef]

8. Wei L, Lian B, Zhang Y, Li W, Gu J, He X, Xie L. Application of microRNA and mRNA expression profiling on prognostic biomarker discovery for hepatocellular carcinoma. BMC Genomics 2014;15:S13. [CrossRef]

9. Viacava Follis A. Centrality of drug targets in protein networks. BMC Bioinformatics. 2021;22:527. [CrossRef]

10. Grinchuk OV, Yenamandra SP, Iyer R, Singh M, Lee HK, Lim KH, et al. Tumor-adjacent tissue co-expression profile analysis reveals pro-oncogenic ribosomal gene signature for prognosis of resectable hepatocellular carcinoma. Mol Oncol 2018;12:89–113. [CrossRef]

11. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics 2007;23:1846–7. [CrossRef]

12. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse. J Open Source Softw 2019;4:1686. [CrossRef]

13. Pagès H, Carlson, M, Falcon S, Li N. Manipulation of SQLite-based annotations in Bioconductor. Available at: https://bioconductor.org/packages/AnnotationDbi. Accessed Jan 3, 2023.

14. 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 Res 2015;43:e47.

15. Law CW, Alhamdoosh M, Su S, Dong X, Tian L, Smyth GK, et al. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. F1000Res 2016;5:ISCB Comm J-1408. [CrossRef]

16. Dolgalev I. msigdbr: MSigDB gene sets for multiple organisms in a tidy data format. Available at: https://cran.r-project.org/package=msigdbr. Accessed Jan 3, 2023.

17. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A, et al. Fast gene set enrichment analysis. bioRxiv. 2021 Feb 1. Doi: 10.1101/060012. [Epub ahead of print].

18. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage Det al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–504. [CrossRef]

19. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: Network analysis and visualization of proteomics data. J Proteome Res 2019;18:623–32. [CrossRef]

20. Makrodimitris S, Reinders M, van Ham R. A thorough analysis of the contribution of experimental, derived and sequence-based predicted protein-protein interactions for functional annotation of proteins. PLoS One 2020;15:e0242723.

21. Assenov Y, Ramírez F, Schelhorn SE, Lengauer T, Albrecht M. Computing topological parameters of biological networks. Bioinformatics 2008;24:282–4. [CrossRef]

22. Therneau TM. Survival: Survival Analysis. Available at: https://cran.r-project.org/package=survival. Accessed Jan 3, 2023.

23. Kassambara A, Kosinski M, Biecek P, Fabian S. Survminer: Drawing survival curves using 'ggplot2'. Available at: https://cran.r-project.org/package=survminer. Accessed Jan 3, 2023.

24. Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res 2019;47:556–60. [CrossRef]

25. Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. PLoS One 2010;5:e13984.

26. Kucera M, Isserlin R, Arkhangorodsky A, Bader GD. AutoAnnotate: A Cytoscape app for summarizing networks with semantic annotations. F1000Res 2016;5:1717. [CrossRef]

27. Xu D, Wu J, Dong L, Luo W, Li L, Tang D, et al. Serpinc1 Acts as a tumor suppressor in hepatocellular carcinoma through inducing apoptosis and blocking macrophage polarization in an ubiquitin-proteasome manner. Front Oncol 2021;11:738607.

28. Kurata M, Okajima K, Kawamoto T, Uchiba M, Ohkohchi N. Antithrombin reduces reperfusion-induced hepatic metastasis of colon cancer cells. World J Gastroenterol 2006;12:60–5.

29. Iwako H, Tashiro H, Amano H, Tanimoto Y, Oshita A, Kobayashi T, et al. Prognostic significance of antithrombin III levels for

30. Rullier A, Senant N, Kisiel W, Bioulac-Sage P, Balabaud C, Le Bail B, et al. Expression of protease-activated receptors and tissue factor in human liver. Virchows Arch 2006;448:46–51.

31. Korita PV, Wakai T, Shirai Y, Matsuda Y, Sakata J, Cui X, et al. Overexpression of osteopontin independently correlates with vascular invasion and poor prognosis in patients with hepatocellular carcinoma. Hum Pathol 2008;39:1777–83. [CrossRef]

32. Noe JT, Mitchell RA. MIF-dependent control of tumor immunity. Front Immunol 2020;11:609948. [CrossRef]

33. Mawhinney L, Armstrong ME, O' Reilly C, Bucala R, Leng L, Fingerle-Rowson G, et al. Macrophage migration inhibitory factor (MIF) enzymatic activity and lung cancer. Mol Med 2015;20:729–35. [CrossRef]

34. Guda MR, Rashid MA, Asuthkar S, Jalasutram A, Caniglia JL, Tsung AJ, et al. Pleiotropic role of macrophage migration inhibitory factor in cancer. Am J Cancer Res 2019;9:2760–73.

35. Meyer-Siegler KL, Cox J, Leng L, Bucala R, Vera PL. Macrophage migration inhibitory factor anti-thrombin III complexes are decreased in bladder cancer patient serum: Complex formation as a mechanism of inactivation. Cancer Lett 2010;290:49–57. [CrossRef]

36. Takafuji V, Forgues M, Unsworth E, Goldsmith P, Wang XW. An osteopontin fragment is essential for tumor cell invasion in hepatocellular carcinoma. Oncogene 2007;26,6361–71.

37. Deyell M, Garris CS, Laughney AM. Cancer metastasis as a non-healing wound. Br J Cancer 2021;124:1491–502. [CrossRef]

38. Priebatsch KM, Kvansakul M, Poon IK, Hulett MD. functional regulation of the plasma protein histidine-rich glycoprotein by Zn2+ in settings of tissue injury. Biomolecules 2017;7:22.

39. Matus CE, Bhoola KD, Figueroa CD. Kinin b1 receptor signaling in skin homeostasis and wound healing. Yale J Biol Med 2020;93:175–85.

40. Mishra A, Suman KH, Nair N, Majeed J, Tripathi V. An updated review on the role of the CXCL8-CXCR1/2 axis in the progression and metastasis of breast cancer. Mol Biol Rep 2021;48:6551–61.

41. Arnold SA, Brekken RA. SPARC: a matricellular regulator of tumorigenesis. J Cell Commun Signal 2009;3:255–73. [CrossRef]

42. Muresan XM, Sticozzi C, Belmonte G, Savelli V, Evelson P, Valacchi G, et al. Modulation of cutaneous scavenger receptor B1 levels by exogenous stressors impairs "in vitro" wound closure. Mech Ageing Dev 2018;172:78–85. [CrossRef]

43. Berger H, Wodarz A, Borchers A. PTK7 faces the Wnt in development and disease. Front Cell Dev Biol 2017;5:31.

44. Ricca TI, Liang G, Suenaga AP, Han SW, Jones PA, Jasiulionis MG. Tissue inhibitor of metalloproteinase 1 expression associated with gene demethylation confers anoikis resistance in early phases of melanocyte malignant transformation. Transl Oncol 2009;2:329–40. [CrossRef]

45. Löffek S, Schilling O, Franzke CW. Series "matrix metalloproteinases in lung health and disease": Biological role of

matrix metalloproteinases: a critical balance. Eur Respir J 2011;38:191–208.

46. Shen H, Wu H, Sun F, Qi J, Zhu Q. A novel four-gene of iron metabolism-related and methylated for prognosis prediction of hepatocellular carcinoma. Bioengineered 2021;12:240–51.

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

48. Gacche RN, Meshram RJ. Targeting tumor micro-environment for design and development of novel anti-angiogenic agents arresting tumor growth. Prog Biophys Mol Biol 2013;113:333–54. [CrossRef]

49. Sharma M, Khan S, Rahman S, Singh LR. The extracellular protein, transthyretin is an oxidative stress biomarker. Front Physiol 2019;10:5. [CrossRef]

50. Xu J, Fang J, Cheng Z, Fan L, Hu W, Zhou F, et al. Overexpression of the Kininogen-1 inhibits proliferation and induces apoptosis of glioma cells. J Exp Clin Cancer Res 2018;37:180.

51. Chen C, Zhao S, Karnad A, Freeman JW. The biology and role of CD44 in cancer progression: therapeutic implications. J Hematol Oncol 2018;11:64. [CrossRef]

52. Senbanjo LT, Chellaiah MA. CD44: A multifunctional cell surface adhesion receptor is a regulator of progression and metastasis of cancer cells. Front Cell Dev Biol 2017;5:18.

53. Hong W, Guan KL. The YAP and TAZ transcription co-activators: key downstream effectors of the mammalian Hippo pathway. Semin Cell Dev Biol 2012;23:785–93. [CrossRef]

54. Shin E, Kim J. The potential role of YAP in head and neck squamous cell carcinoma. Exp Mol Med 2020;52:1264–74

55. Zhang J, He X, Wan Y, Zhang H, Tao T, Zhang M, et al. CD44 promotes hepatocellular carcinoma progression via upregulation of YAP. Exp Hematol Oncol 2021;10:54.

56. Preca BT, Bajdak K, Mock K, Sundararajan V, Pfannstiel J, Maurer J, et al. A self-enforcing CD44s/ZEB1 feedback loop maintains EMT and stemness properties in cancer cells. Int J Cancer 2015;137:2566–77. [CrossRef]

57. Li J, Zhang Y, Ruan R, He W, Qian Y. The novel interplay between CD44 standard isoform and the caspase-1/IL1B pathway to induce hepatocellular carcinoma progression. Cell Death Dis 2020;11:961. [CrossRef]

58. Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. Cell 2009;139:871–90. [CrossRef]

59. Paszek MJ, DuFort CC, Rossier O, Bainer R, Mouw JK, Godula K, et al. The cancer glycocalyx mechanically primes integrin-mediated growth and survival. Nature 2014;511:319–25.

60. Gonzalez-Avila G, Sommer B, Mendoza-Posada DA, Ramos C, Garcia-Hernandez AA, Falfan-Valencia R. Matrix metalloproteinases participation in the metastatic process and their diagnostic and therapeutic applications in cancer. Crit Rev Oncol Hematol 2019;137:57–83. [CrossRef]

61. Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts

of extracellular matrix remodelling in tumour progression and metastasis. Nat Commun. 2020 Oct 9;11(1):5120.

62. Seubert B, Grünwald B, Kobuch J, Cui H, Schelter F, Schaten S, et al. Tissue inhibitor of metalloproteinases (TIMP)-1 creates a premetastatic niche in the liver through SDF-1/CXCR4-dependent neutrophil recruitment in mice. Hepatology 2015;61:238–48. [CrossRef]

63. Li X, Sun X, Kan C, Chen B, Qu N, Hou N, et al. COL1A1: A novel oncogenic gene and therapeutic target in malignancies. Pathol Res Pract 2022 236:154013. [CrossRef]

64. Zhang Z, Wang Y, Zhang J, Zhong J, Yang R. COL1A1 promotes metastasis in colorectal cancer by regulating the WNT/PCP pathway. Mol Med Rep 2018;17:5037–42.

65. Simons M, Mlodzik M. Planar cell polarity signaling: from fly development to human disease. Annu Rev Genet 2008;42:517–40. [CrossRef]

66. Harrold JM, Ramanathan M, Mager DE. Network-based approaches in drug discovery and early development. Clin Pharmacol Ther 2013;94:651–8. [CrossRef]

67. Mosesson MW. Fibrinogen and fibrin structure and functions. J Thromb Haemost 2005;3:1894–904. [CrossRef]

68. Zhang X, Wang F, Huang Y, Ke K, Zhao B, Chen L, et al. FGG promotes migration and invasion in hepatocellular carcinoma cells through activating epithelial to mesenchymal transition. Cancer Manag Res 2019;11:1653–65.

69. Yue H, Hu Z, Hu R, Guo Z, Zheng Y, Wang Y, et al. ALDH1A1 in cancers: Bidirectional function, drug resistance, and regulatory mechanism. Front Oncol 2022;12:918778.

70. Ren L, Yi J, Li W, Zheng X, Liu J, Wang J, et al. Apolipoproteins and cancer. Cancer Med 2019;8:7032–43. [CrossRef]

71. Li H, Long J, Xie F, Kang K, Shi Y, Xu W, et al. Transcriptomic analysis and identification of prognostic biomarkers in cholangiocarcinoma. Oncol Rep 2019;42:1833–42. [CrossRef]

72. Ma XL, Gao XH, Gong ZJ, Wu J, Tian L, Zhang CY, et al.. Apolipoprotein A1: a novel serum biomarker for predicting the prognosis of hepatocellular carcinoma after curative resection. Oncotarget 2016;7:70654–68. [CrossRef]

73. Hoshida Y, Fuchs BC, Tanabe KK. Prevention of hepatocellular carcinoma: potential targets, experimental models, and clinical challenges. Curr Cancer Drug Targets 2012;12:1129–59.

74. Tai CS, Lin YR, Teng TH, Lin PY, Tu SJ, Chou CH, et al. Haptoglobin expression correlates with tumor differentiation and five-year overall survival rate in hepatocellular carcinoma. PLoS One 2017;12:e0171269. [CrossRef]

75. Wang YB, Zhou BX, Ling YB, Xiong ZY, Li RX, Zhong YS, et al. Decreased expression of ApoF associates with poor prognosis in human hepatocellular carcinoma. Gastroenterol Rep (Oxf) 2019;7:354–60. [CrossRef]