Primary liver cancer is the second leading cause of cancer-related death worldwide and therefore a major public health challenge. We review hypotheses of the cell of origin of liver tumorigenesis, and clarify the classes of liver cancer, based on molecular features and how these affect patient prognosis. Primary liver cancer comprises hepatocellular carcinoma (HCC), intrahepatic cholangiocarcinoma (iCCA), and other rare tumors, notably fibrolamellar carcinoma and hepatoblastoma. The molecular and clinical features of HCC vs iCCA are distinct, although they have overlapping risk factors and pathways of oncogenesis. A better understanding of the cell types originating liver cancer can aid in exploring molecular mechanisms of carcinogenesis and therapeutic options. Molecular studies have identified adult hepatocytes as the cell of origin. These cells have been proposed to transform directly into HCC cells (via a sequence of genetic alterations), to de-differentiate into hepatocyte precursor cells (which then become HCC cells that express progenitor cell markers), or to trans-differentiate into biliary-like cells (which give rise to iCCA). Alternatively, progenitor cells also give rise to HCCs and iCCAs with markers of progenitor cells. Advances in genome profiling and next-generation sequencing have led to the classification of HCCs based on molecular features, and assigned them to categories such as proliferation-progenitor, proliferation-transforming growth factor beta, and Wnt-catenin beta 1. iCCAs have been assigned to categories of proliferation and inflammation. Overall, proliferation subclasses are associated with a more aggressive phenotype and poor outcome of patients, although more specific signatures have refined our prognostic abilities. Analyses of genetic alterations have identified those that might be targeted therapeutically, such as fusions in the FGFR2 gene and mutations in genes encoding isocitrate dehydrogenases (in approximately 60% of iCCAs) or amplifications at 11q13 and 6p21 (in approximately 30% of HCCs). Further studies of these alterations are needed before these can be used as biomarkers in clinical decision.
Title: Liver Cancer Cell of Origin, Molecular Class, and Effects on Patient Prognosis

Daniela Sia¹, Augusto Villanueva¹, Scott L. Friedman¹ and Josep M. Llovet¹,²,³

¹ Mount Sinai Liver Cancer Program (Divisions of Liver Diseases, Hematology and Medical Oncology, Department of Medicine), Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, USA.
² Liver Cancer Translational Research Laboratory, BCLC, Liver Unit, CIBEREHD, IDIBAPS, Hospital Clinic, University of Barcelona, Catalonia, Spain.
³ Institució Catalana de Recerca i Estudis Avançats, Barcelona, Catalonia, Spain.

Grant support: DS is the recipient of the Andrea Marie Fuquay Memorial Research Fellowship (Cholangiocarcinoma Foundation). AV is the recipient of the American Association for the Study of Liver Diseases Foundation Alan Hofmann Clinical and Translational Award. SLF is supported by NIH grants DKRO156621 and AA. 020709. J.M.L. has grants from the US National Cancer Institute (NCI) (P30CA165979), the European Commission Horizon 2020 (HEP-CAR 667273–2), the Samuel Waxman Cancer Research Foundation, the Grant I+D Program (SAF2013-41027) and the Asociación Española Contra el Cáncer (AECC).

Disclosures: AV and JML are co-inventors of a patent: “Compositions, Kits, and methods for identification, assessment, prevention and therapy of hepatic disorders”.

Total Word counts: current 6543

References: current 176

Tables: 2

Figures: 5
Address for correspondence:

Josep M. Llovet, MD
Professor of Medicine
Mount Sinai Liver Cancer Program
Division of Liver Diseases
Icahn School of Medicine at Mount Sinai
1425 Madison Avenue, Box 11-23
New York 10029, NY
Phone: 1-2126599503
Fax: 212-849-2574
Email: Josep.Llovet@mssm.edu
ABSTRACT

Primary liver cancer is the second leading cause of cancer-related death worldwide and therefore a major public health challenge. We review hypotheses of the cell of origin of liver tumorigenesis, and clarify the classes of liver cancer, based on molecular features and how these affect patient prognosis. Primary liver cancer comprises hepatocellular carcinoma (HCC), intrahepatic cholangiocarcinoma (iCCA), and other rare tumors, notably fibrolamellar carcinoma and hepatoblastoma. The molecular and clinical features of HCC vs iCCA are distinct, although they have overlapping risk factors and pathways of oncogenesis. A better understanding of the cell types originating liver cancer can aid in exploring molecular mechanisms of carcinogenesis and therapeutic options. Molecular studies have identified adult hepatocytes as the cell of origin. These cells have been proposed to transform directly into HCC cells (via a sequence of genetic alterations), to de-differentiate into hepatocyte precursor cells (which then become HCC cells that express progenitor cell markers), or to trans-differentiate into biliary-like cells (which give rise to iCCA). Alternatively, progenitor cells also give rise to HCCs and iCCAs with markers of progenitor cells. Advances in genome profiling and next-generation sequencing have led to the classification of HCCs based on molecular features, and assigned them to categories such as proliferation–progenitor, proliferation–transforming growth factor beta, and Wnt–catenin beta 1. iCCAs have been assigned to categories of proliferation and inflammation. Overall, proliferation subclasses are associated with a more aggressive phenotype and poor outcome of patients, although more specific signatures have refined our prognostic abilities. Analyses of genetic alterations have identified those that might be targeted therapeutically, such as fusions in the FGFR2 gene and mutations in genes encoding isocitrate dehydrogenases (in approximately 60% of iCCAs) or amplifications at 11q13 and 6p21 (in approximately 30% of HCCs). Further studies of these alterations are needed before these can be used as biomarkers in clinical decision making.
Liver cancer is the second most common cause of cancer-related death worldwide and one of the few neoplasms whose incidence and mortality have been steadily increasing\(^1,2\), with the highest increase in mortality in the United States (US) during the last 2 decades (Figure 1\(^3\)). Liver cancer comprises a heterogeneous group of malignant tumors with different histologic features and unfavorable prognosis that range from hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (iCCA) to mixed hepatocellular-cholangiocarcinoma (HCC-CCA), fibrolamellar HCC, and the pediatric neoplasm hepatoblastoma\(^4,5\). Among these, HCC and iCCA are the most common primary liver cancers, whereas the other neoplasms, including mixed HCC-CCA tumors\(^5\), account for less than 1% of cases. The burden of liver cancer is increasing globally—there could be 1 million cases by 2030\(^6\). It is not clear how direct-acting antiviral (DAA) agents, which can cure hepatitis C virus (HCV) infection, will affect the burden of HCC. It has been estimate that curing more than 90% of patients with HCV infection would eliminate 15% of HCC cases in the US\(^7\). However, there is debate over the effects of DAA agents on progression of HCC\(^8-11\).

HCC alone accounts for 90% of all primary liver cancer cases with nearly 800,000 new annual cases\(^2\). Its highest incidence is in Asia and Sub-Saharan Africa due to the high prevalence of hepatitis B virus (HBV) infection\(^5\). Unlike other cancers, the main risk factors associated with HCC are well defined and include viral hepatitis (B and/or C), alcohol abuse and non-alcoholic fatty liver disease in patients with metabolic syndrome and diabetes. Other co-factors of HCC development, such as aflatoxin B1 and tobacco, increase the incidence of the disease if other common risk factors are present\(^12\).

iCCA is the second most common liver cancer, with the highest incidence in Southeast Asia (30–40 cases/10\(^5\) inhabitants) and low incidence in Western countries (fewer than 5 cases/10\(^5\) inhabitants)\(^13\). Nevertheless, steady increases have been reported\(^13,14\). Risk factors for iCCA development include primary sclerosing cholangitis (PSC), biliary duct cysts, hepatolithiasis and parasitic biliary infestation with flukes—an etiology prevalent in Asia and linked to a specific molecular fingerprint\(^13\). More recently, shared risk factors with HCC have also been identified, such as HBV and HCV—particularly for iCCAs that develop in cirrhotic liver\(^15\).

HCC and iCCA have been considered to be independent tumors that originate from distinct cell populations. However, more recently, some have been recognized as tumor subtypes of a continuum spectrum of diseases. We review the theories behind the cell(s) of origin of liver
cancer, describe emerging molecular classes, link these classes with their etiology and prognosis, and define pathways for future translation.

**CELL(S) OF ORIGIN**

Parenchymal (hepatocytes and cholangiocytes) and non-parenchymal cells (fibroblasts, stellate cells, Kupffer cells, and endothelial cells) form the basic hepatic structure (Figure 2); the existence of stem cells in adult liver has been heavily debated. Hepatocytes constitute 60%–80% of the total liver mass. Architecturally, these cells are organized in lobules, which can be further divided in functional regions or zones. Liver zonation is particularly relevant for hepatocytes, since it affects their function without altering their phenotype. Hepatocytes are frequently polyploid (4N, 8N, etc.); polyploid cells make up 50% of human liver and 90% of mouse liver.

Liver tumors are heterogeneous in morphology, within the same and between different tumors. Some subtypes of HCCs and iCCAs have stem cell features, such as HCCs with CK19-positive cells. The rare, mixed HCC-CCAs have cells with a phenotype that is intermediate between hepatocytes and cholangiocytes. These observations have led to several hypotheses about liver cancer’s cell(s) of origin. Hepatic progenitor cells might generate primary liver tumors because, during liver development, hepatocytes and cholangiocytes each arise from a common progenitor (hepatoblasts). Or, HCC and iCCA might be distinct tumors that originate from mature hepatocytes and cholangiocytes, respectively.

We review findings from recent studies indicating that adult hepatocytes are the source of HCCs, because they can directly degenerate into HCC following sequential genomic insults, differentiate into precursor cells that ultimately can transform into HCC cells with markers of progenitor cells, and trans-differentiate into biliary-like cells able to transform into iCCA (Figure 3). In parallel, progenitor cells may also give rise to HCCs and iCCA with progenitor-like features, whereas adult cholangiocytes, which lack the plasticity and transforming capacity of hepatocytes, can only give rise to iCCAs.

**Progenitor cells and hepatocytes in liver regeneration**

Adult stem cells or somatic stem cells are defined as undifferentiated cells with limitless replicative potential. They can differentiate into all or some specialized cell types, and their primary role is to sustain physiological tissue turnover and to direct tissue repair upon different
types of injury\textsuperscript{27}. Stem cell compartments have been successfully identified in several adult tissues with fast turnover, including the gastrointestinal tract, skin, and bone marrow. Because of their self-renewal potential and long lifespan, these cells are thought to be more prone to malignant transformation\textsuperscript{28, 29} and have been identified as the cell of origin of cancer in several organs, for example skin and intestine\textsuperscript{30-32}. Unlike these organs, however, cells of the adult liver have relatively slow turnover (a period of several months), with an average hepatocyte life span ranging from 200 to 300 days\textsuperscript{33, 34}. In this context, the contribution of stem cells to liver turnover is unclear (reviewed in ref. \textsuperscript{35}).

Hepatocytes are quiescent in normal liver, but their turnover in the liver dramatically accelerates upon hepatic injury or after a significant reduction of liver mass such as following partial hepatectomy. Under these conditions, adult hepatocytes have enormous proliferative potential (capacity to replicate more than 50 times), a feature that has enabled human liver organ donor transplantation as a standard procedure to treat end-state liver disease\textsuperscript{33-35}. There is evidence that regeneration after partial hepatectomy relies predominantly on the replication of hepatocytes, rather than stem cells. However, when the replicative capacity of hepatocytes is severely impaired, such as in patients with acute liver failure or chronic hepatitis, epithelial cells with intermediate hepatocyte–cholangiocyte phenotypes emerge and expand\textsuperscript{36-38}. These cells, also known as oval cells, are thought to localize to the peri-portal area in the canals of Hering and are considered to be bi-potential progenitors\textsuperscript{35, 39} (Figure 2). The capacity of these cells to repopulate the damaged liver has been demonstrated in mice with massive liver damage, restoring more than 80% of hepatocytes\textsuperscript{40}. Other studies from liver injury models have supported the liver stem cell hypothesis, showing proliferation of duct-like cells, frequently termed the ductular reaction, in the portal zone of the hepatic lobule\textsuperscript{41-43}. However, this feature is controversial\textsuperscript{44, 45}.

Other cells have been implicated in regeneration, repair, and eventually carcinogenesis. HCV-related chronic injury can induce ductular metaplasia and proliferation, with mature hepatocytes trans-differentiating to bi-potential oval cells via epigenetic re-programing\textsuperscript{25}. These cells are different from the classical oval cells of the canals of Hering, and may have a more relevant role in hepatocarcinogenesis. However, other sources of precursors have been identified. These include specialized peri-central liver cells, capable of self-renewal in the uninjured liver under the influence of endothelial Wnt signaling\textsuperscript{46}. These pericentral cells express the early liver progenitor marker TBX3\textsuperscript{47} and are diploid. Hybrid periportal cells express hepatocyte markers,
along with low levels of SOX9 and several bile duct genes, are able to repopulate the healthy and diseased liver mass. It is not clear whether these cell populations are bona fide stem or progenitor cells, or distinct hepatocyte subpopulations, and whether they contribute to tumorigenesis.

**Progenitor cells as precursors**

Extensive experimental studies support the hypothesis that progenitor cells might originate liver cancer (Table 1). For example, mice with genetic alterations that affect the Hippo pathway within the liver expand progenitor-like cells and subsequently develop HCC, iCCA, and mixed HCC-CCA. Similarly, liver-specific deletion of the neurofibromatosis type 2 (NF2) tumor suppressor gene in developing or adult mice leads to expansion of progenitor cells without affecting differentiated hepatocytes; these mice develop HCC and iCCA. Furthermore, any mouse hepatic cell, including hepatic progenitors, hepatoblasts, and hepatocytes, that express activated oncogenes (such as H-RAS or SV40LT) can undergo transformation to develop into iCCAs or HCCs (Table 1). There could also be subpopulations of stem or progenitor cells in peribiliary glands, located throughout the biliary tree, that give rise to iCCAs or fibrolamellar HCC. Proteins involved in stem cell renewal and cell fate decisions include MET, which primarily induces hepatocyte differentiation, and EGFR and Notch, which promote cholangiocyte specification. Activation of Notch in mice leads to development of HCC and iCCA. In addition, mutations in genes encoding isocitrate dehydrogenases, frequently detected in iCCAs but rarely in HCCs, inhibit hepatocyte differentiation and promote oval cell proliferation and biliary transformation following liver injury, in cooperation with KRAS mutations (Table 1). Similarly, KRAS mutations and homozygous PTEN deletion in embryonic bi-potential progenitor cells cooperate to induce onset of iCCA. Overall, these experimental studies implicate progenitor cells in the development of the two most common primary liver cancers.

**Adult hepatocytes as cell of origin**

There is significant experimental evidence implicating adult hepatocytes as the cell of origin in liver cancer. Studies using fate-tracing systems have observed that HCC does not originate from progenitor cells, but rather from hepatocytes in both hepatotoxin-induced and carcinogen-free (Mdr2 knock-out) models (Table 1). In particular, these data indicate that Fox1+ cells, which express the progenitor markers EPCAM, SOX9, PROM1, do not contribute to HCC tumorigenesis, in agreement with other studies showing that biliary cells do not give
rise to HCC. In addition, a recent study has demonstrated that hepatocyte-specific p62 expression promotes c-MYC induction, mTORC1 activation and HCC initiation.

Adult hepatocytes, however, have additional properties consistent with a cell that originates cancer. These include their remarkable plasticity, which entails their capacity to de-differentiate into a progenitor state and then can restore the hepatocyte pool. Mature hepatocytes can de-differentiate following loss of the tumor suppressor TP53, which yields nestin-positive progenitor-like cells that can expand and generate primary liver cancers after acquiring lineage-specific oncogenic lesions, such as mutations in WNT (in HCC) or Notch (in iCCA) (Table 1). In addition, adult hepatocytes can also trans-differentiate into biliary-like cells (Figure 3) that can degenerate into iCCA. Upon injury, Notch activation leads to reprogramming of mature hepatocytes into cells that closely resemble biliary epithelial cells, to induce iCCA. Notch can do this on its own or in cooperation with AKT (Table 1).

There is evidence that different cell types can serve as the cell of origin for primary liver cancers (Figure 3, Table 1). Two complementary features of HCC and ICCAs could account for their phenotypic complexity, their molecular features (pathways of activation, mutations, etc.), and the types of cells that become transformed. For example, development of HCC could involve specific molecular alterations in adult hepatocytes (Figure 3), whereas iCCA could originate from cholangiocytes, probably as a result of chronic biliary damage due to liver flukes, PSC, or other insults.

However, adult hepatocytes have also been implicated in the formation of other tumor types. First, tumors with progenitor cell phenotype could derive from progenitor-like cells, either because they are bona fide progenitor cells or following de-differentiation of adult hepatocytes. Transformation of progenitor cells could occur at any stage of hepatic development—from progenitor to adult liver cells—giving rise to tumors of different morphologies, ranging from HCC and iCCA with stem-cell phenotypes to mixed HCC-CCAs (Figure 3). Clinical and pathology analyses of mixed HCC-CCAs have indicated an origin from the progenitor cell compartment. Finally, mature hepatocytes can trans-differentiate, generally as a result of viral infection, to biliary-like cells that give rise to iCCA.

Regardless of their cell of origin (mature hepatocytes vs progenitor cells), liver cancers with stem cell features have a more aggressive clinical behavior and worse prognosis than those without stem cell features. However, these are difficult to study due to discrepancies...
between mice and humans. Studies are needed to determine under what circumstances adult hepatic cells, rather than progenitor cells, give rise to liver tumors and vice versa, and to fill the gap between findings from animal models (cell fate tracing studies) and studies patients with severe liver injury. It will also be important to determine the dominant reprogramming events involved in tumorigenesis in different species. Studies should explore which cells are most affected by oncogene mutations associated with HCC, such as those in the TERT promoter or CTNNB1.

**Immune Response**

Regardless of the cell type that becomes transformed to initiate tumorigenesis, the developing tumor also requires a specific microenvironment. This is particularly relevant to liver tumors, 90% of which develop under conditions of chronic inflammation\(^2\). Factors that contribute to tumor formation include changes in the liver's extracellular matrix, signaling between parenchymal and non-parenchymal cells, and immune dysfunction\(^7\). An altered immune response is an important factor in carcinogenesis.\(^6\) Cancer cells produce growth and angiogenic factors that promote tumor growth, invasion, and metastasis. iCCAs develop under conditions of chronic inflammation, due to viral infections (in about 20% of cases), chronic fluke infestation, PSC, and hepatolithiasis\(^13\).

Chronic inflammation contributes to liver carcinogenesis through a cycle of cell death and regeneration, leading to production of cell survival and proliferation signals that promote formation of regenerative nodules, dysplasia, and cancer\(^7\). In this context, the phenotype of liver tumors could depend on interactions between oncogenes and the immune microenvironment. In a recent animal study, different types of liver tumors, such as HCCs, iCCAs, and mixed HCC-iCCA arise via activation of distinct oncogenes, such as AKT1 and CTNNB1, in combination with inflammatory microenvironment conditions, such as carbon tetrachloride or 3,5-diethoxycarbonyl-1,4-dihydrocollidine\(^7\). Tumors that develop via activation of the same oncogenes can still have different transcriptomes, depending on level of inflammation and features of the microenvironment.

Several mechanisms have linked the immune system with liver cancer development. First, secretion of cytokines by immune cells, such as tumor necrosis factor and IL6, can activate inflammatory signaling pathways in hepatocytes, via JAK–STAT, nuclear factor (NF)-kB, etc., to promote cell proliferation and survival \(^7\). A model of HCC tumorigenesis has been proposed in
which leukocytes that infiltrate the liver organize into ectopic lymphoid-like structures, providing microniches that contain malignant hepatocytic progenitor cells\(^7\). The role of the adaptive immune system is becoming apparent from studies of animal models of liver cancer development with non-alcoholic steatohepatitis. Factors that activate inflammation, such NF-κB and insulin receptor, contribute to carcinogenesis in patients with non-alcoholic fatty liver disease\(^7\). CD8+ and CD4+ T cells and natural killer cells\(^8\) also promote oncogenesis.

The interaction between inflammation and iCCA development has been demonstrated in mice in which biliary epithelial cells that express transgenic AKT and YAP develop only iCCA upon promotion of IL33-mediated biliary tract inflammation\(^8\). Other inflammatory mediators that contribute to oncogenesis include IL6 and iNOS\(^8\). IL6 is highly expressed in human CCAs and promotes cell survival in a STAT3-dependent manner\(^8\). Interestingly, a gene expression signature associated with IL6–STAT3 signaling has been observed in a subset of human iCCAs\(^8\).

Normally, nascent transformed cells are eliminated by both the innate and adaptive immune systems in a process called immune surveillance. Cancer cells express antigens that induce adaptive responses mediated by T cells\(^8\). In the chronically inflamed liver, immune surveillance prevents proliferation and expansion of cancer cells. Some, but not all, recent reports have warned about an unexpectedly higher incidence of HCC occurrence and recurrence in patients treated with DAAs for HCV infection\(^8\)-\(^11\). If confirmed, the potential role of immune system should be explored, because rapid clearance of HCV in patients with chronic infection for years might affect immune surveillance.

**Immune therapies**

The immune microenvironment of HCC is characterized by the co-existence of immune responses leading to both immunogenicity and tolerance\(^8\). HCCs promote immunotolerance through the release of cytokines such as interleukin 10 (IL10) and transforming growth factor beta (TGFB). The tumors and tumor-infiltrating immune cells also express immune regulators such as programmed cell death protein 1 (PD1), its ligand PDL1, and cytotoxic T-lymphocyte protein 4 (CTLA4), which suppress the anti-tumor immune response, allowing tumor growth and progression\(^8\). Tumors with immune infiltrates have been associated with longer survival
times and reduced relapse after resection and transplantation\textsuperscript{2, 88}. In addition, some HCC cases of spontaneous regression have been linked to immune-mediated mechanisms\textsuperscript{2}. These observations establish the rationale for immunotherapy of HCC. Promising responses have been reported with nivolumab, a monoclonal antibody directed against PD1, in a phase 2 trial\textsuperscript{89}. It appears that a T-cell response, features of inflammatory cells (PD1, PDL1 expression), number of mutations in tumor cells, and gene expression profiles might predict patient responses\textsuperscript{50-93}. Using source separation techniques to dissect gene expression patterns of tumor cells from those of tumor-infiltrating immune cells, an immune-specific subclass has been identified in 25% of HCC patients\textsuperscript{94}. This subgroup is characterized by activation of immune cells, expression of PD1 and PDL1, and enrichment in signatures of response to immunotherapies. Furthermore, 2 robust microenvironment-based types have been identified, with either active or exhausted immune responses; these findings could be used to improve design of clinical trials of immunotherapies\textsuperscript{94}. Similarly, a subtype of iCCA characterized by high mutation load and increased expression of immune checkpoint genes has been identified that may predict response to immunotherapies\textsuperscript{95}.

**Classification Based on Molecular Features**

Characterization of molecular subtypes and/or oncogene-addiction loops has generated treatment algorithms for several types of cancer, including breast cancer; these have improved patient outcomes. High-throughput technologies, including single-nucleotide polymorphism arrays, combined with transcriptomic and exome sequencing analyses, have identified molecular subtypes of liver cancer, with distinct oncogene signaling pathways and recurrent mutations\textsuperscript{96}. Even though molecular classifications have not affected decision making for patients with liver cancer, correlations between these subclasses and clinical and pathology characteristics (such as risk factors, outcome, etc.) have been proposed.

**Reaching a consensus**

The development of a molecular classification of HCC has paralleled the progress in genomic profiling technologies, and is extensively reviewed elsewhere\textsuperscript{97}. Initial studies relied on differences in gene expression patterns determined by microarrays, whereas more recent, sequencing-based studies used mutation signatures to classify HCCs\textsuperscript{3}. Unsupervised clustering of gene expression patterns accurately identified 2 major subgroups (reviewed in\textsuperscript{2, 97}). These 2 classes, broadly termed proliferation and non-proliferation, correlate with clinical and pathology features, etiology, and patient outcomes\textsuperscript{97}. Tumors of the proliferation class are highly
heterogeneous with significant enrichment in signaling pathways related to proliferation, such as insulin like growth factor 1 (IGF1) and mechanistic target of rapamycin (MTOR) or stem cell features, such as Notch. This class also contains most gene expression patterns associated with tumor recurrence and shorter survival time (Figure 4). Epigenetic features, such as expression patterns of micro RNAs and DNA methylation patterns, also associate with these subclasses. Chromosomal aberrations are also frequent in HCCs—mostly in chromosomes 1 and 8—which correlate with transcriptome-based subclasses. High-level amplifications have been reported at Chr. 11q13 (at CCND1 and FGF19) and 6p21 (at VEGFA); gains in 11q13 are enriched in tumors from the proliferation subclass (Figure 4). Tumors in the non-proliferation subclass tend to retain hepatocyte-like features; a subset has activation of the canonical WNT signaling pathway, mostly via mutations in CTNNB1. This class contains many less-aggressive, well-differentiated tumors with low levels of alpha fetoprotein, compared to the proliferation class. In parallel, molecular analyses of the adjacent non-tumor tissues have provided valuable insights into mechanisms that promote hepatocarcinogenesis, including aberrant EGF signaling and activation of inflammation pathways (involving NF-kB and IL6).

Next-generation sequencing (NGS) has provided a more detailed picture of molecular alterations in HCCs. This technique can be used to define mutation signatures—recurrent patterns of nucleotide substitutions across different samples. These data provide for an additional layer of stratification beyond transcriptome-based classes. Interestingly, mutation signatures can reflect the exposure history of a tumor. In HCC, specific mutation signatures are associated with environmental exposures such as smoking, alcohol, aflatoxin B1 or aristolochic acid (Figure 4). Data from exome sequencing of large patient cohorts have confirmed the most prevalent mutations in HCC pathogenesis (such as the TERT promoter, CTNNB1, TP53, AXIN1, ARID1A), but also uncovered novel low-frequency events and implicated mutations in non-coding DNAs. In addition, NGS has also provided a map of viral integrations in HCC, mostly related to HBV and more recently for the adeno-associated virus type 2 (AAV2). Recurrent clonal HBV integrations affect TERT, MLL4 and CCNE1, some of which may impact patient survival. AAV2-mediated insertional mutagenesis has been reported in 5% of HCC, mostly in non-cirrhotic patients without a clear etiological factor, predominantly affecting TERT, KMT2B, CCNA2 and CCNE.
Obstacles to clinical implementation

Ultimately, molecular classification of both the tumor and its microenvironment will translate into more effective clinical decision-making and treatment selection. This means optimizing clinical staging systems, both in their ability to predict prognosis and in their capacity to stratify patients based on response to therapies. Despite substantial progress, clinicians cannot yet use molecular information to guide therapy, primarily because there are not yet biomarkers that predict the response to a given therapy, unlike other malignancies. For instance, lung cancer patients with ALK rearrangements have a significantly better response to ALK inhibition with crizotinib than those without ALK rearrangements.111

There are several reasons that molecular information on tumors is not incorporated into clinical practice. Unlike other solid tumors, the most frequent mutations in HCC are un-targetable, including those in the TERT promoter (60%–70%), TP53 (25%–50%) and CTNNB1 (25%–30%). Furthermore, the few potentially actionable events uncovered have not yet been evaluated in clinical trials that include biomarkers. Only recently have some potential biomarkers been tested in early clinical trials, such as FGFR4 inhibitors in patients with high level amplifications of Chr11q13, locus of FGF19. Finally, a therapeutic response to blocking an oncogene in a few patients will be lost among many patients who lack a responsive phenotype if the trial has not been enriched to include those patients whose tumors harbor targetable defects. A recent report suggests that TSC2 loss predicts response to the MTOR inhibitor everolimus in different experimental models of HCC, yet there is no clear biomarker of treatment response in the phase 3 clinical trial. Indeed, the two systemic therapies that increase survival in HCC patients in first (i.e., sorafenib) and second line (i.e., regorafenib) have been tested in ‘all-comers’, with no patient enrichment strategy.

Options to improve the current trial strategy could include conducting clinical trials with enriched populations, based on biomarkers that predict response (FGF19, MET). Or they could investigate non-mutated de-regulated pathways as targets, such as IGF2, or NOTCH signaling. It might also be possible to use alternative anti-tumor approaches, such as immune checkpoint inhibitors, or strategies to antagonize the most prevalent genetic defects found in HCCs, such as mutations in TERT, TP53 and CTNNB1.
Molecular Subclasses of iCCA

There is no effective treatment for iCCA. Reasons include our poor understanding of its pathogenesis and the lack of studies conducted specifically in patients with iCCA, apart from patients with all biliary tract cancers\textsuperscript{13, 14}; this is important because iCCAs are characterized by a distinct set of mutations\textsuperscript{14}. iCCAs were only recently recognized as a distinct entity with both its own unique staging system (see the 7th edition of the American Joint Committee on Cancer Staging Manual\textsuperscript{118}) and ad hoc practice guidelines [see guidelines of the International Liver Cancer Association (ILCA)\textsuperscript{13}]. The classification of iCCA’s distinct phenotype has led to the first studies that characterize its genomic landscape in order to generate an initial molecular classification. Other CCA, such as perihiliar (pCCA) and distal (dCCA) are different molecular entities\textsuperscript{14}.

Two main molecular subtypes of iCCA, proliferation and inflammation, have been proposed based on an integrative analysis of whole-genome expression patterns, chromosomal aberrations, mutation, and alterations in signaling pathways in a large cohort from USA and Europe\textsuperscript{85}. The inflammation subclass was characterized by activation of inflammatory pathways, overexpression of cytokines, and STAT3 activation, whereas the proliferation subclass was characterized based on activation of oncogene signaling pathways (KRAS, EGFR, and Notch) and HCC gene expression signatures associated with short survival times and earlier recurrence\textsuperscript{85}. While neither of the two subclasses displayed specific broad chromosomal gains and losses, the proliferation subclass is enriched in focal aberrations including DNA amplifications at 11q13.2 (locus of CCND1 and FGF19), and deletions at 14q22 (locus of SAV1). The proliferation class also includes a subtype with stem cell–like features\textsuperscript{119}, chromosomal instability\textsuperscript{85} and mutations in isocitrate dehydrogenase genes\textsuperscript{63} (Figure 5). Clinical characteristics such as moderate/poor differentiation, intra-neural invasion and poor outcome are significantly more frequent in the proliferation class. Interestingly, there is a genomic resemblance observed between the proliferation subclass of iCCA and HCC subtypes with poor prognosis (i.e. cluster A\textsuperscript{19}, G3\textsuperscript{120}) and stem cell features (S2)\textsuperscript{121, 122} (Figure 5). This evidence further supports the presence of a common cell of origin in liver cancers.

An independent whole-transcriptomic analysis of a large cohort including iCCA and extrahepatic cholangiocarcinomas also confirmed the existence of the two prognostic CCA molecular subtypes\textsuperscript{123}. In this study, two main classes were reported: a poor prognosis group characterized by enrichment of RAS mutations, MET and EGFR overexpression, and a good
prognosis subgroup characterized by signatures of immune response. Genomic overlap between the poor prognosis group and the proliferation subclass was observed.

Similar to HCC, NGS studies have generated insights into the molecular subtypes of iCCA. The most common genetic aberrations include FGFR2 gene fusions (~25%) and mutations in KRAS (20%), IDH1/2 (~20%) and chromatin remodeling genes [ARID1A, BAP1, PBRM1, ~30%] (reviewed elsewhere). Besides IDH1 and IDH2, KRAS, and EGFR mutations, which are enriched in the proliferation subclass, no significant distribution of other molecular aberrations has been reported in any of the iCCA molecular subtypes.

The incidence of iCCA varies significantly among geographic regions. Incidence is highest in Asian countries, where infections with liver flukes such as Opisthorchis viverrini and Clonorchis sinensis are the major risk factor. In non-endemic regions, other factors, including hepatolithiasis and PSC, likely contribute to risk. Associations between different risk factors and mutations detected in iCCAs were explored in 209 CCAs from Asia and Europe. Mutations in BAP1 and IDH1/2 were found in a higher proportion of non-O viverrini-related iCCAs (22% vs 3.2%), whereas mutations in TP53 were found in a higher proportion of tumors from patients with O viverrini infection (45.2% vs 7.4%)..

Although there have been other studies of large cohorts of patients, it has been difficult to draw conclusions from these due to the different types of neoplasms included (iCCAs, extrahepatic cholangiocarcinomas, and gallbladder tumors).

Importantly, multiple FGFR2 gene fusions have been discovered in approximately 25% of iCCAs (reviewed in ref.); these are used to define tumor subtypes. Tumors with FGFR2 fusions and KRAS mutations or BAP1 mutations have been reported. No significant differences in patient age or outcome, or tumor differentiation or stage, were found in comparing patients with or without the FGFR2 fusion. One study reported a higher frequency of HCV infection in patients with FGFR2 fusions, although this finding has not been confirmed. Early-phase clinical trials testing selective FGFR inhibitors in advanced iCCA patients carrying alterations in FGFR2 are underway (NCT02150967 and NCT01752920). These studies reported an objective response rate of approximately 15% (a signal of efficacy) and manageable safety profile. The discovery that FGFR2 fusions can promote development of iCCAs has been quickly translated to clinical trials. Fusions of the FGFR2 gene might be used as biomarker to select patients for treatment with these agents.
Molecular Features of Uncommon Tumors

Fibrolamellar HCC (FLC) is a rare primary liver cancer with few treatment options that typically affects children and young adults without background liver disease\textsuperscript{60}. FLC is now considered a unique molecular entity within primary liver malignancies\textsuperscript{138}. FLC is caused by an approximate 400 kb deletion in chromosome 19 that incorporates the first exon of \textit{DNAJB1} into all but the first exon of \textit{PRKACA}\textsuperscript{139}. This fusion is detected in 70\%\textsuperscript{138} to 100\% of FLCs\textsuperscript{139}. In addition, exome sequencing and RNA sequencing studies found de-regulation of additional actionable candidates such as \textit{ERBB2} and \textit{AURKA}\textsuperscript{140}. Unlike HCC or iCCA, there are no prevalent mutations in more than 10\% of patients. Unsupervised clustering of gene expression data identified FLC molecular subclasses in a cohort of 58 FLCs, 2 of them associated with proliferative and inflammatory molecular traits. The same study provided a prognostic 8-gene expression signature\textsuperscript{138}. FLC has less chromosomal aberrations compared to HCC or iCCA without recurrent high-level amplifications or deletions.

Hepatoblastoma is the most frequent primary liver tumor in children younger than 5 years old. Like FLC, background liver disease is rare in these patients. WNT signaling plays a major role, with \textit{CTNNB1} mutations (70\%) as the most frequently reported molecular event. Integrative analysis of gene expression and chromosomal aberrations in 24 hepatoblastoma samples distinguished 2 classes, with different DNA gains, histological phenotype and clinical outcome\textsuperscript{141}. The recent use of patients-derived xenografts (PDX) in hepatoblastoma enabled the identification of novel therapeutic targets such as \textit{NRAS} mutations\textsuperscript{142,143}.

Mixed HCC-iCCA is a rare neoplasm accounting for less than 1\% of all primary liver cancers\textsuperscript{5}. Diagnosis is mainly based on histological examination and requires features of HCC and iCCA. According to the 2010 World Health Organization classification, there is a classical type, characterized by areas of typical HCC and iCCA with a transition area, and a stem cell-feature type\textsuperscript{144}. The stem cell-feature type can be further divided into typical, intermediate, and cholangiolocellular carcinoma (CLC) subtypes. Gene expression profiling of small series of HCC-CCA samples indicated that this tumor type might share common characteristics with poorly differentiated HCC and iCCA with stem cell traits\textsuperscript{119,122,145}. A comprehensive integrative genomic analysis of 18 mixed HCC-CCA subtypes defined 3 molecular types (CLC, stem cell, and classical subtype)\textsuperscript{146}. In particular, the CLC subtype, despite being grouped, based on histology, within the stem cell-feature subtype, has a unique molecular profile, with biliary traits,
low levels of chromosome instability, and activation of TGFB and inflammation-related signaling pathways\textsuperscript{146}.

**Challenges in tumor heterogeneity**
Liver tumors have a high degree of molecular heterogeneity, not only among patients with the same types of tumors, but also within regions within the same tumor or tissue\textsuperscript{147}. Expert recommendations and consensus definitions have been established to better address the extent and role of cancer heterogeneity in oncology research\textsuperscript{148}.

From a clinical perspective, molecular heterogeneity should be considered within the concepts of trunk and branch, and driver and passenger mutations. Trunk mutations (mutations that occur early in tumorigenesis) can be found in every cancer cell of a tumor \textsuperscript{148}. Branch mutations are found in only a subgroup of tumor cells; these are acquired at later stages of tumor development, frequently following treatment, and are associated with resistance. Studies of other malignancies have indicated that trunk mutations promote tumor progression, and are therefore the best targets for therapies. Trunk mutations identified in single-biopsy should be the first targets of molecular therapies \textsuperscript{2}. Finally, passenger mutations account for the most common type of mutations, but their relevance to tumor development is marginal. However, they are often used to define a tumors’ clonal features and anti-tumor response.

**Prognostic Markers**
Molecular biomarkers have many applications in cancer. The can improve clinical decision making by helping to predict a patient’s outcome, select the patients most likely to respond to a specific treatment, and avoid treatments not likely to be effective. Unlike other cancers such as breast\textsuperscript{149} and colorectal\textsuperscript{150} cancers, HCC has no molecular markers that have been incorporated into clinical management.

As a general rule, clinical practice guidelines include recommendations based on a specified set of metrics, which rate the level of evidence and the strength of recommendations for each statement; the higher the evidence, the stronger the recommendation. For example, sorafenib is recommended for treatment of patients with advanced-stage HCC; a strong recommendation was derived from a prospective, double-blind, randomized, placebo-controlled trial (the highest level of evidence)\textsuperscript{151}. These metrics are less clear when evaluating biomarkers for HCC. Initial recommendations for biomarkers were developed mostly to address technical aspects of the
assay such as the Tumor Marker Utility Grading System. This was later improved upon by the REMARK guidelines for reporting prognostic biomarker in oncology, which also has addressed issues related to clinical utility. The PROGRESS initiative was the most recent attempt to provide a methodological framework for biomarker research.

The highest level of evidence in biomarker research (Level A) comes from randomized trials. A large prospective study of 10,253 women demonstrated that a 21-gene expression signature could identify patients with early-stage breast cancer at high risk for recurrence. Level B evidence is provided by companion studies of biomarkers in randomized clinical trials, such as those reported for HCC in the Sorafenib HCC assessment randomized protocol and everolimus for liver cancer evaluation trials. Neither of these studies identified markers associated with response to either sorafenib or everolimus, but they confirmed the prognostic role of levels of angiopoietin 2 and vascular endothelial growth factor in plasma of patients with advanced-stage HCC. Level C evidence has been provided most frequently for markers of HCC, from studies of archived samples, with extensive validation (reviewed elsewhere). Most level C studies have evaluated prognostic markers, identifying patients with the proliferation subclass of tumor, either S1 (WNT and TGFB expression), or the S2 (progenitor cell) subtype, with or without amplification of FGF19. The S2 subtype expresses many makers of progenitor cell cells and high levels of alpha fetoprotein. Tumors in the S1–S2 subgroups are poorly differentiated and patients have worse outcomes than patients with other HCC subtypes. Level D evidence comes from studies of convenience, in which specimens were collected for unknown reasons and happen to be available for assay.

Gene signatures associated with risk of HCC development may have a more direct application. For instance, a 186-gene signature in liver tissues was able to predict occurrence of HCC in patients with hepatitis C-related early-stage cirrhosis, and might be used to select patients for trials of chemopreventive agents (Figure 4, Table 2). However, chemoprevention studies are a challenge, because it takes a long time to collect data on a low number of events. Studies that included populations of patients at high-risk for HCC may shorten trial time and identify effective agents more quickly. For predictive biomarkers in advanced HCC, ongoing clinical trials include populations selected based on high expression of MET in tumor cells (trial of tivantinib, NCT01755767) or high serum level of alpha fetoprotein (trial of ramucirumab, NCT02435433), both as second-line therapies.
Molecular markers of iCCA

Even fewer biomarkers of iCCA have been developed than for HCC; there is only level C data for the efficacy of a biomarker for this tumor type. iCCA is one of the few solid tumors for which no systemic molecular therapy has been approved by the Food and Drug Administration\textsuperscript{13}. Surgery is the only potentially curative option, although chemotherapy with gemcitabine plus cisplatin is the standard for patients with advanced-stage iCCA\textsuperscript{162}. According to the ILCA practice guidelines, no studies have provided level A or B evidence for a marker of this tumor—some data from low-quality studies have led to weak recommendations\textsuperscript{13}. These studies have been retrospective in design, without extensive validation in independent cohorts. Some gene expression and microRNA patterns have been associated with features of iCCA\textsuperscript{85, 119, 123}, as well as mutations in KRAS and IDH1 or IDH2 and focal deletions at 14q22.\textsuperscript{85, 128, 129} (Table 2). Gene expression patterns have been reported to identify an iCCA proliferation subclass of tumor\textsuperscript{85}, patients with a poor prognosis, \textsuperscript{85, 123} and tumors with a stem cell–like subtype\textsuperscript{119}, who have aggressive disease and worse outcome.

Features of cancer cells in a tumor determine not only the tumor’s growth and progression, but also affect the tumor microenvironment (non-cancer cells and supporting stroma)\textsuperscript{76}. Some studies associated features of the iCCA microenvironment with patient outcome\textsuperscript{123, 163-165}. Two studies identified a stromal signature associated with shorter survival time by combining laser capture microdissection with gene expression profile analysis\textsuperscript{123, 165}. Although many studies have validated gene expression signatures associated with patient outcomes, the clinical significance of microRNA signatures is not yet clear, due to the limited number of large comprehensive profiling studies\textsuperscript{119, 166-168}.

Mutations in oncogenes such as KRAS and BRAF can be used to determine prognosis for patients with pancreatic or colorectal cancers\textsuperscript{169-171}. However, there is controversy over whether these mutations can predict outcomes of patients with iCCA. Some studies have reported significant associations between mutations in KRAS\textsuperscript{128, 129, 131} or chromatin-remodeling genes (BAP1, PBMR1, and ARID1A)\textsuperscript{125, 172} and either shorter disease-free survival or shorter overall survival, respectively. It is not clear whether mutations in IDH1 or IDH2 associate with overall survival\textsuperscript{125, 129, 173}.

Most biomarker assays use cells or tissue collected by invasive biopsy procedures or from resection specimens, but these specimens are rarely collected from unresectable or metastatic
iCCAs—the tumors of most patients in clinical trials for iCCA. Strategies should be developed to assess circulating markers (called a liquid biopsy), from tumor cells and cell-free nucleic acids in blood\textsuperscript{174}. Detection of molecular signatures or aberrations in the blood could significantly facilitate implementation of prognostic markers. Liquid biopsies might also be used to monitor response to treatment, especially in clinical trials of selective inhibitors of IDH1 and IDH2 (NCT02073994, NCT02273739) and FGFR2 (NCT02150967)\textsuperscript{14,136}.

Future Directions

In the last decade, there have been tremendous advances in our understanding of the cellular and molecular complexity of liver cancers. Sophisticated functional studies and NGS-based research studies have provided information about the cells that form liver tumors and the proteins and pathways involved in hepatocarcinogenesis. These studies support a model of multiple cells of tumor origin. Most of the studies, using in vivo lineage tracing models, revealed an unexpected plasticity of mature hepatocytes\textsuperscript{24, 25, 44, 52, 72} that could be involved in their transformation into cancer cells. Hepatic progenitor cells might also become transformed to give rise to different subtypes of HCC, iCCA, and mixed HCC-iCCA with progenitor markers.

Studies of humanized mice to mimic the onset of liver cancers in multiple settings may help address existing challenges. In addition, cell- and lineage-tracing models should improve so that they recapitulate the severity of liver injury observed in humans, and must clarify if hepatic progenitor cells can replenish the liver mass after marked liver injury, as demonstrated in zebrafish models\textsuperscript{175}. Finally, linking the main genomic hits at pre-neoplastic and early stages of human cancer development with the cell of origin is also an area where further studies are required.

Identifying the cells that give rise to liver tumors and elucidating the different classes of tumors, based on their molecular features, could lead to new approaches to treatment and prognosis. Our knowledge of the molecular mechanisms of HCC and iCCA pathogenesis has not yet been translated into clinical tools, because few studies have linked specific molecular classes of tumors or aberrations to responses to therapy. It is important to incorporate biomarker studies into trials, as in the trials of patients with iCCAs with FGFR2 aberrations\textsuperscript{136}. The successes of immune therapies in different solid tumors, including HCC,\textsuperscript{176} indicate the need to focus greater attention on the hepatic tumor microenvironment and interactions between the tumor and the immune response; we also need to identify markers of response to these therapies. Learning
more about the effects of tumor heterogeneity, and the use of liquid biopsies to determine the molecular and genetic features of patients' tumors, could overcome barriers to personalized treatment of liver cancer\textsuperscript{174}. 
FIGURE LEGENDS

Figure 1: Mortality trends of patients with different malignancies in the USA from 1990 through 2009 (reprinted from ref 3). Changes in cancer mortality, among tumor types, in the US. Mortality from liver and bile duct cancers is increasing more rapidly than any other cancer, in men and women. Data obtained from the 2013 American Association for Cancer Research Cancer Progress Report.

Figure 2: Structure of the liver lobule and location of candidate cell of origin of liver cancer. Intrahepatic organization of the liver lobule and the localization of hepatic cells that form liver tumors. Different types of cancer (i.e. HCC, iCCA, mixed HCC-iCCA) can originate in the liver depending on the transformation event and the cell type undergoing neoplastic transformation. Hepatic stem or progenitor cells are believed to reside within the most terminal branches of the biliary tree (referred to as Canals of Hering). Upon injury hepatocytes can undergo reversible ductal metaplasia and de-differentiate into hepatocyte-derived progenitor-like cells.

Figure 3: Multiple cells of origin of primary liver cancers. HCC and iCCA can develop via transformation of mature hepatocytes and cholangiocytes, respectively. There is evidence that hepatic progenitor cells (HPCs), their intermediate states, or de-differentiated hepatocytes can form liver tumors with progenitor-like features, including mixed HCC-CCA, such as cholangiolocellular carcinoma (CLC). Mature hepatocytes can be also re-programmed into cells that closely resemble biliary epithelial cells and contribute to development of iCCA.

Figure 4: Molecular biomarkers of HCC. Genetic features and gene expression of the tumor or its surrounding microenvironment correlate with clinical and pathology features, etiology, and patient outcomes. The proliferation subclass is characterized by proliferation signaling pathways, which involve proteins such as MTOR, Notch, and TGFB and several gene expression signatures associated with poor outcomes of patients (such as the G3 and 5-gene signatures). The non-proliferation subclass is characterized by mutations in CTNNB1, the S3 gene expression signature, and the classical WNT signaling pathway. Specific mutations associated with etiology (smoking and aflatoxin B exposure) have also been described. An expression pattern of 186 genes in the liver tissue surrounding the HCC can identify patients HCC at higher risk for tumor recurrence after resection.
Figure 5: Main characteristics of subclasses of iCCA. Analyses of gene expression patterns have shown that iCCAs can be assigned to the proliferation or inflammation subclasses. The proliferation class is characterized by chromosome instability, activation of pathways related to cell cycle progression, mutations in oncogenes and shorter survival times of patients. Tumors of the inflammation subclass have gene expression patterns associated with activation of an immune response and less aggressive clinical behavior. FGFR2 fusion events are evenly distributed between these subclasses.

ACKNOWLEDGMENTS
The authors would like to thank Jill Gregory for her contribution in the design work of the figures.
REFERENCES


Notes: Author names in bold designate shared co-first authorship
Table 1. Experimental models supporting the different hypotheses regarding the cell of origin in primary liver cancers

<table>
<thead>
<tr>
<th>Hypothesis tested</th>
<th>Method</th>
<th>Animal models</th>
<th>Altered pathway</th>
<th>Tumor Type</th>
<th>Proposed cell of origin</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPC/oval cells as cell of origin</td>
<td>Clonal analyses</td>
<td>GEMM</td>
<td>Sav1fl/fl, mst1fl/fl, and mst2fl/fl</td>
<td>BMI1, WNT</td>
<td>mixed HCC-iCCA</td>
<td>HPCs</td>
</tr>
<tr>
<td></td>
<td>GEMM</td>
<td>WW45flox/floxAlbumin-Cre</td>
<td>HIPPO</td>
<td>HCC, iCCA</td>
<td>oval cells</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Multi-lineage differentiation</td>
<td>GEMM</td>
<td>Alb-Cre;N/2lox/lox</td>
<td>HIPPO</td>
<td>mixed HCC-iCCA</td>
<td>oval cells</td>
</tr>
<tr>
<td></td>
<td>GEMM</td>
<td>NotchIC::AlbCre</td>
<td>H-Ras, SV40LT</td>
<td>HCC, iCCA</td>
<td>HPCs</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>GEMM</td>
<td>Alb-Cre;LSL-IDH2R172K;LSL-KrasG12D</td>
<td>NF2</td>
<td>HCC, iCCA, mixed</td>
<td>oval cells</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>GEMM</td>
<td>LSL-KrasG12D -PtenfloX</td>
<td>Kras, KRAS</td>
<td>iCCA</td>
<td>hepatoblasts, cholangiocytes</td>
<td>62</td>
</tr>
<tr>
<td>Hepatocytes as cell of origin</td>
<td>GEMM</td>
<td>AllpCre+Trp53Δ2-10Δ2-10</td>
<td>TP53, WNT</td>
<td>HCC</td>
<td>hepatocytes-derived HPC</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>GEMM</td>
<td>Alb-Cre;Trp53Δ2-10Δ2-10</td>
<td>-</td>
<td>ICCA</td>
<td>hepatocytes-derived HPC</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>GEMM</td>
<td>(Mst1/Mst2 DKO, YAP KO, DKO Yap+/-, and tTKO</td>
<td>HIPPO</td>
<td>HCC, iCCA</td>
<td>hepatocytes-derived HPC</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>liver injury (BDL, DDC, MDA, CDE); GEMM (Mdr2Δ-/-)</td>
<td>H-Ras, SV40LT</td>
<td>HCC</td>
<td>HPC, hepatoblasts, cholangiocytes</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>liver injury (DEN); GEMM (Mdr2Δ-/-; PTEN-/-)</td>
<td>MDR</td>
<td>HCC</td>
<td>hepatocytes</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DEN + TAA</td>
<td>-</td>
<td>HCC, HPA</td>
<td>hepatocytes</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tsc1/Sqstm1Δhep and Sqstm1Δhep/MUP</td>
<td>p62</td>
<td>HCC</td>
<td>hepatocytes</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Alb-CreERT2;R26RlacZ/+ and CK19-CreERT2;R26RlacZ/+</td>
<td>NOTCH</td>
<td>iCCA</td>
<td>hepatocytes</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hydrodinamic tail injection</td>
<td>NOTCH, AKT</td>
<td>iCCA</td>
<td>hepatocytes-derived biliary-like cells</td>
<td>73</td>
</tr>
<tr>
<td>Cholangiocytes as cell of origin</td>
<td>lineage-tracing system</td>
<td>GEMM</td>
<td>Liver injury (TAA) and CK19CreERTeYFPp53fl</td>
<td>TP53</td>
<td>iCCA</td>
<td>cholangiocytes</td>
</tr>
<tr>
<td></td>
<td>lineage-tracing system</td>
<td>GEMM</td>
<td>LSL-KrasG12D -PtenfloX</td>
<td>KRAS, PTEN</td>
<td>iCCA</td>
<td>hepatoblasts, cholangiocytes</td>
</tr>
</tbody>
</table>

**Abbreviations:** HPC, Hepatic progenitor cells; GEMM: genetically engineered mouse model; HCC, hepatocellular carcinoma; HPA, hepatocellular adenoma; iCCA, intrahepatic cholangiocarcinoma; mixed HCC-iCCA, mixed hepatocellular-cholangiocarcinoma; KO, knock-out; DKO, double knock-out; TKO, triple knock-out; DEN, N-nitrosodiethylamine; TAA, Thioacetamide; BDL, bile duct ligation; MDA, methylene dianiline; CDE, choline-deficient ethionine-supplemented; DDC, 3,5-diethoxycarbonyl-1,4-dihydrocollidine.
<table>
<thead>
<tr>
<th>Signature</th>
<th>Cohort (etiology)</th>
<th>Number of patients tested</th>
<th>Clinical endpoint</th>
<th>REMARK recommendations</th>
<th>Level of evidence (Simon et al. JNCI 2009)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hepatocellular Carcinoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mRNA-based</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-gene signature</td>
<td>HCC (Alcohol, HCV, HBV)</td>
<td>189/434</td>
<td>survival</td>
<td>OK</td>
<td>C</td>
<td>100</td>
</tr>
<tr>
<td>EpCAM-signature</td>
<td>HCC (HBV)</td>
<td>40/238</td>
<td>survival</td>
<td>OK</td>
<td>C</td>
<td>160</td>
</tr>
<tr>
<td>186-gene signature (adjacent tissue)</td>
<td>HCC (HCV)</td>
<td>82/441</td>
<td>survival, HCC development</td>
<td>OK</td>
<td>C</td>
<td>103, 158</td>
</tr>
<tr>
<td>miRNA based</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-miRNA signature</td>
<td>HCC (HBV)</td>
<td>131/110</td>
<td>venous metastasis, survival</td>
<td>OK</td>
<td>C</td>
<td>161</td>
</tr>
<tr>
<td>Down-regulation miR-26a</td>
<td>HCC (HBV)</td>
<td>241/214</td>
<td>survival</td>
<td>OK</td>
<td>C</td>
<td>159</td>
</tr>
<tr>
<td>DNA methylation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36 CpG DNA methylation signature</td>
<td>HCC (HCV)</td>
<td>221/83</td>
<td>survival</td>
<td>OK</td>
<td>C</td>
<td>102</td>
</tr>
<tr>
<td>Protein markers (Plasma)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ang2, VEGF</td>
<td>HCC (HCV)</td>
<td>491</td>
<td>survival</td>
<td>OK</td>
<td>B</td>
<td>157</td>
</tr>
<tr>
<td>Cholangiocarcinoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mRNA-based</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>poor prognosis/proliferation class</td>
<td>iCCA</td>
<td>119/104</td>
<td>survival and recurrence</td>
<td>further validation required</td>
<td>C</td>
<td>85</td>
</tr>
<tr>
<td>poor prognosis group</td>
<td>iCCA and eCCA</td>
<td>52/52</td>
<td>survival and recurrence</td>
<td>further validation required</td>
<td>C</td>
<td>123</td>
</tr>
<tr>
<td>Stem cell-like subtype</td>
<td>iCCA and mixed HCC-CCA</td>
<td>23/68</td>
<td>survival</td>
<td>further validation required</td>
<td>C</td>
<td>119</td>
</tr>
</tbody>
</table>

**Abbreviations:** HCC, hepatocellular carcinoma; iCCA, intrahepatic cholangiocarcinoma; mixed HCC-iCCA, mixed hepatocellular-cholangiocarcinoma; eCCA, extrahepatic cholangiocarcinoma.
Figure 4

Gene expression microarray

Molecular subclasses
- Proliferation subclass
  a) AKT/MTOR, MET, TGFβ, IGF, NOTCH signaling
  b) Cluster A, G3, S1-2, 5-gene, Late-TGFβ signatures
  c) HLA 11q13 (FGF19/CCND1)
  d) Poor outcome
- Non-Proliferation subclass
  a) Classical WNT (CTNNB1 mutations)
  b) Cluster B, S3, WNT/CTNNB1, Poly 7
  c) Chr7 gain
  d) Good outcome

Outcome prediction

Mutational profile
- Candidate drivers/therapy

Mutational signature
- Etiology (i.e. smoking, aflatoxin B1, aristolochic acid, etc)

Molecular subtype
- High risk of recurrence

Adjuvant therapy
### iCCA Molecular Classes

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Proliferation Class</th>
<th>Inflammation Class</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subtype</strong></td>
<td>Progenitor-like iCCA</td>
<td>Inflammatory pathways (Interleukins/chemokines), STAT3 activation</td>
</tr>
<tr>
<td><strong>Signalling Pathways</strong></td>
<td>KRAS/RAF/ERK, IGF1R, EGFR, NOTCH, HER2, MET, PI3K/AKT/mTOR</td>
<td></td>
</tr>
<tr>
<td><strong>Prognostic Signatures</strong></td>
<td>Poor prognostic signatures (i.e. G3, S1-2, Cluster A, stroma, iCCA-like HCC)</td>
<td>Stem cell like iCCA</td>
</tr>
<tr>
<td><strong>DNA Methylation</strong></td>
<td>Hyper-methylation</td>
<td></td>
</tr>
<tr>
<td><strong>Oncogenic Mutations</strong></td>
<td>KRAS, BRAF, EGFR</td>
<td>IDH1/2</td>
</tr>
<tr>
<td></td>
<td>Chromatin remodelling genes (ARID1A, BAP1, PBMR1), TP53</td>
<td></td>
</tr>
<tr>
<td><strong>Structural Aberrations</strong></td>
<td>Chr 11q13 amplif. (CCND1, FGF19)</td>
<td>Chr14q22.1 (SAV1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Focal deletion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FGFR2 translocations</td>
</tr>
<tr>
<td><strong>Clinical Features</strong></td>
<td>Intra-neural invasion</td>
<td>Well differentiated</td>
</tr>
<tr>
<td></td>
<td>Moderate/poorly differentiated</td>
<td>Good prognosis</td>
</tr>
<tr>
<td></td>
<td>Poor prognosis (survival/recurrence)</td>
<td></td>
</tr>
</tbody>
</table>