Original Article

A genetic database can be utilized to identify potential biomarkers for biphenotypic hepatocellular carcinoma-cholangiocarcinoma

Shaffer R. S. Mok, Sachin Mohan, Navjot Grewal, Adam B. Elfant, Thomas A. Judge

Division of Gastroenterology and Liver Diseases, Department of Medicine, Cooper Medical School of Rowan University, MD Anderson Cancer Center at Cooper, Mount Laurel, NJ, USA

Contributions: (I) Conception and design: SR Mok; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: SR Mok, S Mohan, N Grewal; (V) Data analysis and interpretation: SR Mok, S Mohan, N Grewal; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Shaffer R. S. Mok, MD, MBS. Division of Gastroenterology and Liver Diseases, Department of Medicine, Cooper University Hospital, MD Anderson Cancer Center at Cooper, 501 Fellowship Road, Suite 101, Mount Laurel, NJ 08054, USA. Email: mok.shaffer@gmail.com.

Background: Biphenotypic hepatocellular carcinoma-cholangiocarcinoma (HCC-CC) is an uncommon primary liver neoplasm. Due to limitations in radiologic imaging for the diagnosis of this condition, biopsy is a common method for diagnosis, which is invasive and holds potential complications. To identify alternative means for obtaining the diagnosis and assessing the prognosis of this condition, we evaluated biomarkers for biphenotypic HCC-CC using a genetic database.

Methods: To evaluate the genetic associations with each variable we utilized GeneCards®, The Human Gene Compendium (http://www.genecards.org). The results of our search were entered into the Pathway Interaction Database from the National Cancer Institute (PID-NCI) (http://pid.nci.nih.gov), to generate a biomolecule interaction map.

Results: The results of our query yielded 690 genes for HCC, 98 genes for CC and 50 genes for HCC-CC. Genes depicted in this analysis demonstrate the role of hormonal regulation, embryonic development, cell surface adhesion, cytokeratin stability, mucin production, metalloproteinase regulation, Ras signaling, metabolism and apoptosis. Examples of previously described markers included hepatocyte growth factor (HGF), mesenchymal epithelial transition (MET) and Kirsten rat sarcoma viral oncogene homolog (KRAS). Novel markers included phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha (PIK3CA), GPC3, choline kinase alpha (CHKA), prostaglandin-endoperoxide synthase 2 (PTGS2), telomerase reverse transcriptase (TERT), myeloid cell leukemia 1 (MCL1) and N-acetyltransferase 2 (NAT2).

Conclusions: GeneCards is a useful research tool in the genetic analysis of low frequency malignancies. Utilizing this tool we identified several biomarkers are methods for diagnosing HCC-CC. Finally, utilizing these methods, HCC-CC was found to be predominantly a subtype of CC.

Keywords: Hepatocellular carcinoma (HCC); cholangiocarcinoma (CC); hepatocellular carcinoma-cholangiocarcinoma (HCC-CC); genetic; biomarker

Submitted Feb 07, 2016. Accepted for publication Mar 09, 2016.
doi: 10.21037/jgo.2016.04.01
View this article at: http://dx.doi.org/10.21037/jgo.2016.04.01

Introduction

Biphenotypic hepatocellular carcinoma-cholangiocarcinoma (HCC-CC) comprises an estimated 1–6.5% of primary liver neoplasms (1). Although this entity does demonstrate pathologic features of its well-established biphenotypic counterparts, the characteristics of this neoplasm make diagnosis challenging using conventional radiologic imaging and serologic markers. Moreover, accurate prognostic information is affected by the low frequency of this tumor, which restricts available clinical information even within

large medical centers (2-20).

There are numerous genetic and risk factors important for hepatocellular carcinoma (HCC) and cholangiocarcinoma (CC) separately, which have yet to be addressed in HCC-CC (17-22). Prior studies have examined the role of cytoskeletal stability, apoptosis, and the inflammatory cascade in HCC-CC (4-16). Additionally, as HCC-CC poses a diagnostic dilemma for radiologists, tumor specific biomarker may assist in the diagnosis of this neoplasm and prognostic assessment (23,24).

In this study, we evaluate the utility of a genetic database to identify potential biomarkers for biphenotypic HCC-CC using shared genetic characteristics.

Methods

Variables

We initially evaluated the pathologic subtypes of HCC-CC, HCC and CC (25,26). Pathologic subtypes for HCC-CC included classical and stem-cell. Pathologic subtypes for HCC included: fibrolamellar, scirrhous, sarcomatoid and lympho-epithelial. Subtypes for CC included: intraductal papillary, intestinal-type, clear, squamous and small cell.

After performing our literature search, we identified risk factors relevant to HCC-CC, HCC alone and CC alone (12-15,17-22). Among these risk factors, cirrhosis, hepatitis B virus (HBV) and hepatitis C virus (HCV) viral infections are evaluated risk factors for HCC-CC (12-15,17-20). For HCC alone, the aforementioned risk factors were included as were the following: portal hypertension, alcoholic fatty liver, aflatoxin, peliosis hepatitis, autoimmune hepatitis (AIH), primary biliary cirrhosis (PBC), granulomatous hepatitis, non-alcoholic fatty liver disease (NAFLD), non-alcoholic steatohepatitis (NASH), hemochromatosis (HCM), glycogen storage disease, Wilson's disease, porphyria cutanea tarda (PCT), alpha-1 antitrypsin, tyrosinemia, portal vein thrombosis, Budd-Chiari syndrome. For CC we evaluated the following variables: primary sclerosing cholangitis (PSC), cystic disease of the liver, biliary cyst, choledochal cyst, Caroli disease, other congenital malformation of the bile ducts, other congenital malformations of the gall bladder, cholangitis, schistosomiasis, opisthorchiasis, clonochiasis, recurrent cholangitis, and biliary stricture.

Genetic database

To evaluate the genetic associations with each variable we utilized GeneCards®, The Human Gene Compendium (http://www.genecards.org) (27-33). Our initial search was performed on HCC alone, CC alone, and finally HCC-CC. Next, we evaluated the risk factors mentioned above utilizing the same compendium database. We then evaluated each possibly biomarkers in accordance to HCC alone, CC alone, and finally HCC-CC. All results were recorded and each gene was scrutinized manually, independent of the search results.

During our manual assessment of each gene we evaluated gene function, pathway and interaction, association with other genes, cellular location, genomic location, and existing therapeutic targets. The results of our search were then entered into the Pathway Interaction Database from the National Cancer Institute (PID-NCI) (http://pid.nci.nih.gov), to generate a biomolecule interaction map.

Results

Overall genetic characteristics

The results of our query yielded 690 genes for HCC, 98 genes for CC and 50 genes for HCC-CC. A summary of the search results for these searches can be visualized in *Table S1*. Genes depicted in this analysis demonstrate the role of hormonal regulation, embryonic development, cell surface adhesion, cytokeratin stability, mucin production, metalloproteinase regulation, Ras signaling, metabolism and apoptosis. *Table 1* depicts the relationship between these genes, the genomic and cellular location. These genes were integrated into a PID-NCI biomolecule interaction map (*Figure S1*), demonstrating an overview of the interactions between each gene for HCC-CC.

Gamma-glutamyl transpeptidase (GGT) appeared to have the highest number of associated risk factors (20), followed by the cytokeratin-related genes (KRT7, 8, 9, 18 and 19) (4, 5, 8, 7 respectively). Alkaline phosphatase (ALP) related genes (*ALPL*, *ALPP*, *ALPPL2*) also had a high number of associated etiologies (4, 7 and 4 respectively) (*Table 1*). The genetic location for genes involved in hormonal regulation demonstrated a genetic location of 11p15. Cell adhesion molecules hepatocyte growth factor (HGF) and mesenchymal epithelial transition (MET) demonstrated a genetic location of 7p and genes involved in mucin production 11p15.5. Overall cytokeratin genes demonstrated a genetic location of 12q13, with the exception of KRT19, with a location of 17q21.

Table 1 Genes and potential biomarkers identified for HCC-CC, HCC, CC and associated conditions

Conditions	Total genes
HCC-CC	50
HCC	690
HCC pathologic subtypes	
Fibrolamellar HCC	5
Scirrhous HCC	2
Sarcomatoid HCC	11
Lympho-epithelial-like HCC	94
HCC risk factors	
Cirrhosis	352
Portal hypertension	30
Infectious	
Hepatitis C	272
Hepatitis B	293
Toxin	
Alcoholic fatty liver	30
Aflatoxin	6
Peliosis hepatis	2
Autoimmune	
Autoimmune hepatitis	206
Primary biliary cirrhosis	118
Granulomatous, hepatitis	11
Metabolic	
Fatty liver disease	125
Nonalcoholic steatohepatitis	9
Hemochromatosis	52
Glycogen storage diseases	52
Wilson disease	26
Porphyria cutanea tarda	20
Alpha-1, antitrypsin	20
Tyrosinemia	15
Vascular	
Portal vein thrombosis	9
Budd-Chiari syndrome	5
CC	98
CC pathologic subtypes	
Intraductal papillary CC	2
Intestinal-type CC	8
Clear cell CC	12
Signet-ring CC	0
Squamous Cell CC	54
Small cell CC	39
CC risk factors	
Primary sclerosing cholangitis	41
Congenital	
Cystic disease of liver	97
Biliary cyst	20
Choledochal cyst	8
Caroli disease	2
Other congenital malformations of bile ducts	1
Other congenital malformations of gallbladder	1
Infectious	
Cholangitis	58
Schistosomiasis	43
Opisthorchiasis	12
Clonorchiasis	12
Recurrent cholangitis	7
Anatomic	
Biliary stricture	2

HCC-CC, hepatocellular carcinoma-cholangiocarcinoma; HCC, hepatocellular carcinoma; CC, cholangiocarcinoma.

Cell adhesion molecules (HGF and MET) are located in the extracellular matrix (ECM) and cell membrane. Cytokeratin molecules are expressed in the cellular membrane, Golgi apparatus, and nucleus, except for KRT19, located in the ECM and cellular membrane. Mucin production genes MUC2 and MUC5AC localize to the ECM. Metalloproteinase, membrane metallo-endopeptidase (MME) and reversion-inducing-cysteine-rich protein with kazal motifs (RECK), were located in the ECM and cell membrane, while MMP7 and tissue inhibitor of metalloproteinases 3 (TIMP3) were located in the ECM and Nucleus. Gene *ALPL*, *ALPP*, *ALPPL2* and *GGT* were located in all cellular locations.

The relationship of these genes to each other is summarized in *Table 1* and *Figure S1*. There appears to be a linkage between the embryonic genes and secretin (SCT). KRT18 also demonstrated some associations with the cell adhesion genes. Not surprising, Ras-signaling genes have a relation with apoptosis genes which in turn overlap with cell adhesion, cytokeratin, metabolism and metalloproteinase genes.

Pathologic subtypes

A complete listing of all genes for each pathologic subtype

is presented in *Table S2*. For each pathologic subtype of HCC-CC, no novel genes could be identified, likely due to extreme search specificity. A detailed examination of HCC subtypes was performed. Fibrolamellar HCC demonstrated no genetic overlap with HCC-CC. Scirrhous HCC possessed two novel genes including MET which shared a common bridge with HCC-CC. Sarcomatoid HCC demonstrated carcinoembryonic antigen-related cell adhesion molecule 3 (CEACAM3) and CHD1 in common with HCC-CC. Among the pathologic subtypes of HCC, lympho-epithelial HCC had the most genetic overlap with HCC-CC possessing both cadherin (CDH1) and MET, along with ALPP, ALPL, ALPPL2, MUC2, MUC5AC, NAT2, PTGS2, SCT, catenin (cadherin-associated protein), alpha 1 (CTNNA1) and vascular endothelial growth factor C (VEGFC). This totaled 12 genes or 24% overlap with HCC-CC.

Next, the pathologic subtypes of CC were evaluated for overlap with HCC-CC. Intraductal papillary CC shares SMAD family member 4 (SMAD4), Intestinal-type CC possessed Kirsten rat sarcoma viral oncogene homolog (KRAS), KRT7, phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha (PIK3CA) and tumor protein p53 (TP53) (8% overlap) in common with HCC-CC. Numerically, signet ring CC was similar to the intraductal papillary subtype, differentially expressing 4 genes (8%) in common with HCC-CC (KRT7, MUC2, CDH1 and CEACAM5). Clear cell CC, small cell CC, and squamous cell CC share 20%, 54%, and 78% gene overlap with biphenotypic HCC-CC, respectively. Novel genes expressed by both clear cell CC and HCC-CC include 8-oxoguanine DNA glycosylase (OGG1), CDH1, fragile histidine triad (FHIT) and SMAD family member 4 (SMAD4) among others for 20% overlap with HCC-CC. Genes shared by both HCC-CC and squamous cell CC involve functions such as cytokeratin stability, apoptosis, Ras-signaling, cell adhesion, embryonic development, metalloproteinase, metabolism and tumor necrosis factor alpha pathways.

Hepatocellular carcinoma (HCC) risk factors

Known risk factors for HCC were evaluated with respect to HCC-CC. *Table 2* depicts the total number of genes available for each risk factor for HCC when evaluated separately. Cirrhosis had the highest number of genes at 352, followed by 293 for HBV and 272 for HCV. Peliosis hepatitis had the least number of associated genes.

We compared the genes in each risk factor with HCC and HCC-CC. The results of this comparison are listed in *Table 3*. Of all the listed risk factors, cirrhosis had the highest number of gene associations with HCC [156] and HCC-CC (19 of 50 or 38%). These genes are summarized in *Table S3*. Most commonly, cirrhosis and HCC-CC differentially expressed genes related to Ras signaling [KRAS, V-Raf-1 murine leukemia viral oncogene homolog 1 (RAF1), RAASF1] and cellular metabolism [ALPP, choline kinase alpha (CHKA), GGT1 and NAT2]. Hepatitis B had 133 genes shared with HCC and 12 with HCC-CC (24%). Genes shared between HBV and HCC-CC were alpha-fetoprotein (AFP), MET, KRT8, KRAS, RAF1, genes for metabolism and apoptosis. Similar to HBV, HCV had a high number of comparable genes when compared with HCC [115] and HCC-CC (9.18%). Similar to HBV, HCV demonstrated genes AFP, MET, KRAS and RAF1. Unlike HBV, HCV also demonstrated KRT18 (v. KRT8 for HBV), MMP7, TIMP3 and only GGT1 for metabolism.

In addition to HCV, autoimmune disease such as AIH and PBC demonstrated several genes in common with HCC (88 for AIH and 42 PBC) and HCC-CC [5 (10%), 9 (18%) respectively]. AIH demonstrated GPC3, KRT8, 18, KRAS, GGT1 in common with HCC-CC. When comparing genes for PBC and HCC-CC, this etiology possessed SCT, HGF, KRT7, KRT19, MUC5AC, MMP7, ALPP, CHKA and myeloid cell leukemia 1 (MCL1). Fatty liver disease had 4 genes in common with HCC-CC and 41 with HCC.

Cholangiocarcinoma (CC) risk factors

When evaluating the genes present in CC, our analysis yielded 98 genes. A summary of CC-related risk factors is summarized in *Table 2*. The highest number of genes among CC risk factors included 97 for cystic disease of the liver, 58 for cholangitis, 43 for schistosomiasis and 41 for PSC. The lowest number of genes present was found in other congenital malformations of the bile ducts and gall bladder.

As in HCC, we then compared the CC risk factors with HCC-CC, depicted in *Table 3*. Among these etiologies, cystic disease of the liver demonstrated 17 genes in common with CC and 13 of 50 with HCC-CC (26%). Such genes included all cytokeratin-related genes along with AFP, CEACAM3, SCT, MME, Ras association (RalGDS/AF-6) domain family member (RASSF1) and NAT2. Biliary cysts, demonstrated 10 genes in common with CC and 8 (16%) with HCC-CC. These results can be summarized in *Table S3*.

The next most common etiology for HCC-CC from CC was cholangitis. This risk factor demonstrated 8 genes in common with CC and 7 (14%) with HCC-CC. This risk

Table 2 Genes and biomarkers identified for HCC, CC and HCC-CC in correlation with pathologic subtypes and risk factors

Variables	HCC-CC	HCC	CC
Pathologic subtypes			
Fibrolamellar HCC	0	5	0
Scirrhous HCC	1	2	2
Sarcomatoid HCC	2	11	2
Lympho-epithelial-like HCC	12	94	12
Intraductal papillary CC	1	1	2
Intestinal-type CC	4	4	8
Clear cell CC	10	10	12
Signet-ring CC	4	4	5
Squamous cell CC	39	39	54
Small cell CC	27	27	39
Risk factor			
Cirrhosis	19	156	21
Portal hypertension	1	8	1
Infectious			
Hepatitis C	9	115	12
Hepatitis B	12	133	17
Cholangitis	7	24	8
Schistosomiasis	0	10	0
Opisthorchiasis	1	3	1
Clonorchiasis	0	1	0
Recurrent cholangitis	1	2	1
Metabolic			
Non-alcoholic fatty liver disease	4	41	4
Nonalcoholic steatohepatitis	1	6	1
Hemochromatosis	0	12	0
Glycogen storage diseases	0	6	0
Wilson disease	0	10	0
Porphyria cutanea tarda	0	4	0
Alpha-1 antitrypsin deficiency	0	4	0
Tyrosinemia	1	4	1
Congenital			
Cystic disease of liver	13	47	17
Biliary cyst	8	10	10
Choledochal cyst	2	3	2
Caroli disease	0	0	0
Other congenital malformations of bile ducts	0	0	1
Other congenital malformations of gallbladder	1	1	1
Autoimmune			
Autoimmune hepatitis	5	88	8
Granulomatous hepatitis	0	4	0
Primary biliary cirrhosis	9	42	10
Primary sclerosing cholangitis	2	15	3
Carcinogen			
Alcoholic fatty liver	2	15	2
Alcohol	8	49	8
Aflatoxin	2	6	2
Peliosis hepatis	0	0	0
Vascular			
Portal vein thrombosis	1	4	1
Budd-Chiari syndrome	0	0	0
Anatomic			
Biliary stricture	0	2	0

HCC-CC, hepatocellular carcinoma-cholangiocarcinoma; HCC, hepatocellular carcinoma; CC, cholangiocarcinoma.

factor appeared to demonstrate several metabolic genes (ALPL, ALPP, ALPPL2 and GGT1), as well as KRT19 and MUC2. PSC demonstrated 2 (4%) genes in common with HCC-CC (KRT19 and GGT1) *vs.* 3 with CC alone. Choledochal cyst as a risk factor also demonstrated 2 (4%) genes in common with HCC-CC (SCT and PTGS2). Opisthochiasis, recurrent cholangitis, and other congenital malformations of the gall bladder demonstrated only a single gene in common with HCC-CC. All other CC-related etiologies had no risk factors in common with HCC-CC depicted in *Table 3* and *Table S3*.

Discussion

Biphenotypic HCC-CC is a unique hepatic neoplasm expressing features of both HCC and CC. This tumor poses a diagnostic dilemma, utilizing radiologic imaging alone, and thus biomarkers would prove a valuable resource in this condition (23,24). To identify potential biomarkers that could assist clinicians in the diagnosis of this unique primary liver cancer, we utilized GeneCards®, the Human Gene Compendium.

Prior studies have validated the use of this database in the study of several medical conditions (27-33). Specifically, this compendium has assisted researchers in identifying

Table 3 Description of genes and potential biomarkers identified for HCC-CC and relationship to each other

Gene function	Gene name	Etiologies with gene (total)	Genetic location	Cellular location	Function
Embryonic	AFP	6	4q13.3	ECM and cytoskeleton	Plasma protein produced by yolk sac
	CEACAM3	2	19q13	Cell membrane	Transmembrane signaling molecule to direct phagocytosis of bacteria
	GPC3	2	Xq26	Cell membrane, cytosol, golgi, lysozome	Core protein anchored to the cytosolic membrane
Hormonal	CCKBR	2	11p15.4	Cell membrane	Receptor for CCKB
	SCT	7	11p15.5	ECM	Endocrine hormone secretin
Cell adhesion	HGF	5	7q21.1	ECM to cell membrane	Tyrosine kinase cellular signaling
	MET	4	7q31	ECM to cell membrane	Tyrosine kinase cellular signaling
Cytokeratin	KRT7	4	12q13.13	Cellular membrane, golgi, nucleus	Neutral Proteins involved in differentiation
	KRT8	5	12q13.13	Cellular membrane, golgi, nucleus	Neutral Proteins involved in differentiation
	KRT18	8	12q13	Cellular membrane, golgi, nucleus	Neutral Proteins involved in differentiation
	KRT19	7	17q21.2	ECM to cell membrane	Neutral Proteins involved in differentiation
Mucin production	MUC2	2	11p15.5	ECM, golgi	High molecular weight glycoprotein in gut barrier function
	MUC5AC	3	11p15.5	ECM	Protein coding gene for ECM
Metalloproteins	MME	2	3q25.2	ECM, cell membrane	Endopeptidase that cleaves several glycoproteins
	MMP7	5	11q21	ECM, nucleus	Breakdown of ECM
	RECK	3	9p13.3	ECM, cell membrane	Acted upon by cancer leading to negative regulation of MMP
	TIMP3	4	22q12.3	ECM, nucleus	Peptidases for degradation of ECM
Ras signaling	KRAS	5	12p12.1	Cell membrane, cytosol, nucleus	Promoter leading GTPase activation
	RAF1	3	3p25	Mitochondrion, cytosol, nucleus	MAP kinase > Ras signaling
	RASSF1	2	3p21.3	Cytoskeleton, nucleus	Tumor suppressor leading to hypermethylation of CpG islands, accumulation of cyclin-D that causes cell cycle arrest
Metabolism	ALPL	4	–	All	Metabolism
	ALPP	7	2q37.1	All	Metabolism
	ALPPL2	4	–	All	Metabolism
	CHKA	3	11q13.1	Cytoskeleton, cytosol	Synthesis of phosphatidylcholine
	GGT1	20	–	All	Metabolism
	NAT2	5	8p22	Cytosol	Deactivate arylamine and hydrazine drugs and carcinogens, also N-Acetyltransferase
	PTGS2	2	1q25	ER, nucleus	COX, prostaglandin biosynthesis
Apoptosis	MCL1	3	1q21	Mitochondrion, cytosol, Nucleus	Encodes anti-apoptotic protein (BCL-2)
	PIK3CA	3	3q36.3	Cell membrane, cytosol	Oncogene ATP to phosphorylate PtdIns
	TERT	3	5p15.33	Nucleus	Maintains length of telomeres

HCC-CC, hepatocellular carcinoma-cholangiocarcinoma; ECM, extracellular membrane; ER, endoplasmic reticulum; MMP, matrix metalloproteinase; GTPase, guanyl triphosphatase; COX, cyclo-oxygenase; BCL-2, B-cell lymphoma 2; ATP, adenosine tri-phosphate.

genes vital to the prognosis and pathogenesis of multiple malignancies, as well as non-neoplastic liver diseases (30-33). A similar concept has been employed in the setting of HCC-CC in a recent abstract analyzed the genetic composition of 15 pathologic specimens of HCC-CC (34). These investigators descriptively identified genetic markers present in their specimens and compared them with HCC, CC and HCC-CC information.

Our study identifies KRAS, MET, PIK3CA and TP53 as potential biomarkers for HCC-CC. Each of these genes is involved in Ras-signaling, a process vital to oncogenesis. MET factor and HGF encode tyrosine kinases, which allow for further cell signaling from the ECM into the cytoplasm. Furthermore, these gene products also interact with PI3K, which via signaling cascades also activates Ras. Both MET and PIK3CA have been evaluated in prior study to help determine prognostic and chemotherapeutic information in non-small cell lung cancer, breast, gastric, among other cancers (35,36). The MET-HGF complex has been evaluated previously in HCC, CC and HCC-CC, the latter via a small study of 30 pathologic specimens (37). This study demonstrated excess expression of c-met solely in the CC portion of combined HCC-CC.

A similar specificity for the CC portion of biphenotypic HCC-CC has also been described in the literature with regard to KRAS (9,38). The KRAS oncogene has previously been suggested as a clinical biomarker for a variety of abdominal neoplasms, including HCC-CC (34,38). Cancers other than HCC-CC associated with differential expression of KRAS, include colorectal and pancreatic cancer. In colon cancer, KRAS expression has been shown to provide useful information regarding treatment strategies (39).

Biomarkers previously associated with HCC-CC were confirmed in this study and included: AFP, CEA, GGT1, cytokeratin, mucin and metalloproteinase genes (7-10,12-15,19). Among these biomarkers, GGT, AFP, and CEA have been non-specific in the setting of HCC-CC (25). In contrast, cytokeratin, mucin production, and metalloproteinase molecules appear to be valuable to the diagnosis and prognosis of HCC-CC (7-10,12-15,19,25). Specifically, they appear to be useful in differentiating between classical and stem cell subtypes of HCC-CC. Such determinations can be facilitated utilizing cytokeratin signaling biomarkers (7-10,12-15,19,25). These serologic markers also appear to be useful in colorectal and breast neoplasms, with similar application (40).

The overlap between HCC-CC and HCC's pathologic subtypes was determined in our analysis to range between 0–24%, while CC subtypes demonstrated an overlap of 2–78%. Given the significantly larger genetic overlap between the subtypes of CC and HCC-CC, it would appear that HCC-CC is more likely a pathologic subtype of CC with features of HCC. Recently, overall survival for HCC and HCC-CC (41,42) has been assessed using the Surveillance, Epidemiology, and End Results (SEER) database. HCC-CC demonstrated an overall 1-, 3-, and 5-year survival rates of 26.5%, 12.4% and 9.2% (43). Additional analysis evaluating post-transplant prognosis documented a survival rate of 46% for HCC-CC as compared with 78% survival for HCC (44). When reviewing these results with prior studies of survival results for HCC and CC, the survival percentiles of HCC-CC are more consistent with CC (41-45). For CC, the survival after transplantation has been estimated at 22–42% of CC and 0-18% for CC without transplantation. The prognostic comparison and genetic overlap with CC suggests a shared pathogenesis for HCC-CC and CC.

Novel serologic markers identified in our analysis involve pathogenic roles in embryonic development, apoptosis and metabolism. The first of these unique genes, GPC3, codes for a cell surface heparin sulfate proteoglycan. This molecule inhibits the dipeptidyl peptidase activity of dipeptidyl peptidase-4 (DPP4), vital in apoptosis and growth regulation of several tissues. This gene has previously been utilized as a serologic marker for prognosis of HCC after curative resection (46). Down-regulation of this gene is correlated with uncontrolled cellular growth. Expression of this gene in the setting of HCC-CC has yet to be evaluated.

Our biomarker analysis also identifies were two genes involved in cellular metabolism. Choline kinase alpha (CHKA), encodes for enzymes that regulate the synthesis of phosphatidylcholine. The second gene, PTGS2, in combination with CHKA, is an additional metabolic target regulating biosynthesis of cyclo-oxygenase 2 (COX-2). The COX-2 enzyme is involved in inflammation and mitogenesis and has been implicated as a serologic marker for predicting the prognosis of prostate, breast and several other malignant conditions (47). Other biomarkers that may be valuable in obtaining prognostic information include two apoptosis genes, MCL1 and telomerase reverse transcriptase (TERT).

Over expression of telomerase reverse transcriptase (TERT) leads to cessation of telomere shortening, hence being associated with oncogenesis. Another apoptotic gene MCL1, encodes for an anti-apoptotic protein, which is a member of the Bcl-2 family. Alternate splicing of this

protein leads to isoform1, which inhibits apoptosis directly. Both molecules appear to be useful in early detection of various malignancies, including HCC (48).

N-acetyltransferase 2 (NAT2) is also identified in our analysis. NAT2 has been implicated in the activation/inactivation of medications, as well as carcinogens. Subsequently, NAT2 may serve as a biomarker to predict the risk for drug induced liver injury (49). In prior studies of CC, genetic polymorphisms and upregulation of NAT2 have correlated with risk for CC (50). Such findings can be correlated with those listed above for KRAS and MET.

The culmination of the above genetic analysis demonstrates the utility of GeneCards in the analysis of low frequency malignancies such as HCC-CC. Not only did this genetic analysis confirm previously documented serologic markers for HCC-CC, it identifies several unique molecular targets, which may be useful in studies evaluating the pathogenesis, diagnosis, and prognosis of HCC-CC. It has also illuminated the similarity of HCC-CC with the pathologic subtypes of CC as compared to HCC. This genetic compendium also permits the creation of a map outlining relationships between several of these genes which may allow a better understanding of the pathogenesis of this rare primary neoplasm.

Despite these findings, potential weaknesses of this study include the retrospective evaluation of data collected from small numbers of patients. However utilizing a vast database, such as GeneCards®, The Human Gene Compendium, allows for an expanded evaluation of a rare disease. This approach has been validated in the past in similar neoplasms as well as more common liver disease. The ability to analyze a complex series of genetic components, which would be otherwise time and labor intensive, is an added benefit of this approach.

Although several novel cellular components and pathways have been identified as potential biomarkers for HCC-CC, the utility of each of these components or the combination of these biomarkers in clinical diagnosis and prognosis have yet to be determined. Nonetheless, employing large relational genetic databases such as GeneCards for an initial analysis will permit more focused investigation into the utility of biomarkers as well as guide studies of pathogenesis and future therapies.

Acknowledgements

None.

Footnote

Conflicts of Interest: AB Elfant is a consultant for Boston Scientific. The other authors have no conflicts of interest to declare.

References

1. Wittekind C, Fischer HP, Ponchon T. Combined hepatocellular and cholangiocarcinoma. In: Hamilton SR, Aaltonen LA, editors. World Health Organization classification of tumours: pathology and genetics of tumours of the digestive system. Lyon: IARC Press, 2000:181-2.
2. Park H, Choi KH, Choi SB, et al. Clinicopathological characteristics in combined hepatocellular-cholangiocarcinoma: a single center study in Korea. Yonsei Med J 2011;52:753-60.
3. Sasaki M, Sato H, Kakuda Y, et al. Clinicopathological significance of 'subtypes with stem-cell feature' in combined hepatocellular-cholangiocarcinoma. Liver Int 2015;35:1024-35.
4. Tang D, Nagano H, Nakamura M, et al. Clinical and pathological features of Allen's type C classification of resected combined hepatocellular and cholangiocarcinoma: a comparative study with hepatocellular carcinoma and cholangiocellular carcinoma. J Gastrointest Surg 2006;10:987-98.
5. Yano Y, Yamamoto J, Kosuge T, et al. Combined hepatocellular and cholangiocarcinoma: a clinicopathologic study of 26 resected cases. Jpn J Clin Oncol 2003;33:283-7.
6. Ng IO, Shek TW, Nicholls J, et al. Combined hepatocellular-cholangiocarcinoma: a clinicopathological study. J Gastroenterol Hepatol 1998;13:34-40.
7. Hauben E, Struyf N, Michielsen P, et al. Cytokeratin profiles and mucin secretion in combined hepatocellular-cholangiocarcinoma. A case report. Pathol Res Pract 1996;192:488-91.
8. Tickoo SK, Zee SY, Obiekwe S, et al. Combined hepatocellular-cholangiocarcinoma: a histopathologic, immunohistochemical, and in situ hybridization study. Am J Surg Pathol 2002;26:989-97.
9. Coulouarn C, Cavard C, Rubbia-Brandt L, et al. Combined hepatocellular-cholangiocarcinomas exhibit progenitor features and activation of Wnt and TGFβ signaling pathways. Carcinogenesis 2012;33:1791-6.
10. Itoyama M, Hata M, Yamanegi K, et al. Expression of both hepatocellular carcinoma and cholangiocarcinoma

phenotypes in hepatocellular carcinoma and cholangiocarcinoma components in combined hepatocellular and cholangiocarcinoma. Med Mol Morphol 2012;45:7-13.
11. Cazals-Hatem D, Rebouissou S, Bioulac-Sage P, et al. Clinical and molecular analysis of combined hepatocellular-cholangiocarcinomas. J Hepatol 2004;41:292-8.
12. Chu KJ, Lu CD, Dong H, et al. Hepatitis B virus-related combined hepatocellular-cholangiocarcinoma: clinicopathological and prognostic analysis of 390 cases. Eur J Gastroenterol Hepatol 2014;26:192-9.
13. Zhan Q, Shen BY, Deng XX, et al. Clinical and pathological analysis of 27 patients with combined hepatocellular-cholangiocarcinoma in an Asian center. J Hepatobiliary Pancreat Sci 2012;19:361-9.
14. Mourad W, Khalaf H, Tulbah A, et al. Clinicopathologic features of hepatic neoplasms in explanted livers: a single institution experience. Ann Saudi Med 2007;27:437-41.
15. Okuda H, Shiratori K, Yamamoto M, et al. Clinicopathologic features of patients with intrahepatic cholangiocarcinoma who are seropositive for alpha-fetoprotein-L3 and those with combined hepatocellular and cholangiocarcinoma. J Gastroenterol Hepatol 2006;21:869-73.
16. Fujii H, Zhu XG, Matsumoto T, et al. Genetic classification of combined hepatocellular-cholangiocarcinoma. Hum Pathol 2000;31:1011-7.
17. Koh KC, Lee H, Choi MS, et al. Clinicopathologic features and prognosis of combined hepatocellular cholangiocarcinoma. Am J Surg 2005;189:120-5.
18. Jarnagin WR, Weber S, Tickoo SK, et al. Combined hepatocellular and cholangiocarcinoma: demographic, clinical, and prognostic factors. Cancer 2002;94:2040-6.
19. Maeda T, Adachi E, Kajiyama K, et al. Combined hepatocellular and cholangiocarcinoma: proposed criteria according to cytokeratin expression and analysis of clinicopathologic features. Hum Pathol 1995;26:956-64.
20. Tomimatsu M, Ishiguro N, Taniai M, et al. Hepatitis C virus antibody in patients with primary liver cancer (hepatocellular carcinoma, cholangiocarcinoma, and combined hepatocellular-cholangiocarcinoma) in Japan. Cancer 1993;72:683-8.
21. European Association For The Study Of The Liver; European Organisation For Research And Treatment Of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2012;56:908-43.
22. Khan SA, Davidson BR, Goldin RD, et al. Guidelines for the diagnosis and treatment of cholangiocarcinoma: an update. Gut 2012;61:1657-69.
23. Fowler KJ, Sheybani A, Parker RA 3rd, et al. Combined hepatocellular and cholangiocarcinoma (biphenotypic) tumors: imaging features and diagnostic accuracy of contrast-enhanced CT and MRI. AJR Am J Roentgenol 2013;201:332-9.
24. Wells M, Venkatesh S, Johnson G, et al. Biphenotypic Intrahepatic Cholangiocarcinoma and Hepatocellular Carcinoma: Imaging Features on MRI, CT, Ultrasound and PET. Radiological Society of North America 2014 Scientific Assembly and Annual Meeting, Chicago IL. http://archive.rsna.org/2014/14007120.html. Accessed January 4, 2015
25. O'Connor K, Walsh JC, Schaeffer DF. Combined hepatocellular-cholangiocarcinoma (cHCC-CC): a distinct entity. Ann Hepatol 2014;13:317-22.
26. Wanless IR. International consensus on histologic diagnosis of early hepatocellular neoplasia. Hepatol Res 2007;37 Suppl 2:S139-41.
27. Rappaport N, Twik M, Nativ N, et al. MalaCards: a comprehensive automatically-mined database of human diseases. Curr Protoc Bioinformatics 2014;47:1.24.1-19.
28. Stelzer G, Dalah I, Stein TI, et al. In-silico human genomics with GeneCards. Hum Genomics 2011;5:709-17.
29. Safran M, Dalah I, Alexander J, et al. GeneCards Version 3: the human gene integrator. Database (Oxford) 2010;2010:baq020.
30. Kim JM, Kim HG, Han JM, et al. The herbal formula CGX ameliorates the expression of vascular endothelial growth factor in alcoholic liver fibrosis. J Ethnopharmacol 2013;150:892-900.
31. Zhou T, Wang T, Garcia JG. Genes influenced by the non-muscle isoform of Myosin light chain kinase impact human cancer prognosis. PLoS One 2014;9:e94325.
32. Sharma G, Agarwal SM. Identification of critical microRNA gene targets in cervical cancer using network properties. Microrna 2014;3:37-44.
33. Baasiri RA, Glasser SR, Steffen DL, et al. The breast cancer gene database: a collaborative information resource. Oncogene 1999;18:7958-65.
34. Keller JW, Doyle MM, Wang-Gilliam A, et al. Analysis of the genomic profile of biphenotypic tumors compared to cholangiocarcinoma and hepatocellular carcinoma. J Clin Oncol 2014;32:abstr 226.
35. Santarpia M, Altavilla G, Margeli M, et al. PIK3CA mutations and BRCA1 expression in breast cancer: potential biomarkers for chemoresistance. Cancer Invest

2008;26:1044-51.
36. Sierra JR, Tsao MS. c-MET as a potential therapeutic target and biomarker in cancer. Ther Adv Med Oncol 2011;3:S21-35.
37. Varnholt H, Asayama Y, Aishima S, et al. C-met and hepatocyte growth factor expression in combined hepatocellular and cholangiocarcinoma. Oncol Rep 2002;9:35-41.
38. Imai Y, Oda H, Arai M, et al. Mutational analysis of the p53 and K-ras genes and allelotype study of the Rb-1 gene for investigating the pathogenesis of combined hapatocellular-cholangiocellular carcinomas. Jpn J Cancer Res 1996;87:1056-62.
39. Cavallini A, Valentini AM, Lippolis C, et al. KRAS genotyping as biomarker in colorectal cancer: a comparison of three commercial kits on histologic material. Anticancer Res 2010;30:5251-6.
40. Ohlsson L, Israelsson A, Öberg Å, et al. Lymph node CEA and MUC2 mRNA as useful predictors of outcome in colorectal cancer. Int J Cancer 2012;130:1833-43.
41. Nathan H, Pawlik TM, Wolfgang CL, et al. Trends in survival after surgery for cholangiocarcinoma: a 30-year population-based SEER database analysis. J Gastrointest Surg 2007;11:1488-96; discussion 1496-7.
42. Weimann A, Varnholt H, Schlitt HJ, et al. Retrospective analysis of prognostic factors after liver resection and transplantation for cholangiocellular carcinoma. Br J Surg 2000;87:1182-7.
43. Wang J, Wang F, Kessinger A. Outcome of combined hepatocellular and cholangiocarcinoma of the liver. J Oncol 2010;2010.
44. Groeschl RT, Turaga KK, Gamblin TC. Transplantation versus resection for patients with combined hepatocellular carcinoma-cholangiocarcinoma. J Surg Oncol 2013;107:608-12.
45. Nagorney DM, Kendrick ML. Hepatic resection in the treatment of hilar cholangiocarcinoma. Adv Surg 2006;40:159-71.
46. Yang SL, Fang X, Huang ZZ, et al. Can serum glypican-3 be a biomarker for effective diagnosis of hepatocellular carcinoma? A meta-analysis of the literature. Dis Markers 2014;2014:127831.
47. Wang D, Guo XZ, Li HY, et al. Prognostic significance of cyclooxygenase-2 protein in pancreatic cancer: a meta-analysis. Tumour Biol 2014;35:10301-7.
48. El-Mazny A, Sayed M, Sharaf S. Human telomerase reverse transcriptase messenger RNA (TERT mRNA) as a tumour marker for early detection of hepatocellular carcinoma. Arab J Gastroenterol 2014;15:68-71.
49. Higuchi N, Tahara N, Yanagihara K, et al. NAT2 6A, a haplotype of the N-acetyltransferase 2 gene, is an important biomarker for risk of anti-tuberculosis drug-induced hepatotoxicity in Japanese patients with tuberculosis. World J Gastroenterol 2007;13:6003-8.
50. Prawan A, Kukongviriyapan V, Tassaneeyakul W, et al. Association between genetic polymorphisms of CYP1A2, arylamine N-acetyltransferase 1 and 2 and susceptibility to cholangiocarcinoma. Eur J Cancer Prev 2005;14:245-50.

Cite this article as: Mok SR, Mohan S, Grewal N, Elfant AB, Judge TA. A genetic database can be utilized to identify potential biomarkers for biphenotypic hepatocellular carcinoma-cholangiocarcinoma. J Gastrointest Oncol 2016;7(4):570-579. doi: 10.21037/jgo.2016.04.01

Table S1 Search strategy for genes and biomarkers identified in HCC, CC and HCC-CC

Gene symbol	Pathologic subtype		
	HCC	HCC-CC	Cholangiocarcinoma
1	MET	MET	MUC1
2	TP53	KRT7	AC019117.1
3	CDH1	KRT19	AFP
4	CTNNB1	VEGFC	ASPH
5	KRT7	CDKN1B	CCNB1
6	PCNA	MUC2	CDKN1B
7	PIK3CA	CEACAM3	CDX2
8	NME1	MUC5AC	CEACAM3
9	NKX2-1	AFP	CEACAM6
10	KRT19	HGF	CFLAR
11	CDKN2A	CDX2	ERBB2
12	RASSF1	PTGS2	HGF
13	CCND1	CCNB1	KRT20
14	CDKN1A	TIMP3	KRT19
15	FHIT	OGG1	KRT7
16	CASP8	CFLAR	LGALS3
17	KRT18	MAGEA3	MAGEA3
18	KRT8	CEACAM6	MCL1
19	PTEN	MCL1	MET
20	VEGFC	ASPH	MUC2
21	CEACAM5	PTPN3	MUC17
22	MKI67	TP53	MUC5AC
23	AXIN1	CDH1	MUC4
24	IFI27	PIK3CA	OGG1
25	BIRC5	RASSF1	PTGS2
26	CDKN1B	FHIT	THBS1
27	TERT	KRT18	TIMP3
28	MUC2	KRT8	VEGFC
29	TYMP	CEACAM5	PTPN3
30	CEACAM3	KRAS	GPC3
31	TGFBR2	TERT	KRAS
32	MMP2	SMAD4	ALPL
33	VEGFA	TNFSF10	ALPP
34	BCL2	MMP7	ALPPL2
35	SMAD4	PSG2	APOC1P1
36	BCL2L1	GPC3	ASB16-AS1
37	TGFA	CTNNA1	
38	JUP	PTK2	BRAF
39	IGF2R	ALPP	C7orf55
40	CYP1A1	MME	CASP9
41	MMP9	CHKA	CCKBR
42	RARB	RAF1	CDH1
43	KRT14	EBAG9	CEACAM5
44	TNFSF10	NAT2	CGB
45	APC	GGT1	CHML
46	BAX	CCKBR	CHKA
47	EGF	RECK	COX16
48	HRAS	ALPPL2	CTNNA1
49	MMP14	ALPL	CXorf40A
50	MMP7	SCT	DAPK1
51	MUC5AC		DMRTC1
52	PSG2	HCC	DYT10
53	SERPINB5		EBAG9
54	ABCB1		EIF4E
55	AFP		FHIT
56	CTNNA1		GGT1
57	GSTP1		KRT18
58	CDKN3		KRT8
59	HGF		KRTCAP3
60	TGFB1		LINC00543
61	BRCA2		LRRC26
62	CDX2		MAGEA10
63	KRAS		MAPK14
64	TIMP2		MIR200C
65	AKT1		MIR214
66	ANGPT2		MME
67	AMACR		MMP7
68	CCNA2		MUC5B
69	GRP		NAT1
70	IGF1R		NAT2
71	PLAU		NCAM1
72	RB1		NRAS
73	SLC2A1		OVCA2
74	CDK2		PET100
75	CDK4		PIK3CA
76	EPCAM		PLA2G4A
77	HIF1A		PNLIP
78	MDM2		PPFIBP2
79	MYC		PSG2
80	PLAUR		PTK2
81	PTCH1		RAF1
82	SETD2		RASSF1
83	PTGS2		RECK
84	MTUS1		SCT
85	CAV1		SLPI
86	CASP3		SMAD4
87	CEACAM7		SPANXA2
88	DNMT1		SSX2B
89	ERBB3		TBC1D3H
90	FASLG		TERT
91	FAS		TFF1
92	KDR		TIGD2
93	PTK2		TMEM139
94	TP73		TMEM256
95	HNF1A		TNFSF10
96	ALPP		TNFRSF1B
97	ABCC1		TNFRSF1A
98	CCNB1		TP53
99	CDH2		
100	CDH3		
101	DPYD		
102	TERC		
103	CTAG1B		
104	CYP2E1		
105	MTOR		
106	BSG		
107	CEACAM1		
108	CTNND1		
109	HBEGF		
110	HSPB1		
111	JUN		
112	LGALS1		
113	MAPK1		
114	MME		
115	SKP2		
116	FN1		
117	TWIST1		
118	CHKA		
119	CRAT		
120	ENG		
121	GSTT1		
122	SERPINB2		
123	TIMP3		
124	AHR		
125	CD34		
126	IGF2		
127	MAGEA4		
128	PTK2B		
129	SERPINB3		
130	SNAI2		
131	CYCS		
132	E2F1		
133	FGF2		
134	GSTM1		
135	ITGB1		
136	MVP		
137	NOTCH1		
138	OGG1		
139	PARP1		
140	PDGFRA		
141	RAF1		
142	STK11		
143	GPC3		
144	BAK1		
145	CFLAR		
146	FGFR2		
147	FOS		
148	HPSE		
149	PRKCA		
150	RARA		
151	RHOA		
152	STAT3		
153	BAGE		
154	IFNA1		
155	MAGEA1		
156	ABCG2		
157	AREG		
158	CD80		
159	CD82		
160	CDK1		
161	CDKN2B		
162	DPP4		
163	EPOR		
164	ERCC1		
165	ETS1		
166	FGF1		
167	FGF3		
168	HDAC9		
169	HLA-A		
170	KRT1		
171	MGMT		
172	PLG		
173	SPINK1		
174	TNFRSF10B		
175	TXN		
176	RHOC		
177	SPP1		
178	ANGPT1		
179	AR		
180	CD46		
181	CSF1R		
182	EBAG9		
183	FGFR1		
184	H19		
185	HSPA1A		
186	ITGAV		
187	KRT13		
188	MAPK8		
189	MAPK3		
190	NAT2		
191	SMAD2		
192	SP1		
193	TAP1		
194	TFAP2A		
195	TGFBR1		
196	TFF3		
197	TIMP1		
198	CDKN1C		
199	GGT1		
200	MAGEA3		
201	MMP12		
202	WNT1		
203	ACP1		
204	ANPEP		
205	ANXA2		
206	CD86		
207	CEACAM6		
208	CXADR		
209	DNASE1		
210	EGR1		
211	EZR		
212	FASN		
213	GJA1		
214	HSPA4		
215	ILK		
216	ING1		
217	IRF1		
218	ITGA6		
219	NEU1		
220	NRP1		
221	PSMB9		
222	PXN		
223	SOD2		
224	TEK		
225	VTN		
226	XIAP		
227	HGFAC		
228	KLF6		
229	MT1G		
230	SOCS1		
231	TNFRSF6B		
232	TPX2		
233	G6PC		
234	ANXA5		
235	APAF1		
236	AURKA		
237	CASP1		
238	CCKBR		
239	CLU		
240	CREB1		
241	CTTN		
242	CYR61		
243	DCN		
244	GAPDH		
245	GNRHR		
246	GZMB		
247	HDAC1		
248	HNF4A		
249	IFNA2		
250	IL6R		
251	MAP2K1		
252	MEN1		
253	NQO1		
254	PTTG1		
255	RECK		
256	SSTR2		
257	TAP2		
258	TGFB2		
259	TGIF1		
260	TNFRSF10A		
261	ABCC2		
262	ADAMTSL1		
263	ALDH2		
264	ANXA1		
265	APOD		
266	BMP6		
267	BMP7		
268	CALR		
269	CCR7		
270	CDC25A		
271	CXCL12		
272	CYP27B1		
273	CYP1B1		
274	DCK		
275	EDNRA		
276	EPHB4		
277	HNF1B		
278	LINC01194		
279	MCL1		
280	MTA1		
281	NFKBIA		
282	NR1H2		
283	PDGFB		
284	POMGNT2		
285	PPP2R4		
286	RUNX3		
287	SMARCA4		
288	ARG1		
289	FOXM1		
290	HIC1		
291	ADCY10		
292	ALPPL2		
293	ATP7B		
294	BAD		
295	CADM1		
296	CD4		
297	CDH13		
298	COPS5		
299	CXCR2		
300	CYP17A1		
301	CYP3A4		
302	DIABLO		
303	DNMT3B		
304	ENO1		
305	EZH2		
306	GDF15		
307	HIST4H4		
308	HSPA5		
309	IRS1		
310	LEPR		
311	PIGR		
312	PIK3CG		
313	SHC1		
314	SMAD7		
315	STMN1		
316	TSG101		
317	VIPR1		
318	ALB		
319	ALDOB		
320	ALDH3A1		
321	CLDN10		
322	DIO3		
323	FABP1		
324	HDGF		
325	HFE		
326	IFNAR2		
327	LAPTM4B		
328	MDK		
329	REG3A		
330	RSF1		
331	SERPINA7		
332	AMFR		
333	AURKB		
334	BCL10		
335	CCNG1		
336	CDC6		
337	CSK		
338	CYP2A6		
339	DDIT3		
340	DIO2		
341	EDNRB		
342	EPAS1		
343	F2R		
344	FGFR4		
345	FST		
346	GJB1		
347	GJB2		
348	HMGA1		
349	IDO1		
350	IFNB1		
351	INSR		
352	KISS1		
353	REG1A		
354	RHOD		
355	S100A8		
356	SPINT1		
357	TCF7L2		
358	VCP		
359	VDR		
360	ZEB2		
361	CCL20		
362	CD81		
363	DLC1		
364	F2		
365	FGL1		
366	FZD7		
367	GLUL		
368	GNMT		
369	GPT		
370	LECT2		
371	MAGEC2		
372	MAT1A		
373	MAT2A		
374	PEG10		
375	PINX1		
376	PSMD10		
377	TAT		
378	UGT1A7		
379	ABCC3		
380	CCR6		
381	CHUK		
382	CIITA		
383	CTLA4		
384	CXCL9		
385	DNAH8		
386	DNMT3A		
387	DUSP1		
388	EEF1A1		
389	EIF2AK2		
390	EPHX1		
391	ETS2		
392	F13A1		
393	FADD		
394	GHRL		
395	HBB		
396	HPN		
397	HSP90B1		
398	IL15		
399	LDLR		
400	MBD4		
401	MGAT5		
402	MTAP		
403	MUTYH		
404	PIK3R1		
405	PIN1		
406	PKM		
407	PNP		
408	POLR3K		
409	PPARD		
410	PPIG		
411	PRDX5		
412	PRKCB		
413	PSEN2		
414	PSMB8		
415	RXRA		
416	SERPINA1		
417	SFRP1		
418	SMAD3		
419	SOAT1		
420	ST6GAL1		
421	TCF4		
422	TEP1		
423	TFDP1		
424	THRB		
425	TK1		
426	TNFSF11		
427	TXNRD2		
428	ADH1B		
429	AIFM1		
430	AKR1A1		
431	ALPL		
432	BID		
433	BUB1		
434	CDH17		
435	CISH		
436	CYP1A2		
437	EIF2AK3		
438	F2RL2		
439	F9		
440	FUCA1		
441	GAL3ST1		
442	GLS2		
443	HK2		
444	IL32		
445	IL1RAPL2		
446	ITGA1		
447	JAG1		
448	JAK1		
449	KLRK1		
450	MAD2L1		
451	MMP15		
452	NOTCH4		
453	NPM1		
454	NR1I2		
455	NR3C1		
456	NTRK2		
457	OXA1L		
458	PARK2		
459	PODXL		
460	PPP2R1B		
461	PRKG1		
461	PRLR		
462	PTPN13		
463	PTPN3		
464	RPS6KB1		
465	RUNX1		
466	SCT		
467	SERPINF1		
468	STAT5B		
469	STC1		
470	SULT2A1		
471	TFPI2		
472	UGT1A		
473	VIP		
474	XPC		
475	PDGFRL		
476	ABCA1		
477	ACP5		
478	ALDH9A1		
479	ARMC10		
480	ASGR2		
481	ASS1		
482	ATF1		
483	BECN1		
484	CBS		
485	CCNE1		
486	CCR1		
487	CD58		
488	CD63		
489	CDKN2C		
490	CSE1L		
491	CSN1S1		
492	CTGF		
493	CYP20A1		
494	EFNB2		
495	EHHADH		
496	EIF6		
497	ELK1		
498	ENPP2		
499	EPHX2		
500	EPRS		
501	F8		
502	F10		
503	FBN1		
504	GCLC		
505	GLS		
506	GRB2		
507	GSTA1		
508	GSTA2		
509	GSTA4		
510	GSTA3		
511	IKBKB		
512	IL6		
513	ITGAL		
514	ITGA5		
515	JUNB		
516	KISS1R		
517	KRIT1		
518	MGAT3		
519	MLXIPL		
520	MMP16		
521	PDLIM5		
522	PLA2G6		
523	PML		
524	POLR2L		
525	PPARA		
526	PRDM2		
527	PROM1		
528	PTMA		
529	RBP4		
530	SCARB1		
531	SHBG		
532	SLC11A2		
533	SLC2A2		
534	SMAD6		
535	SOCS3		
536	SOX4		
537	SRD5A2		
538	SRF		
539	SULF1		
540	TLR3		
541	TNFRSF9		
542	TSPO		
543	UGT1A1		
544	UGT1A9		
545	XBP1		
546	YBX3		
547	ABCB7		
548	ABCB4		
549	AC004862.6		
550	ACSL4		
551	ADORA2A		
552	AICDA		
553	AHSG		
554	AKR1B10		
555	ANAPC11		
556	ANXA10		
557	AOC3		
558	APOBEC1		
559	ARHGAP1		
560	ASPH		
561	ATF6		
562	BCR		
563	BDNF		
564	BHMT		
565	BNIPL		
566	C19orf80		
567	C9orf78		
568	CAPN2		
569	CCL15		
570	CCL19		
571	CCL21		
572	CCL5		
573	CCNC		
574	CCT2		
575	CD14		
576	CHP2		
577	CLEC12A		
578	COPA		
579	CREB3L3		
580	CX3CR1		
581	CYP2C19		
582	CYP3A7		
583	DACT1		
584	DEK		
585	DEPDC5		
586	DKK1		
587	DLAT		
588	DLGAP5		
589	DNAJA3		
590	DNAJB1		
591	ECI2		
592	EHD4		
593	EIF2D		
594	ELANE		
595	ETHE1		
596	F2RL3		
597	F7		
598	FAH		
599	FBP1		
600	FOXP3		
601	GOT1		
602	GOT2		
603	HCCAT4		
604	HCCAT3		
605	HCCAT5		
606	HBE1		
607	HCRP1		
608	HEIH		
609	HEPN1		
610	HHCM		
611	HK1		
612	HLA-DRB1		
613	HLF		
614	HULC		
615	IGF2BP2		
616	INS		
617	JUND		
618	KIR3DL1		
619	KLHDC2		
620	KLKB1		
621	KTN1		
622	LCO		
623	LINC00261		
624	LIPC		
625	MAGEE1		
626	MAGEC3		
627	MAGED2		
628	MAGEE2		
629	MARVELD2		
630	METTL21B		
631	METTL21A		
632	MTO1		
633	NDNL2		
634	NFE2L2		
635	NR1I3		
636	NXF1		
637	OTC		
638	PAGE5		
639	PEMT		
640	PHF20		
641	PNOC		
642	POLR1C		
643	PPAT		
644	PPBP		
645	PRDX4		
646	PRKCI		
647	PRKACA		
648	PSD3		
649	PSMG2		
650	PYCARD		
651	RBM39		
652	RGN		
653	RPLP0		
654	SARNP		
655	SEPP1		
656	SHC3		
657	SIRPA		
658	SLC17A5		
659	SLC25A13		
660	SLC25A47		
661	SLC28A1		
662	SLCO1B3		
663	SLCO1B1		
664	SMYD3		
665	SOAT2		
666	SQSTM1		
667	SRRD		
668	STARD13		
669	TCN1		
670	TF		
671	TFDP3		
672	TH		
673	TGS1		
674	TLR2		
675	TLR4		
676	TMEM176A		
677	UAP1		
678	UGT2B7		
679	URGCP		
680	URI1		
681	UROD		
682	UTP6		
683	VPS37A		
684	VPS54		
685	YY1AP1		
686	ZC4H2		
687	ZDHHC2		

Table S2 Search strategy for genes and biomarkers identified in the pathologic subtypes of HCC, CC and HCC-CC[a]

Gene symbol	Pathologic subtype									
	Fibrolamellar HCC	Scirrhous HCC	Sarcomatoid HCC	Lympho-epithelial-like HCC	Intraductal papillary CC	Intestinal-type CC	Signet ring CC	Clear cell CC	Squamous cell CC	Small cell CC
1	DNAJB1	MET	CDH1	CDH1	ERBB2	KRAS	KRT20	MUC1	TP53	TP53
2	MARVELD2	TGFB1	PCNA	MET	SMAD4	ERBB2	KRT7	KRT7	MUC1	MUC1
3	MTO1		NKX2-1	MUC2		KRT7	MUC2	OGG1	KRT7	KRT7
4	PAGE5		CDKN2A	ABCB1		BRAF	CDH1	AFP	KRT19	MET
5	PRKACA		CEACAM3	CASP8		PIK3CA	CEACAM5	MET	CDH1	ERBB2
6			VEGFA	VEGFC		TFF1		MME	CDKN1B	NCAM1
7			JUP	JUP		TNFRSF1A		CDH1	MET	KRAS
8			FAS	APC		TP53		FHIT	KRT8	OGG1
9			CTNND1	STAT3				ALPP	FHIT	KRT19
10			PDGFRA	MUC5AC				BRAF	VEGFC	BRAF
11			IFNA2	CTNNA1				RASSF1	ERBB2	CDKN1B
12				TGFB1				SMAD4	CCNB1	KRT20
13				AMACR					KRT20	FHIT
14				HIF1A					KRAS	VEGFC
15				MDM2					MCL1	KRT18
16				PLAUR					OGG1	KRT8
17				PTGS2					LGALS3	CEACAM5
18				ALPP					DAPK1	CEACAM3
19				CDH2					CEACAM5	RASSF1
20				LGALS1					EIF4E	LGALS3
21				TEK					RASSF1	MME
22				IL6					CEACAM3	PTGS2
23				IL6R					TERT	HGF
24				AREG					HGF	MUC4
25				CD80					BRAF	TERT
26				DPP4					CGB	PIK3CA
27				FGF1					PTK2	PTK2
28				HDAC9					TIMP3	TNFSF10
29				TXN					KRT18	MAGEA3
30				RHOC					CFLAR	DYT10
31				CSF1R					THBS1	RAF1
32				NAT2					RECK	CEACAM6
33				TFF3					NAT2	CHKA
34				TIMP1					ALPP	MMP7
35				CD86					RAF1	PSG2
36				DNASE1					MUC4	CCKBR
37				EGR1					PTGS2	SLPI
38				HSPA4					MMP7	PLA2G4A
39				IRF1					PSG2	SCT
40				NEU1					CASP9	
41				IFNB1					PIK3CA	
42				SOCS1					TNFSF10	
43				CASP1					CDX2	
44				IFNA2					MAGEA3	
45				TGIF1					MUC2	
46				CCL20					MUC5AC	
47				CTLA4					SMAD4	
48				SERPINA1					CHKA	
49				BMP7					EBAG9	
50				CXCL12					CCKBR	
51				CYP27B1					CTNNA1	
52				NFKBIA					TNFRSF1B	
53				NR1H2					MIR214	
54				IL1RAPL2					TFF1	
55				VIP						
56				ALPPL2						
57				CD4						
58				ENO1						
59				PIGR						
60				SMAD7						
61				ALB						
62				REG3A						
63				TLR4						
64				F2R						
65				IDO1						
66				S100A8						
67				VDR						
68				PLA2G6						
69				CHUK						
70				CIITA						
71				F13A1						
72				IL15						
73				SOAT1						
74				TNFSF11						
75				ALPL						
76				NR1I2						
77				NR3C1						
78				SCT						
79				CCL5						
80				CD14						
81				ELANE						
82				FOXP3						
83				HLA-DRB1						
84				TLR2						
85				ITGAL						
86				PPARA						
87				SOCS3						
88				XBP1						
89				AOC3						
90				CCL19						
91				CCL21						
92				DLAT						

[a], HCC-CC subtypes stem-cell and classic yielded no results. HCC-CC, hepatocellular carcinoma-cholangiocarcinoma; HCC, hepatocellular carcinoma; CC, cholangiocarcinoma.

Table S3 Genes and potential biomarkers present for each risk factor for HCC-CC

Gene function	Gene name	General		Infectious					Metabolic			Congenital				Autoimmune			Carcinogen			Vascular
		Cirrhosis	PHTN	Hepatitis C	Hepatitis B	Cholangitis	Opisthorchiasis	Recurrent cholangitis	NAFLD	NASH	Tyrosinemia	Cystic disease of liver	Biliary cyst	Choledochal cyst	Other congenital malformations of gallbladder	AIH	PBC	PSC	Alcoholic fatty liver	Alcohol	Aflatoxin	PVT
Embryonic	AFP	F																				
	CEACAM3																					
	GPC3																					
Hormonal	CCKBR																					
	SCT																					
Cell Adhesion	HGF																					
	MET																					
Cytokeratin	KRT7																					
	KRT8																					
	KRT18																					
	KRT19																					
Mucin Production	MUC2																					
	MUC5AC																					
Metalloproteins	MME																					
	MMP7																					
	RECK																					
	TIMP3																					
Ras Signaling	KRAS																					
	RAF1																					
	RASSF1																					
Metabolism	ALPL																					
	ALPP																					
	ALPPL2																					
	CHKA																					
	GGT1																					
	NAT2																					
	PTGS2																					
Apoptosis	MCL1																					
	PIK3CA																					
	TERT																					

HCC-CC, hepatocellular carcinoma-cholangiocarcinoma; HCC, hepatocellular carcinoma; CC, cholangiocarcinoma; ECM, extracellular membrane; ER, endoplasmic reticulum; MMP, matrix metalloproteinase, GTPase, guanyl triphosphatase; COX, cyclo-oxygenase; BCL-2, B-cell lymphoma 2; ATP, adenosine tri-phosphate; PHTN, portal hypertension; NAFLD, non-alcoholic fatty liver disease; NASH, non-alcoholic steatohepatitis; AIH, autoimmune hepatitis; PBC, primary biliary cirrhosis; PSC, primary sclerosing cholangitis; PVT, portal vein thrombosis.

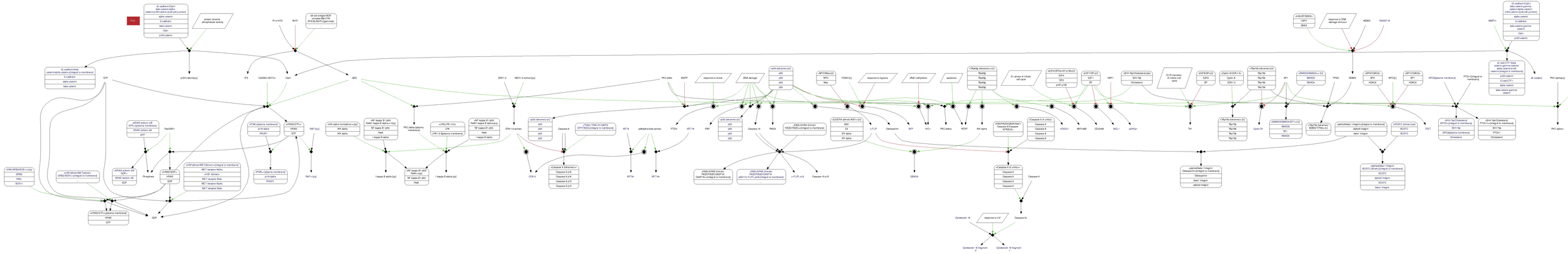

Figure S1 Gene interaction map. Pathway interaction database from the national cancer institute, demonstrating the interaction between all genes identified for biphenotypic hepatocellular carcinoma-cholangiocarcinoma. Purple, no dominant characteristics of hepatocellular carcinoma or cholangiocarcinoma; Red, dominant characteristics of hepatocellular carcinoma; Blue, dominant characteristics of cholangiocarcinoma.