Chapter 6
Molecular Subtypes and Genomic Signatures of Hepatocellular Carcinoma for Prognostication and Therapeutic Decision-Making

Sung Hwan Lee, Sun Young Yim, Jae-Jun Shim, and Ju-Seog Lee

EPCAM

Introduction

New technological developments have frequently preceded major advances in biomedical research and medicine [1]. For example, the development of fluorescent DNA sequencing technique made it possible to establish the large-scale high-throughput technology needed for the human genome sequencing. Also, polymerase chain reaction (PCR), fluorescent DNA sequencing, and other techniques have enabled the discovery of over 6000 Mendelian disease genes [2]. The advent of the DNA sequencing technologies has now made it possible to measure every alteration in human genome and expression of all coding and noncoding genes in different tissues under variety of conditions. This high-throughput technology has therefore afforded biomedical scientists a unique opportunity to integrate the descriptive characteristics (i.e., "phenotype") of a biological system under study with the genomic readout (i.e., mutations, copy number alteration, and RNA expression). The opportunity to contemplate the integrated view of biological systems has provoked a shift in biological sciences away from the classical reductionism to systems biology [1, 3, 4]. The systems approach to a disease is based on the hypothesis that disease processes perturb the regulatory network of genes and proteins in a way that differ from the respective normal counterpart. Consequently, by using multiparametric measurements, it may be possible to transform current

S. H. Lee · S. Y. Yim · J.-J. Shim · J.-S. Lee (✉)
Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
e-mail: jlee@mdanderson.org

Y. Hoshida (ed.), *Hepatocellular Carcinoma*, Molecular and Translational Medicine, https://doi.org/10.1007/978-3-030-21540-8_6

diagnostic and therapeutic approaches and enable a predictive and preventive personalized medicine [4].

The application of next-generation DNA sequencing technologies to characterize tumors at the gene level has had significantly impacted clinical oncology [5–7]. In particular, global gene expression analysis of various human tumors has resulted in the identification of gene expression patterns or signatures related to tumor classification, disease outcome, and response to therapy. The microarray and DNA/RNA sequencing technologies have also been used to investigate the mechanism of action of specific cancer therapeutics.

Clinical Staging of Hepatocellular Carcinoma

It is well established that cancer even in the same tissue is a very heterogeneous disease that differs widely in clinical outcomes and in response to therapy. It is now clear that this heterogeneity is due to different molecular defects that can induce similar tumor phenotypes. Although histopathological and biochemical markers constitute important tools for identifying groups of tumors that differ with respect to prognosis and responses to treatments, the genes and molecular pathways associated with these markers have not been comprehensively defined. Global gene expression analysis of human tumors has already revealed the identification of gene expression patterns or signatures related to tumor classification, prognosis, and response to therapy [8, 9].

The goal of all staging systems is to separate patients into groups with homogeneous prognosis, which then form the bases for the selection of most appropriate treatments. Much work has been devoted to establishing prognostic models for hepatocellular carcinoma (HCC) by using clinical information and pathological classification in order to provide information at diagnosis on both survival and treatment options [10–15]. Although much progress has been made, many issues still remain unresolved. For example, a staging system that reliably separates patients with early HCC as well as intermediate to advanced HCC into homogeneous groups with respect to prognosis does not exist. This is of particular importance because the natural course of early HCC is unknown and the natural progression of intermediate and advanced HCC are known to be quite heterogeneous [15]. This is especially troublesome since the accuracy of imaging techniques is rapidly evolving and affording detection of early HCC nodules [16, 17]. Although the pathological diagnosis of high-grade dysplastic nodules (DN) and early HCC is at present controversial, it is likely that many HCCs evolve from the DN [18]. However, prognostic predictions based on morphological characteristics of these early lesions are still tentative. Due to early surveillance program and improvement in imaging systems, early-stage and small HCCs at diagnosis dramatically increased [19]. Prognosis of early-stage HCC is not well understood, and conventional parameters such as number of tumors and size of tumors may not well account for response to treatment and prognosis. Thus, future classification will have to identify new more relevant vari-

ables that discriminate between patients with small early HCC without any vascular or extrahepatic extension.

Molecular Profiling of Hepatocellular Carcinoma

Numerous studies dealing with gene expression profiling of HCC have appeared during the last 15 years (Table 6.1). The molecular profiling of HCC presents challenges that are not commonly seen in other human tumors. This is primarily due

Table 6.1 Clinically relevant genomic subtypes of HCC

First author	Year	Primary endpoint	Number of genes	Experimental platform	Reference
Norio Iizuka	2003	Prediction of early intrahepatic recurrence of HCC after curative resection	12	Oligonucleotide microarray	[25]
Yukinori Kurokawa	2004	Prediction of early recurrence of HCC	20	PCR-based array	[26]
Qing-hai Ye	2003	Prediction of hepatitis B virus-positive metastatic HCC	153	cDNA microarray	[27]
Stephanie Roessler	2010	Prediction of tumor relapse in early-stage HCC Patients	161	NCI oligo set microarray	[28]
Hyun Goo Woo	2008	Prediction of HBV-related HCC	628	Affymetrix U133A 2.0 array	[29]
Beatriz Mínguez	2011	Prediction of microscopic vascular invasion in HCV-related HCC	35	Affymetrix HG-U133 plus 2 array	[32]
Jean–Charles Nault	2013	Prediction of HCC recurrence after curative resection	5	Affymetrix HG133A array	[34]
Ju-Seog Lee	2004	Classification of prognostic subclass and prediction of overall survival in HCC	406	Oligonucleotide microarray	[8]
Ju-Seog Lee	2004	Comparison of the molecular features of mouse and human HCCs	329	Oligonucleotide microarray	[42]
Ju-Seog Lee	2006	Prediction of the cellular origin of the tumor (hepatobalst vs. hepatocyte)	907	Oligonucleotide microarray	[9]
Sandrine Boyault	2007	Transcriptome classification of HCC and potential therapeutic target	16	Affymetrix HG133A array	[48]
Taro Yamashita	2008	Classification system defined by EpCAM and AFP to reveal HCC subtypes	29	cDNA and Oligo microarray	[46]

(continued)

Table 6.1 (continued)

First author	Year	Primary endpoint	Number of genes	Experimental platform	Reference
Taro Yamashita	2009	EpCAM-positive HCC as a tumor-initiating cells with stem/progenitor cell features	793	cDNA and Oligo microarray	[47]
Derek Y. Chiang,	2008	Molecular classification of HCC using copy number alteration and gene expression data	~1000	Affymetrix HG-U133 plus 2	[49]
Hoshida	2009	Molecular subclasses of human HCC	619	cDNA & Oligo microarray	[54]
Hyun Goo Woo	2010	Identification of a cholangiocarcinoma-like gene expression trait in HCC	581	Affymetrix U133A 2.0 array	[52]
Xin-Rong Yang	2010	Investigate the prognostic values of putative hepatic stem/progenitor cell in HCC	14	Real-time qRTePCR analysis	[36]
Soomi Kim	2012	Prediction of overall survival in HCC	65	Illumina microarray platform	[41]
TCGA	2017	Comprehensive and integrative genomic characterization of HCC	528	Illumina Hiseq	[56]
TCGA	2017	Characterization of the clinical, pathological, and molecular features of nonproliferative HCCs	550	Various microarray, Illumina Hiseq	[58]
Yujin Hoshida	2008	Genomewide expression profiling of HCC correlated with survival outcome	186	Illumina DASL	[62]
Lindsay Y King	2015	Genomic and clinical prognostic index for hepatitis C-related early-stage cirrhosis	186	Illumina DASL	[65]
Ji-Hoon Kim	2014	Development of genomic predictor for identifying late recurrence and its clinical implications	233	Illumina HumanHT-12	[66]

to the complex pathogenesis of this cancer [20, 21]. HCC arises most commonly in cirrhotic livers following infection with HBV or HCV. However, HCC can also occur under variety of other conditions such as hemochromatosis, excessive alcohol consumption, and nonalcoholic steatohepatitis. Each of these conditions represents complex and different constellations of chromosomal aberrations and genetic and epigenetic alterations as well as changed molecular pathways. Nevertheless, global gene expression profiling, because of its extraordinary power of resolution, may currently be the most appropriate technology platform to molecularly resolve the complex pathogenesis of HCC. These datasets represent an impressive progress in the use of gene expression profiling in elucidating the molecular pathogenesis of

HCC and hold the promise of improving the diagnostic and prognostic prediction for HCC patients. The dataset is also large enough to warrant a critical examination of reproducibility and validation of the molecular classification of HCC and the predictive expression "signatures" (or markers) generated by the global gene expression profiling of HCC.

For the analysis of cancer genomic data, two general methods have been applied to uncover molecular subtypes significantly associated clinical outcomes [22]. Supervised approach intends to find a set of variables such as expressed genes or mutation frequency from tumors on the basis of which one can reliably predict clinical outcomes such as survival, recurrence, response to treatments, or any class of interest in patients. Unsupervised approach intends to find either a completely novel subset (or cluster) of patients that are not recognized previously or to uncover similarity among group of patients that were considered as clinically different ones. The goal is to find more details about underlying biology of tumors that are clinically different and identify robust biomarkers that can reliably classify patients for better management.

Prognostic Subtypes of Hepatocellular Carcinoma Redefined by Supervised Approaches

HCC recurrence is a serious complication following resection of the primary tumor and happens in 50% of cases 3 years after the operation [23]. In 75% of the cases, this is due to intrahepatic metastasis, whereas the remaining 25% are due to de novo HCC [24]. The major histopathological features that predict HCC recurrence are vascular invasion, degree of differentiation of the tumor, and multinodularity [23].

Several studies have employed supervised approach to gene expression profiling data to address the issue of HCC recurrence following resection and intrahepatic metastasis.

Iizuka et al. investigated mRNA expression data from 33 HCC tumors as training set with use of early version of oligonucleotide microarrays with only 6000 genes [25]. The training set was used in a supervised learning manner to construct a predictor with 12 genes. The predictive performance of the system was then compared on a blinded set of samples from 27 newly enrolled patients. The system correctly predicted early intrahepatic recurrence or nonrecurrence in 25 (93%) of 27 samples in the blinded set and had a positive predictive value of 88% and a negative predictive value of 95%. This study was the first one that demonstrated the potential of prognostic values of genomic data from HCC tumors.

In another study, Kurokawa et al. addressed the issue of intrahepatic recurrence by analyzing gene expression using a quantitative reverse transcriptase polymerase chain reaction (qRT-PCR)-based array platform of 3072 genes in 100 HCC patients [26]. The authors selected 92 genes that demonstrated distinct expression patterns differing significantly between recurrence and recurrence-free cases. Using the 20

top-ranked genes (from the 92 selected), the predictor correctly predicted the early intrahepatic recurrence for 29 of 40 cases within the validation group, with the odds ratio of 6.8 (95% CI 1.7–27.5, $p = 0.01$). The 2-year recurrence rates in the patients with the good signature and those with the poor signature were 29.4% and 73.9%, respectively. The authors further showed (using multivariate Cox analysis) that the 20-gene molecular signature was an independent indicator for recurrence (hazard ratio 3.82, 95% CI 1.44–10.10, $p = 0.007$).

Ye et al. analyzed the expression profiles of 67 primary and metastatic HCC samples from 40 patients [27]. Using a supervised machine learning algorithm, the authors generated a 153-gene-expression signature that permitted classification of metastatic HCC patients and patient survival. The authors further showed that the gene-expression signature of primary HCCs with accompanying metastasis was very similar to that of their corresponding metastases, implying that genes favoring metastasis progression were initiated in the primary tumors. Furthermore, osteopontin, which was identified as a lead gene in the signature, was overexpressed in metastatic HCC and an osteopontin-specific antibody effectively blocked HCC cell invasion in vitro and inhibited pulmonary metastasis of HCC cells in nude mice. This metastatic gene signature was further redefined and validated in follow-up study [28]. In multivariate analyses including various clinical risk factors and clinical staging, the metastasis signature was an independent prognostic predictor, especially applicable to early recurrence, and a poor prognostic factor mainly associated with metastatic dissemination of HCC cells but not late recurrence, an outcome contributed mainly by high carcinogenic activities of diseased livers.

Woo et al. applied similar approach to identify genes whose expression is significantly associated with early recurrence after curative-intent treatment [29]. Authors selected 628 genes as classifiers from gene expression data from 65 HBV-associated HCC tumors by using univariate Cox proportional hazard model. Prognostic significance of the recurrence signature was validated in independent cohort of HCC patients. Gene network analysis with the recurrence signature revealed that SP1 transcription complex might be prominent common regulators of genes that differed in expression between high risk and low risk of early recurrence.

Vascular invasion is significantly associated with recurrence and poor clinical outcome after curative treatment of HCC such as resection and liver transplantation [30]. Meta-analysis with 1500 HCC patients further supported that the presence of vascular invasion is a critical factor for selection of patients for curative treatment in addition to size and number of tumors [31]. Minguez et al. developed 35-gene-based predictor that can identify HCC patients with vascular invasion [32]. Interestingly, high expression of CD24, an adhesion receptor of activated endothelial cells and platelets, was significantly associated with vascular invasion [33]. In contrast, many of metabolic genes such as *GLYAT*, *UGT2B15*, *CYP3A4*, and *ADH4* were under-expressed in the tumors with vascular invasion.

Nault et al. carried out stepwise analysis of gene expression data from HCC to identify prognostic gene set. By analyzing gene expression data from previous studies, they first selected 103 candidate genes for further selection with qRT-PCR experiments. By applying univariate Cox analysis and a stepwise forward selection

and backward elimination approach to expression data of 103 genes from training datasets, they identified a panel of five genes (*TAF9*, *RAMP3*, *HN1*, *KRT19*, and *RAN*) showing the strongest prognostic relevance and constructed five-gene scores for validation in independent cohort of HCC patients [34]. The five genes reflected different signaling pathways deregulated in poor prognostic tumors. *KRT19* is related to the stem cell and progenitor feature and known to be associated with poor prognosis of HCC [9]. Authors suggested that the five-gene score could be used for better selection of patients for liver transplantation, for example, by extending the Milan criteria to good prognostic tumors even if tumor size is more than 5 cm [35]. Yang et al. used similar approach with genes related to hepatic stem and progenitor cells to identify prognostic genes [36]. In this study, the expression and clinical significance of putative hepatic stem cell genes and tumor angiogenesis-related genes were investigated by real-time RT-PCR first and later by immunohistochemistry in three independent cohorts of patients with HCC.

Prognostic Subtypes of Hepatocellular Carcinoma Uncovered by Unsupervised or Semi-supervised Approaches

By applying unsupervised analysis of global gene expression data from human HCC, Lee et al. identified two distinctive subclasses that are highly associated with the survival of the patients: subtype A and B represent tumors with a poor and better prognosis, respectively [8]. A limited number (1016 gene features representing 947 unique genes) of genes were identified that both predicted the length of survival of the HCC patients and provided new molecular insights into the pathogenesis of HCC. Because application of a knowledge-based annotation of the 947 genes revealed that cell proliferation is the best characteristic of subtype A, it was named National Cancer Institute Proliferation (NCIP) signature. Subtype A also displayed higher expression of genes involved in ubiquitination and histone modification. It is well established that the ubiquitin system is frequently deregulated in cancers [37] and has been proposed as a possible predictive marker for recurrence of human HCC [38, 39]. The predictive power of the NCIP signature was validated in independent HCC datasets [9, 40, 41]. Cross-species comparison of the signature also revealed mouse models best mimicking human subtypes A and B [42].

The hepatic stem (HS) cell subtype of HCC was defined as gene expression patterns resembling those found in fetal hepatic stem cells [9]. Interestingly, HS subtype is a subset of previously recognized poor prognostic subtype A of NCIP. Gene network analysis of HS signature revealed that AP1 transcription factors such as *FOS*, *FOSL2*, and *JUNB* are highly activated in HS subtype. Shared gene expression patterns of the HS subtype and hepatic stem cells suggest that this subtype of HCC may arise from adult hepatic progenitor cells. Further support for this idea is supplied by the finding that expression of well-known markers of hepatic oval cells, such as *KRT7*, *KRT19*, and *VIM*, is found in the HS subtype of HCC [43].

Hepatic stem cell-like subtype of HCC was also uncovered by independent study. Epithelial cell adhesion molecule (EpCAM) was predominantly expressed in hepatic progenitor cells or hepatic stem cells [44, 45]. In attempt to find subset of HCC with stem cell characteristics, Yamashita et al. identified 70 EpCAM-coexpressed genes in EpCAM-positive HCC for construction of prediction model [46]. Prognostic significance of EpCAM-positive HCC subtype was validated in large independent cohorts. Based on the EpCAM signature, which may be related to different liver cell lineages, authors proposed the four subtypes of HCC: EpCAM-positive and AFP-positive HCC as hepatic stem cell-like HCC, EpCAM-positive and AFP-negative as bile duct epithelium-like HCC, EpCAM-negative and AFP-positive HCC as hepatocytic progenitor-like HCC, and EpCAM-negative and AFP-negative HCC as mature hepatocyte-like HCC. Markers for hepatic progenitor cells such as KRT19 and KIT are more abundantly expressed in hepatic stem cell-like HCC, whereas mature hepatocyte-specific genes such as *CYP3A4* are more abundantly expressed in mature hepatocyte-like HCC. Later study demonstrated that EpCAM-positive HCC is highly invasive and EpCAM is account for invasiveness of these cancer cells [47].

Another unsupervised approach revealed six genomic subtypes of HCC (G1–G6) [48]. Each subtype showed characteristic genetic alterations. The tumors in G1–G3 subtypes were associated with high chromosomal instability compared to the tumors in G4–G6 subtypes. Among the frequently mutated genes, *CTNNB1* mutations were enriched in the G5–G6 subtypes, while mutations in *TP53* genes were significantly associated with the G2–G3 subtypes. *PIK3CA* mutations were associated with the G2 subgroup. Hypermethylation on promoters of *CDH1* and *CDKN2A* were most frequently observed in the G5–G6 and G3 subtypes, respectively.

Integrative analysis of genomic copy number alteration with mRNA expression data from HCC tumors uncovered five genomic subtypes [49]. A subtype is a unique subclass of HCC characterized by polysomy of chromosome 7 and the concomitant overexpression of many genes in this chromosome. Intriguingly, these tumors lack gains of chromosome 8q, which are the second most frequent chromosomal alterations in hepatocellular carcinomas and include the known oncogenes *MYC*, *PTK2*, and *COPS5* [50, 51]. *CTNNB1*-activated subtype was enriched for gain-of-function mutations in *CTNNB1* (mostly located in exon 3). Interferon (IFN)-related subtype overexpressing several IFN-stimulated genes was associated with smaller tumor size.

Woo et al. carried out semi-supervised analysis with pooled gene expression data from HCC and cholangiocarcinoma. They discovered that the subset of HCC tumors is highly similar to cholangiocarcinoma and named them cholangiocarcinoma-like HCC (CLHCC) [52]. Tumors in CLHCC subtype are characterized by high expression of markers for hepatic progenitor cells such as *KRT19*, *EPCAM*, and *PROM1*. As expected, CLHCC subtype is significantly associated with poor prognosis, and it was validated in multiple independent cohorts of HCC patients. The CLHCC tumors are enriched with the proliferation, metastasis/adhesion, and development-related functions reflecting their aggressive phenotype. Biological

characteristics of CLHCC signature are also well overlapped with multiple embryonic stem cell signatures as well as hepatic stem cell signature [9, 53].

Meta-analysis of gene expression data from eight independent patient cohorts uncovered three HCC subtypes (S1, S2, and S3), each correlated with clinical parameters such as tumor size, extent of cellular differentiation, and serum AFP levels [54]. Of the three subtypes, S1 and S2 subtypes are associated with poor prognosis of HCC patients and S3 subtype is characterized by less aggressive features, including preserved hepatocyte function, smaller and more differentiated tumor, and better prognosis. S1 subtype is characterized by activation of TGF-β pathway and CLHCC gene signature [9, 52, 55], while S2 subtype is characterized by stem cell markers such as *EPCAM*, *AFP*, and *GPC3*, activation of *IGF2* pathway, and relative suppression of interferon target genes and hepatoblastoma-like gene signature [46, 47]. A vascular invasion gene signature [32] is more strongly associated with the S2 subtype. Interestingly, a subset of the S3 subtype HCC is characterized by gain-of-function mutations in *CTNNB1*. S2 subtype is further characterized by proliferation as well as *MYC* and *AKT* activation, and S3 was associated with hepatocyte differentiation.

Kim et al. carried out meta-analysis with two prognostic gene expression signatures to find limited number of genes whose expression is significantly associated with the prognosis of HCC patients [41]. Of 1016 NCIP genes and 628 recurrence genes from previous studies [8, 29], only 65 genes were shared by both gene lists. For easier translation of prognostic genome signatures to clinics, author generated recurrence-risk scores based on expression of 65 genes. The risk score was developed using Cox coefficient values of 65 genes in the training set, and its robustness was validated in test sets. The risk score was a highly significant predictor of overall survival and recurrence-free survival. In multivariate analysis, the risk score was a significant risk factor among clinical variables examined together. Interestingly, authors found that a high risk score was significantly associated with activation of *AKT* and *IGF1R*, whereas a high frequency of mutations of *CTNNB1* was significantly associated with a low risk score.

In recent analysis of HCC genome data from The Cancer Genome Atlas (TCGA) project, investigators found that HCC with *IDH1/2* mutations has very unique gene expression [56]. Interestingly, many HCC tumors without IDH mutations have IDH signature, and those with IDH signature (IDH-like subtype) showed significantly poor survival after treatment. When compared with other molecular subtypes of HCC, the IDH-like HCC exhibited the highest similarity to an HS [9]. These samples exhibited similarity to Hoshida's S2 subtype [54] and CLHCC subtype [52] and had high risk scores based on a gene expression of 65 genes [41].

By applying iCluster approach that integrates all of genomic data including somatic mutation, copy number alteration, mRNA expression, miRNA expression, and DNA methylation data [57], TCGA investigator uncovered three genomic subtypes: iC1, iC2, and iC3 [56]. iC1 subtype is characterized by clinical associations with younger age, Asian ethnicity, and female gender. These tumors exhibited features such as higher tumor grade and presence of macrovascular invasion. Molecular correlations with iC1 included a low frequency of *CDKN2A* silencing,

CTNNB1 mutation, and *TERT* promoter mutations accompanied with low *TERT* expression. This subclass was associated with overexpression of proliferation marker genes such as *MYBL2*, *PLK1*, and *MKI67*. iC2 and iC3 subtypes exhibited a high frequency of *CDKN2A* silencing, *TERT* promoter mutations, *CTNNB1* mutations, and *HNF1A* mutation. Correlation with clinical variables reveals association of iC2 subtype with low-grade tumors and less microvascular invasion. iC3 subtype is characterized by a higher degree of chromosomal instability with distinct 17p loss, high frequency of *TP53* mutation, and hypomethylation of multiple CpG sites. When compared with Hoshida's 3 genomic subtypes, iC1 is highly similar to Hoshida S2 subtype whereas iC3 is highly similar to Hoshida S3 subtype.

Recent meta-analysis with pooled HCC gene expression data revealed four subtypes of HCC that are well associated with liver zonation program: periportal (PP) subtype, perivenous (PV) subtype, extracellular matrix (ECM) subtype, and stem cell (STEM) subtype [58]. PV subtype is enriched for somatic mutations in *CTNNB1* and expresses many genes involved in liver zonation such as *GLUL*, *HAL*, and *VNN1*. Likewise, PP-type HCCs expressed a host of amino acid-degrading enzymes, such as *ARG1* and *GLS2*, which were major hubs in the periportal gene network in liver. STEM subtype is highly related to previously recognized HS subtype [9].

Nontumor Genomic Signatures

It has long been recognized that survival prediction of HCC patients is more challenging than with most other cancers. This is, in case of HBV and HCV, the consequence of the underlying viral-driven nonneoplastic disease, i.e., chronic hepatitis and cirrhosis that can inflict functional impairment on the liver that may affect the outcome of the HCC patients [59]. In HCC, two distinct types of recurrence are known. Early recurrence arises from primary cancer cells disseminating to the surrounding liver and is usually observed within the first 2 years after surgery. In contrast, late recurrence, which is typically observed more than 2 year after surgery, appears to be a result of chronic liver damage known as the "field effect" and produces de novo tumors that are independent of resected primary tumors [60]. The two types of recurrence have different clinical courses and probably appear in distinct biological contexts [61]. For better disease management, it is therefore important to uncover the biological characteristics of each type of recurrence and to develop distinct molecular prognostication systems that can identify patients at high risk for either type.

Hoshida et al. characterized gene expression data from nontumor surrounding tissues from HCC patients to uncover critical genes that might reflect field effect in liver leading to HCC development later [62]. By applying leave-one-out cross-validation procedure, authors identified 186 genes whose expression is significantly associated with survival of HCC patients. Prognostic significance of the signature was validated in large independent cohort of HCC. In particular, while the signature is not associated with early recurrence after surgery, it was significantly associated

with late recurrence. Genes upregulated in poor prognostic subtype include those related to interferon signaling, activation of NFκB, and TNFα signaling pathway. Interestingly, the downstream targets of IL6 were strongly associated with the signature, which is consistent with the finding that IL6 plays key roles in protecting mice from chemically induced HCC development [63]. The 186-gene signature was further validated in more relevant clinical setting. It was significantly correlated with long-term outcomes including HCC development of patients with hepatitis C-related early-stage cirrhosis [64, 65]. Therefore, the signature might be used to identify patients with cirrhosis in most need of surveillance and strategies to prevent the development of HCC.

Kim et al. identified gene set whose expression is significantly associated with hepatic injury and regeneration (HIR) in human liver [66]. When applied to gene expression data from nontumor surrounding tissues of HCC patients, HIR signature was significantly associated with late recurrence. In contrast, tumor-derived 65-gene recurrence score [41] was only associated with early recurrence. Gene network analysis revealed that STAT3 might be key upstream regulator of HIR signature. Activation of STAT3 in HCC patients with high risk for late recurrence was validated by immunostaining of surrounding liver tissues. The outcomes of analysis strongly suggested that early and late recurrences are clinically different entities with distinctive biological characteristics. Thus, separate rational treatment recommendations should be developed for better management of HCC patients. For example, patients at high risk of late recurrence may benefit from the use of JAK/STAT pathway inhibitors after surgical resection. Because current staging systems and biomarkers are limited in their ability to assess patients' risk of recurrence and their potential benefit from adjuvant therapy, two genomic predictors specific for early and late recurrence may represent tools that could help refine treatment decisions based on molecular characteristics.

Conclusion and Perspective

Comprehensive molecular and genomic analyses of large cohorts of HCC have now uncovered clinically relevant genomic subtypes, characteristic genomic alterations associated with subtypes, and genomic predictors of these subtypes. The results from analysis of genomic data have started to impact both clinical decision-making in oncology and advanced our understanding of cancer biology, as well as facilitated the development of more effective therapies.

While most of these findings are very encouraging, there are substantial gaps in translating genomic subtypes to clinics. While many of discovered genomic subtypes are clinically relevant, its clinical utility is hampered by discrepant results, which are probably due to difference in technological platforms, patient population, preparation and processing of samples, and classification algorithms. However, some of genomic subtypes were repeatedly discovered by independent studies. For example, HS subtype from NCI study is subset of poor prognostic subtype of NCIP

classification and highly similar to EpCAM-positive subtype, CLHCC subtype, and IDH-like HCC [56]. PV subtype from meta-analysis study is subset of Hoshida's S3 subtype and highly similar to CTNNB1 subtype from Barcelona group's study [49]. Albeit the similarity among independent classifications still remains superficial level, these similarities clearly suggest that it is possible to find consensus among different genomic classification methods that are clinically significant and biologically meaningful.

Key limitation of genomic subtype in HCC is that they do not provide clinically actionable information that is essential for personalized treatment of patients. Although sorafenib, a multi-kinase inhibitor [67], is approved for first-line treatment of advanced HCC patients more than 10 years ago [68, 69], there are no studies demonstrating association of genomic subtypes with treatment response to sorafenib yet. Likewise, many of targeted drugs approved for treatment of HCC patients lacks biomarkers reflecting genomic subtypes. Therefore, it will be important to collect tumors in prospective clinical trials to connect genomic subtypes and response to treatment.

Finally, another limitation is lack of actionable targets in subtypes. Many known drivers of HCC such as *CTNNB1*, *TERT*, *MYC*, and *YAP1* have been considered to be undruggable targets. Furthermore, key drivers or therapeutic targets are not fully discovered yet in some genomic subtypes. Therefore, it is important to establish preclinical models that faithfully recapitulate pathogenesis of subtypes. Animal models that recapitulate human's physiology and clinical setting have been crucial for understanding hepatocarcinogenesis and improving the treatment of HCC. The perfect animal model should reproduce natural history, etiology, and pathology of human HCC that would allow not only to uncover molecular mechanisms of HCC development over time but also to examine and evaluate potential novel therapeutic approaches in preclinical setting.

References

1. Sander C. Genomic medicine and the future of health care. Science. 2000;287(5460):1977–8.
2. Stenson PD, Mort M, Ball EV, Evans K, Hayden M, Heywood S, et al. The Human Gene Mutation Database: towards a comprehensive repository of inherited. Hum Genet. 2017;136(6):665–77.
3. Guttmacher AE, Collins FS. Realizing the promise of genomics in biomedical research. JAMA. 2005;294(11):1399–402.
4. Tian Q, Price ND, Hood L. Systems cancer medicine: towards realization of predictive, preventive. J Intern Med. 2012;271(2):111–21.
5. Dienstmann R, Rodon J, Barretina J, Tabernero J. Genomic medicine frontier in human solid tumors: prospects and challenges. J Clin Oncol. 2013;31(15):1874–84.
6. Macconaill LE, Garraway LA. Clinical implications of the cancer genome. J Clin Oncol. 2010;28(35):5219–28.
7. Mendelsohn J. Personalizing oncology: perspectives and prospects. J Clin Oncol. 2013;31(15):1904–11.

8. Lee JS, Chu IS, Heo J, Calvisi DF, Sun Z, Roskams T, et al. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. Hepatology. 2004;40(3):667–76. https://doi.org/10.1002/hep.20375.
9. Lee JS, Heo J, Libbrecht L, Chu IS, Kaposi-Novak P, Calvisi DF, et al. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. Nat Med. 2006;12(4):410–6. https://doi.org/10.1038/nm1377.
10. Calvet X, Bruix J, Gines P, Bru C, Sole M, Vilana R, Rodes J, et al. Prognostic factors of hepatocellular carcinoma in the west: a multivariate. Hepatology. 1990;12(4 Pt 1):753–60.
11. Chevret S, Trinchet JC, Mathieu D, Rached AA, Beaugrand M, Chastang C. A new prognostic classification for predicting survival in patients with. J Hepatol. 1999;31(1):133–41.
12. Investigator C. A new prognostic system for hepatocellular carcinoma: a retrospective study of. Hepatology. 1998;28(3):751–5.
13. Okuda K, Ohtsuki T, Obata H, Tomimatsu M, Tomimatsu M, Okazaki N, Hasegawa H, Nakajima Y, et al. Natural history of hepatocellular carcinoma and prognosis in relation to. Cancer. 1985;56(4):918–28.
14. Tan CK, Law NM, Ng HS, Machin D. Simple clinical prognostic model for hepatocellular carcinoma in developing. J Clin Oncol. 2003;21(12):2294–8.
15. Marrero JA, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM, et al. Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the American Association for the Study of Liver Diseases. Hepatology. 2018;68(2):723–50. https://doi.org/10.1002/hep.29913.
16. Kim CK, Lim J, Lee WJ, Lee WJ. Detection of hepatocellular carcinomas and dysplastic nodules in cirrhotic liver. J Ultrasound Med. 2001;20(2):99–104.
17. Lencioni R, Cioni D, Bartolozzi C. Tissue harmonic and contrast-specific imaging: back to gray scale in ultrasound. Eur Radiol. 2002;12(1):151–65.
18. Kojiro M, Roskams T. Early hepatocellular carcinoma and dysplastic nodules. Semin Liver Dis. 2005;25(2):133–42.
19. Leykum LK, El-Serag HB, Cornell J, Papadopoulos KP. Screening for hepatocellular carcinoma among veterans with hepatitis C on disease. Clin Gastroenterol Hepatol. 2007;5(4):508–12.
20. Thorgeirsson SS, Grisham JW. Molecular pathogenesis of human hepatocellular carcinoma. Nat Genet. 2002;31(4):339–46. https://doi.org/10.1038/ng0802-339.
21. Marquardt JU, Andersen JB, Thorgeirsson SS. Functional and genetic deconstruction of the cellular origin in liver cancer. Nat Rev Cancer. 2015;15(11):653–67 LID. https://doi.org/10.1038/nrc4017.
22. Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. Nat Rev Genet. 2015;16(6):321–32. https://doi.org/10.1038/nrg3920.
23. Forner A, Llovet JM, Bruix J. Hepatocellular carcinoma. Lancet. 2012;379(9822):1245–55.
24. Cheung ST, Chen X, Guan XY, Wong SY, Tai LS, Ng IO, So S, et al. Identify metastasis-associated genes in hepatocellular carcinoma through. Cancer Res. 2002;62(16):4711–21.
25. Iizuka N, Oka M, Yamada-Okabe H, Nishida M, Maeda Y, Mori N, Takao T, et al. Oligonucleotide microarray for prediction of early intrahepatic recurrence of. Lancet. 2003;361(9361):923–9.
26. Kurokawa Y, Matoba R, Takemasa I, Nagano H, Dono K, Nakamori S, Umeshita K, et al. Molecular-based prediction of early recurrence in hepatocellular carcinoma. J Hepatol. 2004;41(2):284–91.
27. Ye QH, Qin LX, Forgues M, He P, Kim JW, Peng AC, Simon R, et al. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using. Nat Med. 2003;9(4):416–23.
28. Roessler S, Jia HL, Budhu A, Forgues M, Ye Q-H, Lee J-S, Thorgeirsson SS, et al. A unique metastasis gene signature enables prediction of tumor relapse in. Cancer Res. 2010;70(24):10202–12.
29. Woo HG, Park ES, Cheon JH, Kim JH, Lee J-S, Park BJ, Kim W, et al. Gene expression-based recurrence prediction of hepatitis B virus-related human. Clin Cancer Res. 2008;14(7):2056–64.
30. Llovet JM, Di Bisceglie AM, Bruix J, Kramer BS, Lencioni R, Zhu AX, Sherman M, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. J Natl Cancer Inst. 2008;100(10):698–711.

31. Mazzaferro V, Llovet JM, Miceli R, Bhoori S, Schiavo M, Mariani L, Camerini T, et al. Predicting survival after liver transplantation in patients with hepatocellular. Lancet Oncol. 2009;10(1):35–43.
32. Minguez B, Hoshida Y, Villanueva A, Toffanin S, Cabellos L, Thung S, Mandeli J, et al. Gene-expression signature of vascular invasion in hepatocellular carcinoma. J Hepatol. 2011;55(6):1325–31.
33. Sagiv E, Memeo L, Karin A, Kazanov D, Jacob-Hirsch J, Mansukhani M, Rechavi G, et al. CD24 is a new oncogene, early at the multistep process of colorectal cancer. Gastroenterology. 2006;131(2):630–9.
34. Nault JC, De Reynies A, Villanueva A, Calderaro J, Rebouissou S, Couchy G, Decaens T, et al. A hepatocellular carcinoma 5-gene score associated with survival of patients. Gastroenterology. 2013;145(1):176–87.
35. Clavien PA, Lesurtel M, Bossuyt PMM, Gores GJ, Langer B, Perrier A. Recommendations for liver transplantation for hepatocellular carcinoma: an. Lancet Oncol. 2012;13(1):e11–22 LID. https://doi.org/10.1016/S470-2045(11)70175-9.
36. Yang XR, Xu Y, Yu B, Zhou J, Qiu S-J, Shi G-M, Zhang B-H, et al. High expression levels of putative hepatic stem/progenitor cell biomarkers. Gut. 2010;59(7):953–62.
37. Pagano M, Benmaamar R. When protein destruction runs amok, malignancy is on the loose. Cancer Cell. 2003;4(4):251–6.
38. Shirahashi H, Sakaida I, Terai S, Hironaka K, Kusano N, Okita K. Ubiquitin is a possible new predictive marker for the recurrence of human. Liver. 2002;22(5):413–8.
39. Tu K, Yang W, Li C, Zheng X, Lu Z, Guo C, Yao Y, et al. Fbxw7 is an independent prognostic marker and induces apoptosis and growth arrest. Mol Cancer. 2014;13:110.
40. Kim JH, Sohn BH, Lee HS, Kim SB, Yoo JE, Park YY, et al. Genomic predictors for recurrence patterns of hepatocellular carcinoma: model derivation and validation. PLoS Med. 2014;11(12):e1001770. https://doi.org/10.1371/journal.pmed.1001770.
41. Kim SM, Leem SH, Chu IS, Park YY, Kim SC, Kim SB, et al. Sixty-five gene-based risk score classifier predicts overall survival in hepatocellular carcinoma. Hepatology. 2012;55(5):1443–52. https://doi.org/10.1002/hep.24813.
42. Lee JS, Chu IS, Mikaelyan A, Calvisi DF, Heo J, Reddy JK, et al. Application of comparative functional genomics to identify best-fit mouse models to study human cancer. Nat Genet. 2004;36(12):1306–11. https://doi.org/10.1038/ng1481.
43. Roskams TA, Libbrecht L, Desmet VJ, Desmet VJ. Progenitor cells in diseased human liver. Semin Liver Dis. 2003;23(4):385–96.
44. Schmelzer E, Wauthier E, Reid LM, Reid LM. The phenotypes of pluripotent human hepatic progenitors. Stem Cells. 2006;24(8):1852–8.
45. Schmelzer E, Zhang L, Bruce A, Wauthier E, Ludlow J, Yao H, Moss N, et al. Human hepatic stem cells from fetal and postnatal donors. J Exp Med. 2007;204(8):1973–87.
46. Yamashita T, Forgues M, Wang W, Kim JW, Ye Q, Jia H, Budhu A, et al. EpCAM and alpha-fetoprotein expression defines novel prognostic subtypes of. Cancer Res. 2008;68(5):1451–61.
47. Yamashita T, Ji J, Budhu A, Forgues M, Yang W, Wang HY, et al. EpCAM-positive hepatocellular carcinoma cells are tumor-initiating cells with stem/progenitor cell features. Gastroenterology. 2009;136(3):1012–24. https://doi.org/10.1053/j.gastro.2008.12.004.
48. Boyault S, Rickman DS, de Reynies A, Balabaud C, Rebouissou S, Jeannot E, Herault A, et al. Transcriptome classification of HCC is related to gene alterations and to new. Hepatology. 2007;45(1):42–52.
49. Chiang DY, Villanueva A, Hoshida Y, Peix J, Newell P, Minguez B, LeBlanc AC, et al. Focal gains of VEGFA and molecular classification of hepatocellular carcinoma. Cancer Res. 2008;68(16):6779–88.
50. Patil MA, Gutgemann I, Zhang J, Ho C, Cheung ST, Ginzinger D, Li R, et al. Array-based comparative genomic hybridization reveals recurrent chromosomal. Carcinogenesis. 2005;26(12):2050–7.
51. Okamoto H, Yasui K, Zhao C, Arii S, Inazawa J. PTK2 and EIF3S3 genes may be amplification targets at 8q23-q24 and are associated. Hepatology. 2003;38(5):1242–9.

52. Woo HG, Lee JH, Yoon JH, Kim CY, Lee HS, Jang JJ, Yi N-J, et al. Identification of a cholangiocarcinoma-like gene expression trait in. Cancer Res. 2010;70(8):3034–41.
53. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, Weinberg RA, et al. An embryonic stem cell-like gene expression signature in poorly differentiated. Nat Genet. 2008;40(5):499–507.
54. Hoshida Y, Nijman SM, Kobayashi M, Chan JA, Brunet JP, Chiang DY, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. Cancer Res. 2009;69(18):7385–92. https://doi.org/10.1158/0008-5472.CAN-09-1089.
55. Coulouarn C, Factor VM, Thorgeirsson SS. Transforming growth factor-beta gene expression signature in mouse hepatocytes. Hepatology. 2008;47(6):2059–67.
56. The Cancer Genome Atlas Research Network. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. Cell. 2017;169(7):1327–41. https://doi.org/10.1016/j.cell.2017.05.046.
57. Shen R, Olshen AB, Ladanyi M. Integrative clustering of multiple genomic data types using a joint latent. Bioinformatics. 2009;25(22):2906–12.
58. Desert R, Rohart F, Canal F, Sicard MA-O, Desille M, Renaud S, et al. Human hepatocellular carcinomas with a periportal phenotype have the lowest. Hepatology. 2017;66(5):1502–18.
59. Sangiovanni A, Del Ninno E, Fasani P, De Fazio C, Ronchi G, Romeo R, Morabito A, et al. Increased survival of cirrhotic patients with a hepatocellular carcinoma detected. Gastroenterology. 2004;126(4):1005–14.
60. Imamura H, Matsuyama Y, Tanaka E, Ohkubo T, Hasegawa K, Miyagawa S, Sugawara Y, et al. Risk factors contributing to early and late phase intrahepatic recurrence of. J Hepatol. 2003;38(2):200–7.
61. Poon RT, Fan ST, Ng IO, Lo CM, Liu CL, Wong J. Different risk factors and prognosis for early and late intrahepatic recurrence. Cancer. 2000;89(3):500–7.
62. Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, Gupta S, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. N Engl J Med. 2008;359(19):1995–2004.
63. Naugler WE, Sakurai T, Kim S, Maeda S, Kim K, Elsharkawy AM, Karin M, et al. Gender disparity in liver cancer due to sex differences in MyD88-dependent IL-6. Science. 2007;317(5834):121–4.
64. Hoshida Y, Villanueva A, Sangiovanni A, Sole M, Hur C, Andersson K, Chung RT, et al. Prognostic gene expression signature for patients with hepatitis C-related. Gastroenterology. 2013;144(5):1024–30.
65. King LY, Canasto-Chibuque C, Johnson KB, Yip S, Chen X, Kojima K, et al. A genomic and clinical prognostic index for hepatitis C-related early-stage. Gut. 2015;64(8):1296–302.
66. Kim JH, Sohn BH, Lee HS, Kim SB, Yoo JE, Park YY, et al. Genomic predictors for recurrence patterns of hepatocellular carcinoma: model. PLoS Med. 2014;11(12):e1001770.
67. Wilhelm S, Carter C, Lynch M, Lowinger T, Dumas J, Smith RA, Schwartz B, et al. Discovery and development of sorafenib: a multikinase inhibitor for treating. Nat Rev Drug Discov. 2006;5(10):835–44.
68. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med. 2008;359(4):378–90. https://doi.org/10.1056/NEJMoa0708857.
69. Cheng AL, Kang YK, Chen Z, Tsao CJ, Qin S, Kim JS, Luo R, et al. Efficacy and safety of sorafenib in patients in the Asia-Pacific region with. Lancet Oncol. 2009;10(1):25–34.