Molecular Genetics and Genomics of Hepatocellular Carcinoma

Dilek Colak¹ and Namik Kaya²

¹Department of Biostatistics, Epidemiology and Scientific Computing, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia
²Department of Genetics, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia

1. Introduction

Hepatocellular carcinoma (HCC) is a major type of liver cancer and third leading cause of cancer-related deaths worldwide. It is often diagnosed at an advanced stage, and hence typically has a poor prognosis. The advances in high-throughput “omics” technologies (genomics, transcriptomics, proteomics) parallel to the availability of high-density microarrays and next-generation sequencing technologies feature important advances in understanding of the complex biological processes underlying tumorigenesis and metastasis of HCC, and uncovering promising biomarkers with clinical potential. Ultimately, the trend will be toward a personalized medicine that will improve diagnosis, treatment and prevention of primary liver cancer. In this chapter, we present an overview of most up-to-date developments regarding these approaches toward an understanding of molecular mechanisms of HCC and for the development of novel biomarkers and cancer therapeutics targets.

2. Background

Hepatocellular carcinoma (HCC) is the most common primary cancer originating in the liver, the fifth most common cancer type, and is the third leading cause of cancer mortality worldwide [1-2]. It is often diagnosed at an advanced stage, leading to poor prognosis. Recent reports show that HCC is becoming more widespread and has dramatically increased in North America, Western Europe and Japan [2-4]. Early detection of HCC, especially detection of early/small HCC, followed by the appropriate treatment would significantly alter the prognosis and reduce the number of tumor-related deaths. Though inspiring progress has been made in understanding the molecular mechanisms of HCC, there is still lack of complete understanding of the disease perhaps due to complexities associated with the HCC such as intricacy of liver transcriptome, viral infection, liver regeneration, and other confounding factors, that have been major limitations for developing useful biomarkers for the detection and early diagnosis as well as identification of novel therapeutic targets for HCC.
The advances in high-throughput “omics” technologies such as genomics, transcriptomics, proteomics, and metabolomics combined with the availability of high-density microarrays and low-cost high-throughput parallel sequencing technologies and their analyses using different bioinformatics tools and algorithms are providing unprecedented biological insights related to HCC. Global molecular profiling studies of HCC are providing a comprehensive view of genomic aberrations and expression changes that occur during the carcinogenic process. Hence, the knowledge gained from continuing research efforts on HCC undoubtedly facilitates the understanding of molecular mechanism of HCC pathogenesis, and to provide the best therapy for each cancer patient and to improve patient management. This approach will create a foundation for personalized therapeutics and treatments and expectantly will be available in the near future alongside the unprecedented advancement of next-generation sequencing technologies. These technologies already began to identify novel genes that may have a driver force for HCC pathobiology [5]. Identification of such driver genes within each tumor will highly likely be a source for the development of novel therapeutic targets for the malignancies for each HCC-affected individual.

Our aim in this chapter is to focus on the current advances in the genomics field of HCC as well as recent progress using next-generation deep sequencing technologies, and the current shift towards integrative approaches using data from these advanced technologies that will help better understanding of HCC and for the development of novel biomarkers and cancer therapeutics targets.

3. Genomic alterations in liver cancer

Current advances in microarray technologies has resulted in high-dimensional genomic data sets and profoundly improved our understanding of genomic imbalances in the context of its role in carcinogenesis; first, with the introduction of copy number variation (CNV) concept in addition to single nucleotide polymorphisms (SNP), and second, with the improved mapping of such CNVs throughout the whole genome of patients versus normal individuals. While very early observations have identified CNVs as “chromosomal polymorphisms” that are several megabases in size, the lower end of the size range of CNVs continues to drop that is consistent with the pace of technological advancement [6]. With their inclusion of coding genes, it is hardly surprising that CNVs play a role in human health and disease, although their role is only recently being recognized, first in the context of Mendelian disorders and more recently in complex diseases [7-8]. The presence of these polymorphisms, either at small (SNPs or mutations) or large (CNVs and CNAs) scale as well as regions comprising loss of heterozygosity (LOH) blocks are, therefore, likely to contribute to cancer formation [9-11]. The genomic modifications in a tumor represents a structural fingerprint that may include the transcriptional control mechanisms and locally impact gene expression levels [10, 12].

Initial array platforms utilizing either spotted clones inserted in bacterial artificial chromosomes (BACs) or in situ synthesized oligos on chip surfaces have been applied to HCC samples to better understand the role of genomic aberrations at the DNA level. These microarray-based assays are called array comparative genomic hybridization (aCGH)
technique since they are a modified version of the comparative genomic hybridization (CGH) approach applied to microarrays. Numerous studies have investigated chromosomal alterations associated with HCC using both CGH and aCGH techniques (as reviewed by Moinzadeh et al. [9]). Moreover, two leading microarray companies have developed similar assays containing only SNP probes. This approach was initiated by Affymetrix Inc. then later applied by Illumina Inc. Both companies have come up with different SNP assays comprising different numbers of unique SNP probe sets. While the aCGH approach provides much higher resolution over standard microscope-based banding techniques in terms of cytogenetics analysis, SNP arrays bring two main advantages over the other techniques including aCGH: LOH and uniparental disomy detection, and more diverse applications, such as utilization in association studies based on both SNP as well as CNV calls.

Later, higher density arrays having hundreds of thousands or even currently more than a million unique probe sets targeting CNVs and SNPs have been employed in HCC research and identified critical regions of the genome likely to be involved in molecular carcinogenesis of HCC. Such critical regions commonly exhibit either deletion or increased gene dosage, leading to changes in DNA copy number variations/polymorphisms (CNVs/CNPs), aberrations/abnormalities (CNAs) or contain LOH blocks in various cancers, including HCC [9, 13-15].

It is plausible that such HCC-specific CNVs and LOH blocks spanning from several kilobases to megabases comprise critical driver genes that may play a leading role in hepatocarcinogenesis and contain the genetic factors involved in HCC [14, 16-17]. In one of those early studies, Luo et al. utilized an integrated approach of DNA and RNA level analyses for HCC, and investigated overlapping genome-wide transcriptomic and genomic alterations among hepatocellular carcinomas (HCC), hepatoblastomas (HPBL), tissue adjacent to HCC and normal liver tissue derived from normal livers and hepatic resections [14]. In their study, genomic imbalances between 27 HCC samples and matching normal controls were determined using low density oligonucleotide arrays. The results indicated that several regions on chromosome 7, 8, 10 and 12 harbor numerous genomic aberrations. Further investigations revealed that many of these changes do not cause remarkable gene expression alterations. However, among other genes, two genes, GPC3 and TIEG, were found to have significant correlation between their copy numbers and expression changes. Further investigations of these two genes in a larger cohort (484 hepatic tissue and normal samples) confirmed the expression differences in HCC samples. Additional studies investigating the role of GPC3 expression in poorer clinical outcome revealed this gene may have a possible role on HCC aggressiveness and therefore may predict the HCC outcome [14].

In a more recent study, Chen et al. employed Affymetrix’s 500K SNP arrays with an average of 6 kb distance between its unique SNP probes to examine 13 different HCC cell lines in addition to some other cancer cell lines as well as 45 archived primary HCCs [18]. Numerous common and novel aberrations were observed in multiple cancer lines confirming previously known HCC-related cytogenetic regions detected by low-resolution methods and refining their breakpoints and boundaries, and also introducing
previously unknown critical genomic regions associated with HCC. Among 653 amplicons and 57 homozygous deletions (HDs) detected by the arrays using different cell lines, 126 amplicons and 6 HDs were selected and tested to identify novel HCC-related genes. Further analysis of such aberrant regions yielded two genes, FNDC3BB and SLC29A2, consistently up-regulated in multiple HCC data sets. Knock-down studies using short hairpin RNAs targeting both genes showed decreased cell proliferation, tumor formation, and anchorage-independent growth in xenograft models in nude mice confirmed a possible pivotal role of these genes in growth and tumor formation in subsets of HCC samples. Up-regulation of either gene is proposed to be activated through STAT3 signaling pathway which is a well-known phenomenon in HCC progression usually triggered by cytokines such as interleukin-6 [19-22].

In another study, Clifford et al. used Affymetrix SNP 6.0 assay comprising probes for detection of CNVs and SNPs, each has more than 900,000 unique oligos, totaling nearly 1.9 million probe sets [23]. In their study, a large number of samples exceeding 1100 cases including histopathologically confirmed HCC and liver cirrhosis (LC) samples as well as normal controls with Korean and Chinese ethnicity were analyzed in two stages; each having different subsets of patients and controls. Based on their analysis, two SNPs were found to diverge significantly between HCC versus LC group and therefore considered a likely factor influencing transitional events from cirrhosis to hepatocellular carcinoma. Interestingly, the first SNP, rs2551677, is not within close proximity of any known gene, the closest gene DDX18 being 175 kb upstream of the SNP. The second SNP, rs2880301, is positioned on intron 1 of TPTE2 encoding a homolog of PTEN tumor suppressor gene and is the first time reported to be associated with carcinogenesis [23]. Additionally, three SNPs (rs9267673, rs2647073, and rs3997872) were found to be strongly associated with HCC only and were not presenting any additive/multiplicative effect. The first SNP, rs9267673, is in close proximity of C2 gene unlike the other two SNPs, rs2647073 and rs3997872, associated with SNPs falling into linkage disequilibrium of two different HLA group genes: The rs2647073 with HLA-DRB1, HLA-DRB6, HLA-DRB5, and HLA-DRA whereas the rs3997872 with HLA-DQA1, HLA-DQB1, HLA-DQA2, and HLA-DQB2 loci. The associations were independently confirmed using TAGMAN assays indicating the validity of the SNP study. When they analyzed probes targeting copy number polymorphisms, eight CNV loci including six germline CNVs were identified to be significantly associated with liver carcinogenesis. One of the germline CNVs showing a high level of association with HCC is located on a small region on p arm of chromosome 1 where no gene is known. Five other CNVs found to be linked to HCC involving KNG1, C4orf29, LARP2, ALDH7A1, PHAX, C5orf48, LMNB1, SRPK2, PUS7, and TMPO genes. Among these CNVs, two involving TRG@ and TRA@ had the strongest association to HCC. Moreover, a functional pathway and network analyses carried out using 1000 most significant SNPs associated with HCC. Among the critical pathways "antigen processing and presentation" is found the most significantly overrepresented pathway with p-value of 1x10^{-11} indicating the strongest association to HCC. Overall, these observations indicate involvement of immune system in constitutional susceptibility to HCC and HCC carcinogenesis which was suggested by clinical observations and animals models previously.

In a recent study, Jia et al. searched for critical somatic CNVs in 58 HCC tumor samples with adjacent non-tumor samples using Affymetrix 6.0 assay and identified 1241 regions [24].
These regions were then interrogated in search of dysregulated genes and 362 differentially expressed genes were identified. Among these, 20 genes were further evaluated functionally and TRIM35, HEY1, and SNRPE were confirmed to be involved in HCC by various functional experiments. Involvement of these genes, TRIM35 as tumor suppressor, and HEY1 and SNRPE as potential oncogenes, in HCC is novel.

4. Global gene expression profiling of liver cancer

During the past two decades, discoveries on the global gene expression profiling technologies emerged one after the other. Subtractive hybridization, differential display, SAGE, microarrays and more recently next generation sequencing techniques appeared as cutting-edge tools to study genome-wide transcriptional profiling differences in nearly all different types of tissues. With the rapid advances in these technologies, the medicine, particularly cancer genomics, is evolving into numerous dimensions.

The microarrays had a great impact from the way we look at the transcriptome and the way we understand the biology and complexity of it. Microarray expression technologies have allowed the simultaneous analysis of thousands of transcripts that cover nearly the entire genome [25]. Hence, gene expression microarrays, that are providing a comprehensive view of the transcriptional changes that occur during the carcinogenic process, have been applied with great success to the molecular profiling of HCC which has resulted in a much more detailed molecular classification scheme as well as in the identification of potential gene signature sets, molecular biomarkers, prediction of early recurrence and patient survival [26-29].

Over the last decade, numerous studies have applied this technology, and identified a number of candidate genes useful as biomarkers in cancer staging, prediction of recurrence and prognosis, and treatment selection. Considering the complexity of the HCC carcinogenesis many genes may be involved in the initiation and progression of the cancer, and therefore a comprehensive expression analysis using microarray technology has great potential to discover new genes involved in carcinogenesis, as well as may highlight the functional modules and pathways altered in HCC. Indeed, some of the new target molecules that were identified using this technology have been used to develop new serum diagnostic markers and therapeutic targets against HCC to benefit patients.

The first report of cDNA microarray analysis of hepatocellular carcinoma (HCC) by Lau et al. [30] studied the gene expression using about 4000 known human genes in 10 pairs of HCC and non-tumorous tissues. Since then numerous studies have been published to date in the context of genome-wide expression profiling of HCC liver. The microarray analyses of HCC highlighted activation of important pathways in liver carcinogenesis, such as wingless-type (WNT), p53, transforming growth factor (TGF)-β, MAPK signalling pathways [31-34] as well as novel genes with altered expression, such as MARKL1, VANGL1, PEG10, BMAL2, HLA-DR, GPC3, and ROBO1.

Over the past 10 years, the microarray-based gene expression profiling has been used to identify gene signatures associated with etiological factors, histological phenotypes, and clinical phenotypes, as well as unveiling novel subtypes of HCC previously unrecognized.
by conventional methods [26, 35-36]. Most cases of HCC originate from chronic liver disease caused by hepatitis viral infection, including hepatitis B virus (HBV) and hepatitis C virus (HCV), exposure to aflatoxin B1 in mold, and alcohol abuse. In this context, gene signatures associated with different etiologies have also been reported [37-39]. Microarray studies indicated that HBV and HCV viral infections lead to the development of liver cancer by different molecular mechanisms [32, 38-39]. Okabe et al. analyzed expression profiles of 20 primary HCCs by using cDNA microarrays consisting of 23,040 genes, and compared HBV- with HCV-related HCC [32]. The authors identified a gene signature that is correlated with the infection status, and found that genes that are involved in drug metabolism and carcinogen detoxification were differentially regulated between HCV-based and HBV-based HCC. In another study, Iizuka et al. [38] performed genome-wide expression profiling 45 HCC (14 HBV- and 31-HCV-associated) and identified 83 genes whose expression significantly differed between the two types of HCCs. The HBV-associated HCC showed significantly up-regulation of imprinted genes (H19 and IGF2) and genes related to signal transduction, transcription, and metastasis. On the other hand, HCV-associated HCC displayed up-regulation of genes related to detoxification and immune response. Delpuech et al. showed that HBV-associated HCC altered different cellular pathways, those controlling apoptosis, p53 signalling and G1/S transition, whereas the HCV-related HCC resulted in an over-expression of the TGF-beta induced gene [31].

Microarray gene expression profiling together with prediction models have been used in numerous studies to identify gene signatures in tumor or surrounding non-tumorous tissues that can predict vascular invasion, metastasis, post-surgical recurrence, survival, and response to therapy. These signatures may aid in identifying patients most likely to benefit from surgery and chemotherapeutic treatment.

Vascular invasion (VI) is an unfavorable prognostic factor for early HCC recurrence. There have been several microarray studies which identified gene signatures that correlated with VI. Ho et al. [40] identified 14 genes correlated with VI, which can classify patients with high or low risk of VI development and recurrence after curative hepatectomy. In another study, Budhu et al. reported a 17-gene signature expressed in noncancerous hepatic tissues with venous metastasis, capable of predicting recurrence after surgical hepatectomy, with 79% accuracy [41]. Similarly, Wang et al. identified a 57-gene signature to predict disease recurrence at diagnosis (84% sensitivity) [42].

Using supervised machine learning methods on the gene expression data, Nam et al. identified 240 genes that classified samples into different histological grades, from low-grade DNAs to primary HCC [43]. Kim et al. reported 44 genes that can discriminate HBV-positive HCC from non-tumor liver tissues [44]. Iizuka et al. [26] reported a gene signature of 12 genes that can predict HCC patients at high risk of early intra-hepatic recurrence (IHR) after curative surgery (93% sensitivity). Similarly, Kurokawa et al. [45] identified a 20-gene signature which could predict early IHR after curative resection. In another study, a 3-gene signature (HLA-DRA, DDX17, and LAPTM5) found to be predictor of recurrence after curative hepatectomy, which predicted early IHR with 81% accuracy in the validation group [46].
Researchers also used microarrays to identify gene signatures as predictors of survival after surgical resection. Lee et al. analyzed the gene expression profile of 91 HCC samples using unsupervised classification approach which divided the patients into two subclasses with significant differences in survival [29]. The authors also identified genes that accurately predicted the length of survival. Functional analyses indicated that genes related to cell proliferation, anti-apoptosis, and cell cycle regulators were found to be predictor of poor prognosis. Other microarray studies reported c-Met- and TGF-beta regulated genes that are highly associated with the length of survival [47-48].

HCC is a challenging malignancy; most cases of HCC are diagnosed in an advanced stage, and, therefore, treatment options are limited. Hence, it is important to diagnose it at early stage. DNA microarray studies attempted to identify markers for early HCC [27, 49]. Recently considerable attention has been placed on global gene expression studies as well as genomic aberrations in order to understand the pathogenesis of HCC, and to look for possible early markers of detection [14, 16-17, 28, 34, 49-51]. Furthermore, combining cross-species comparative and/or functional genomics approaches from human and animal models of HCC along with genomic DNA copy number alterations enhances the ability to identify robust predictive markers for HCC [13, 36, 52-54]. Thus, characterization of diverse HCC subgroups using the array technologies together with improved analytical approaches are crucial for better management of the disease, especially in the era of personalized medicine approach in HCC treatment.

5. Global miRNA expression profiling of HCC

One of the most important findings of the analysis on the human genome is identification of a significant number of sequences encoding non-coding RNA molecules such as small nucleolar RNAs and microRNAs also known as miRNAs [55-56]. MicroRNAs, single-stranded RNAs typically 21-23 nucleotide long, are untranslated molecules that have capability to bind complementary sequences resulting in their silence, therefore, regulating the expression of their target genes. Some of these molecules have currently been intensely studied and their biogenesis, structure and function are now known and some other small regulatory RNAs are yet to be discovered. Among these different RNAs species, miRNAs holds special attention due to its properties and potential use as therapeutic targets for cancer. Currently around 6000 miRNAs from multiple species have been annotated in different databases. These miRNAs target and regulate around 30% of all protein coding genes in mammals. It has been shown that the miRNAs regulate processes essential to differentiation, apoptosis, cell growth, adhesion, and cell death [57-58]. Recently, due to its oncogenic and tumor suppression activities, these molecules exploited for various cancers, including HCC [59-63].

The genomic instability, transcriptional regulation, and epigenetic alteration have been identified to contribute to the abnormal expression of miRNAs in HCC. Furthermore, the aberrant expression of certain miRNAs is correlated with clinical features of HCC, indicating their potential to serve as diagnostic and prognostic biomarkers of HCC [64-65]. Some aberrantly expressed miRNAs may have a direct role in liver tumorigenesis, and could promote differentiation, cell cycle progression, angiogenesis and invasion, such as
mir-221 and mir-21 [59]. Murakami et al. identified eight miRNAs with altered expression in HCC, which discriminated HCC samples from non-tumor with 97.8% accuracy [66]. Similarly, Huang et al. identified 24 abnormally expressed miRNAs [67]. Toffanin et al. studied miRNA profiling of HCC samples that was previously profiled for mRNA and copy number (CN) changes [65]. The authors identified three subclasses of HCC based on miRNA profiles. The other studies identified miRNA signatures that predicted metastasis potential, recurrence and survival [64, 68-69]. Since the miRNAs are stable in blood, more recently, the circulating miRNAs have been reported as diagnostic markers for various cancers, including the HCC [70-72]. Therefore, identification of miRNAs and their protein-coding target genes is important to understand the mechanisms of hepatocarcinogenesis, and reveals new biomarkers for diagnosis, prognosis and therapeutic targets.

6. Deep sequencing of HCC using next-generation sequencing technologies

Current advances in genomics technologies have been first seen as revolution of microarrays and then recently appeared as high-throughput parallel sequencing techniques. The resolute advance of fluorescence-based standard Sanger technique seemingly stretched to its limits for technical enhancements. At the same time soaring demand for low-cost and high-output sequencing has driven the development of superior technologies that allow massively parallel sequencing processes, producing millions and billions of sequences at once [73-74]. Therefore, it was inevitable to see the replacement of standard sequencing methods to newly emerging advanced sequencing technologies called next generation sequencing technologies. These technologies initially appeared as relatively high-cost difficult techniques for practical use and believed to be useful for only whole genome sequencing of different species. But soon these perceptions were evaded radically. Today, the state of DNA sequencing technologies is in a greater flux than ever before. With this foreseeable evolution, comes new possibilities not only in the field of large-scale genomic sciences from medicine to agriculture and plant sciences coupled with new challenges in data storage and analysis, but also for practical use such as clinical utilization for routine diagnostics [74-78]. Currently, several methods are already established, made significant impact on the field of genomics by having reputable track record for many different published applications and some are in the process of building confidence, some are yet to be tested and perfected [78-84].

The first study of HCC using the next-generation sequencing technology for deep sequencing appeared most recently [5]. Using Illumina’s Genome Analyzer Ilx system, also called GAIlx, Totoki et al. sequenced genomic libraries from a normal Japanese male and hepatitis C-positive HCC sample. Both samples’ sequence reads had an almost complete match to a human reference sequence covering 99.79% and 99.69% for lymphocyte (normal male) and HCC sample genomes, respectively. Nearly ~3 million nucleotide variations were recorded from each genome, yielding 84,555 bases more variations in lymphocyte genome perhaps due to presence of chromosomal alterations in tumor genome and 11,731 of these changes in HCC were somatically acquired. There were several interesting results related to nucleotide changes in the study. First, it was found that occurrence of somatic substitutions was varied between genic and intergenic regions, significantly lower in the genic regions (consisting of coding and noncoding exons, and introns) in comparison to its counterpart.
intergenic regions. This was explained either by negative selection of lethal mutations in the
genic regions or by the existence of specific molecules responsible for the repair of
transcribed region. Second, presence of germline variations was significantly lesser in the
coding regions relative to the non-coding regions. Third, the ratio of nonsynonymous to
synonymous variations (N/NS) either somatic or germline origins differed in HCC and was
significantly lower than that of somatically originated substitutions. To explain this, authors
highlighted the influence of positive selections happening in exons causing survival of
tumor cells or favored negative selection of somatic variations over germline substitutions
on the coding exons. Fourth, the preferred somatic substitutions included T>C/A>G and
C>T/G>A transitions. Fifth, in addition to 81 confirmed somatic substitutions common to
both genomes (all in protein coding regions), 670 small deletions and insertions were
identified and seven of which were validated. Among these variations, some of the changes
seemed more critical since they were located on the previously annotated tumor suppressor
genes for HCC and other cancer types. Moreover, authors decided to resequence exons
potentially harboring malignant changes in 96 HCC and control samples as well as 21 HCC
cell lines. These efforts yielded two critical somatic mutations p.Phe190Leu and p.Gln212X
in LRRC30.

Besides the nucleotide changes, small deletions and insertions, 22 verified chromosomal
rearrangements were identified. These rearrangements were mostly intra-chromosomal and in
close proximity with some known copy number regions. These chromosomal rearrangements
led four different fusion transcripts that involve transcriptional regulation of BCORL1-ELF4
[5]. Then using the deep whole exome sequencing approach (76X or more coverage) a
nonsense mutation in TSC1 gene was also identified in a subset of tumor cells.

As demonstrated in this study, further next-generation sequencing studies have the
potential to reveal novel genes/mutations and likely critical pathways that can be utilized
for the biomarker discovery and identification of novel therapeutic targets for HCC. Besides,
the next-generation technologies have already been proven to be useful for genomic studies
on some cancers [85-94]. Moreover, once affordable prices are reached, such next generation
sequencing techniques will create an amazing opportunity to look for genome-wide DNA
and/or RNA level differences and methylation patterns in many cancer types at an
affordable cost and will open doors for daily diagnostics and personalized medicine [86]
[95-96].

7. Animal models and comparative genomics of HCC

Developing animal models of HCC provide an experimental ground for dissecting the
 genetic and biological complexities of human cancer and contribute to our ability to
 identify and characterize pathogenic modifications relevant to various stages of cancer
development and progression. [97-98]. Several models of constitutive, conditional and
inducible models of HCC were developed inducing genetic manipulations and
investigating the genetic changes. The results usually are comparable to that found in
humans [99]. Each model appears to have its own advantages and disadvantages [100].
Recent studies, including our own, demonstrated the usefulness of modeling human
cancer in diethylnitrosamine (DEN)-induced in rats [54] as well as in genetically
eengineered mice [97, 101].
The recent studies have used cross-species comparative genomics approach, that identifies genes that are conserved in animal models of cancer and in human cancer, that would facilitate the identification of critical regulatory modules conserved across species in the expression profiles and to understand the molecular pathogenesis of various cancers, including HCC [36, 54, 101-103]. The cross-species comparative analysis of animal models and human HCCs would provide new therapeutic strategies to maximize the efficiency of treatments.

8. Integrative and comparative analyses of HCC for identification of novel therapeutic targets and biomarker discovery

It has been shown that CNAs have clear impact on expression levels in a variety of tumors [9, 13, 15]. The presence of such CNAs and LOH may contribute to cancer formation [9-11]. Integrating the gene expression with the CNA data reveals the chromosomal regions with concordantly altered genomic and transcriptional status in tumors [12, 52, 104]. The pattern of genomic modifications in a tumor represents a structural fingerprint that may include the transcriptional control mechanisms and locally impact gene expression levels [10, 12]. Therefore, focusing on differentially-expressed genes with concomitant altered DNA copy number may identify novel early HCC markers of malignant transformation, progression and survival [17].

The studies using integrative analysis of genomic aberrations with the expression profiling demonstrated the usefulness of this approach to identify the likely drivers of cancer [105] and helped better understand the processes affected by the drivers/passenger factors and led to obtain novel insights into pathobiology of HCC [17, 54, 105].

In this context, we performed cross-species and integrative genomic analysis to identify potential biomarker genes for early HCC [54]. In this study, we first developed a rat model of early HCC as well as liver regeneration post-hepatectomy and compared them to normal liver using a microarray approach. We then performed a cross-species comparative analysis coupled with CNAs of early human HCCs to identify the critical regulatory modules conserved across species. We identified 35 gene signature conserved across species, with more than 50% mapping to human CNA regions associated with HCC [54]. Combining cross-species comparative and/or functional genomics approaches from human and animal models of HCC along with genomic DNA copy number alterations enhances the ability to identify robust predictive markers for HCC [13, 36, 52-54].

9. Future directions

Elucidating the molecular pathogenesis of HCC on human samples has been an onerous task due to certain limitations such as varying etiologies among studied patients, changes likely to arise during the different stages of the disease or progression of HCC, and heterogeneity of the disease. Moreover, the success of studies is hampered by the fact that hepatic transcriptome is among the most complex of any organ, and the study of tumor formation in liver can be thorny and complicated by the continuous change of the transcriptome during liver regeneration after hepatectomy. Besides, cancer progresses through a series of histopathological stages during which genetic alterations accumulate
and, in consequence, the pattern of genetic expression changes complicates the interpretation of the genetic changes in human HCC. These limitations have hampered development of proper therapeutics that was further complicated by recurrences even after aggressive local therapies.

The advances in high-throughput “omics” and next-generation sequencing technologies have been providing unprecedented biological insights related to pathogenesis of HCC. Undoubtedly, comparative and integrative genomics approaches are promising to lead to novel and robust biomarkers for improved diagnosis, prognosis, and treatment of HCC. The systems approach via the integration of data reflecting alterations at genomic, transcriptomic, proteomic, and epigenomics levels will ultimately converge toward a personalized medicine that will improve diagnosis, treatment and prevention of liver cancer.

10. References


[96] Mardis ER: The $1,000 genome, the $100,000 analysis? *Genome Med* 2010, 2:84.


This book is oriented towards clinicians and scientists in the field of the management of patients with liver tumors. As many unresolved problems regarding primary and metastatic liver cancer still await investigation, I hope this book can serve as a tiny step on a long way that we need to run on the battlefield of liver tumors.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
