Liver cancer (also known as hepatic cancer or primary hepatic cancer) is a malignant cancer originating in the liver which severely affects public health and quality of life. Globally, liver cancer is the 5th most common cancer in male patients and the seventh most common in female patients. The majority of liver cancer cases occur in the developing countries, especially in Southeast Asia and sub-Saharan Africa where there is a high prevalence of hepatitis B virus (HBV) infection (1).

Hepatocellular carcinoma (HCC) and epigenetics

HCC is the most common histological subtype of liver cancer, accounting for 70–85% of total liver cancer cases (2,3). In 2008, around 700,000 HCC cases were diagnosed worldwide (4). In the past 20 years, the HCC incidence in the U.S. increased by three times with a 5-year survival rate lower than 12%, and HCC related to persistent hepatitis C virus (HCV) infection has become fastest-growing cause of cancer-related death (1).

In China, HCC is the third most common cause of cancer-related death in both male and female patients (5). According to previous studies, the inhibited growth rate of HCC in China is attributable to several different advances in public health, including effective control of contagious HBV and HCV infection, reduction of food contamination by aflatoxin, improved quality of drinking water, and medical injection procedures which have been more standardized and safer. After the Chinese government initiated HBV vaccination to infants and children in areas with high risk of HCC in 1986, fatality due to HCC decreased by 95% over the following 15 years among those...
aged 0–19 (6-10). However, the overall incidence of HCC is still rising, due to various factors including unhealthy lifestyle (smoking, heavy drinking, input of high-energy food), and non-alcoholic fatty liver disease.

The pathogenesis of HCC involves interaction between environmental and genetic factors (1,11,12). The major factors in HCC development include HBV infection, HCV infection, alcohol abuse, non-alcoholic fatty liver disease, and exposure to environmental carcinogens (e.g., aflatoxin B1). Among the less common factors are hereditary hemochromatosis, α1-antitrypsin deficiency, autoimmune hepatitis, hematoporphyrinism, and Wilson’s disease (1,13). The precise molecular mechanisms of HCC are very complicated and heterogeneous, impeding clinical diagnosis and treatment (3).

To date, the main issues with HCC treatment are insufficient early detection and rapid tumor development. Despite the fact that clinical and imaging criteria for early HCC diagnosis have been available, and that surgical dissection and liver transplant have been commonly performed as mature therapeutic approaches, the sensitivity and specificity of existing tumor biomarker tests (e.g., Alpha-fetoprotein) are still relatively low when screening for HCC. Most patients are diagnosed and treated at later stages of HCC, which greatly limits the selection of available treatments and leads to poor prognoses (3,14-16). In addition, due to patient variability in tumor development rates, HCC patients receiving the same treatment may exhibit different prognoses and survival rates.

Reliable prognosis prediction and classification/categorization cannot be obtained based on current diagnostic and evaluating methods (17). Therefore, improving early detection can facilitate the selection of treatment options, which may help to achieve better clinical outcomes for HCC patients.

In 1942, Waddington discovered the phenomenon of inconsistency between genotype and phenotype, giving rise to the discipline of epigenetics which plays a critical role in the regulation of gene expression (18). The key aspects of epigenetics include DNA methylation, genomic imprinting, maternal effects, gene silencing, nucleolar dominance, activation of dormant transposons, and RNA editing. Epigenetics studies address how environmental factors interact with the genome and generate heritable modification of gene expression and cellular phenotype without altering DNA sequence. Epigenetic modification can be affected by cellular metabolism and extracellular environments, such as pathogenic factors. Therefore, studying epigenetics is critical to understanding the mechanisms underlying the pathogenesis and progression of human diseases (17-19).

DNA methylation abnormalities, which play an important role in the onset of tumor development (19-28), have been observed in multiple cancers including HCC. According to several studies, abnormal DNA methylation can be used as a molecular biomarker for tumor screening and prognosis prediction (29-32). Screening for DNA methylation abnormalities, in combination with integrated analysis of gene expression, can be used to identify the epigenetic events driving tumorigenesis (33,34). Several groups have reported that the of the unique and specific arrangement of DNA methylation abnormalities is correlated to HCC diagnosis and prognosis (13,28,35-38). Regional DNA hypermethylation is also known to occur during the early stages of HCC; thus, studying the association between DNA methylation levels and HCC staging may be useful for clinical diagnosis and prediction (39).

Emerging techniques for examination of targeted genome regions/loci

Three known kinds of DNA methylation modification are 5-methylcytosine (5mC), present in eukaryotes; N6-methyladenosine, mainly seen in bacteria and pathogenic organisms; and 7-methylguanine, which rarely occurs. In vivo formation of 5mC is catalyzed by DNA methyltransferases (DNMTs) using S-adenosyl methionine as a methyl provider, and selectively attaching a methyl group to the 5th carbon atom of the pyrimidine ring of cytosine (40). In eukaryotes the majority of 5mC formation occurs in the genomic regions that are abundant with cytosine-phosphate-guanine dinucleotides, also known as CpG islands. CpG islands are characterized by high CG percentages (usually >50%) and 200–500 base pairs.

Examination of DNA methylation can be divided into several categories: global DNA methylation analysis targeting the whole genome, gene-specific or locus-specific DNA methylation analysis, and methylation pattern/profiling analysis. Different approaches and techniques provide DNA methylation information on different levels, depending upon the goals of examination (41).

Techniques for global DNA methylation analysis include high-performance liquid chromatography (HPLC) (42), high-performance capillary electrophoresis (HPCE) (43), thin-layer chromatography (TLC) (44), methyl group acceptance assay (45), chloroacetaldehyde assay (46), and immunochemical analysis (47).

Techniques for gene-specific DNA methylation analysis
include methylation-sensitive restriction enzyme-Southern (MSRE-Southern) (48), methylation-sensitive restriction endonuclease-PCR (49,50), methylation-specific PCR (51), methylation-sensitive single-nucleotide primer extension (52), direct bisulfite genomic sequencing (49), methylight PCR (53,54), methylation-specific single-strand conformation analysis (55-57), methylation-specific denaturing gradient gel electrophoresis (58), methylation-specific melting curve analysis (59), methylation-specific denaturing high-performance liquid chromatography (60), and methylation-specific oligonucleotide microarray (61).

Methylation pattern/profiling analysis includes restriction landmark genomic scanning (62), methylated CpG island amplification (63)-representational difference analysis (64) methylation-sensitive arbitrarily primed PCR (65), and differential methylation hybridization (66,67).

Traditional methods of HCC methylation analysis target specific genes and loci (e.g., oncogenes and tumor suppressor genes) and quantitatively measure methylation levels (68-70). For example, in combined bisulfite restriction analysis (COBRA), researchers process DNA samples with sodium bisulfite, amplify target sequences with PCR, and digest PCR products with specific restriction enzymes. Electrophoresis, probe hybridization, and quantitative scanning are then used to obtain the methylation ratio in the tested samples. Gene methylation levels can be determined from analysis of PCR bands, as methylated bands cannot be digested by restriction enzymes (71). The COBRA technique has been widely used for DNA methylation screening due to its many strengths—simplicity, efficiency, high compatibility, and low requirements for sample sequences availability. However, the constraints imposed by specific restriction enzyme sites do not allow for determination of complete genomic DNA methylation signature. Gene-specific or locus-specific DNA methylation analysis can provide methylation information about specific sites known to be associated with HCC. Methylation analysis, in combination with integrated analysis of gene expression, may also facilitate exploration of potential gene functions and possible epigenetic cancer drivers.

According to a study using a bisulfite sequencing technique to measure methylation levels in HCV-infected liver tissues, the promoter regions of two Wnt inhibitor genes—SFRP2 and DKK1—were hypermethylated in precancerous stages of chronic HCV infection and cirrhosis, which led to the silencing of Wnt and eventually the induction of HCC. The methylation level of the SFRP2 promoter region was highest in HCC, followed by cirrhosis and chronic hepatitis, and lowest in normal liver (72).

Nishida et al. performed methylation-specific PCR and observed the epigenetic inactivation of some tumor suppressor genes (HIC1, GSTP1, SOCS1, RASSF1, CDKN2A, APC, RUNX3, PRDM2) in the early stages of HCC. This specific subset of tumor suppressor genes was significantly hypermethylated, and their methylation frequencies were associated with time-to-HCC occurrence in chronic hepatitis C patients (CHC). Thus, the altered methylation of these genes may play a critical role in the early steps of carcinogenesis, and can provide potential prognostic value for determining the risk of CHC developing into HCC (68). Some tumor suppressor genes (CASP8, MINT31, PTGS2, CACNA1G) were found to present unaltered methylation levels in the early steps of HCC, but continuously increased methylation along with the progression of HCC. In contrast, another seven tumor suppressor genes (RASSF2, MINT2, MINT2, RPRM, SFRP2, CDH1, DCC) were merely hypermethylated in the highly aggressive (high grade) stages of HCC, suggesting that different tumor suppressor genes function at different stages or times of HCC development (68).

Several previous studies have reported that the expression levels of oncogenes (c-Myc, cyclinD1, β-catenin) and tumor suppressor genes (P16INK4A, P53, E-cadherin, DLC-1, pRb) are decreased to different extents in HCC (74). In human HCC the ubiquitous activation of Ras and JAK/STAT signaling pathways may be caused by the coincident abnormal methylation of inhibitors of Ras and JAK/STAT (75). Using methylation-specific PCR to analyze the promoter regions of 105 genes, Calvisi et al. observed the epigenetic silencing of Ras pathway inhibitors (SOCS-4,5, SPRY1, SPRY2, DAB2) and activation of Ras gene and its downstream genes (ERK, AKT, RAL). The down-regulation of the Ras inhibitors SPRY1, SPRY2, DAB2, SOCS4, SOCS5 and angiogenesis inhibitors BNIP3, BNIP3L, IGFBP3, EGLN2 were shown to be correlated with poor prognoses in HCC.
patients, indicating a relationship between site-specific DNA hypomethylation and tumor development (76). At the same time, the Ras signaling pathway promotes tumor angiogenesis through activating transcription of target genes—VEGF, HIF-1α, IL-8 (77,78). The expression levels of Ras and its downstream effectors are highest in HCC cells, followed by para-tumor tissues, and lowest in normal liver, suggesting important roles in the induction of angiogenesis in human HCC. Meanwhile, the DNA methylation index in HCC patients is inversely correlated with survival period, linking the consistently increased methylation of promoter regions in tumor suppressor genes with the progression of HCC (see Table 1).

Another important effector of HCC development and progression is tissue inhibitor of metalloproteinase-3 (TIMP-3), which is a matrix metalloproteinase (MMP) that can inhibit cellular growth and migration and prevent the formation of vasculature. TIMP-3 plays important roles in multiple cellular physiological processes including cell proliferation, migration, differentiation, angiogenesis, apoptosis, and so forth. Mannello et al. found that the hypermethylation of CpG islands of the TIMP-3 sequence in HCC patients caused its downregulation through epigenetic inactivation, and facilitated the metastasis of cancer cells (109).

The LINE-1 retrotransposon composes approximately 17% of the human genome, and its methylation level imposes great influence on global genomic methylation (110,111). Previous literature has suggested that hypomethylation of LINE-1 is closely related to poor prognoses in different cancers, including cancers of the colon, esophagus, stomach, and ovary (112-115). Several recent studies have discovered the significant hypomethylation of promoter regions of LINE-1 in the early stages of HCC (111,112,116-118). The activation of LINE-1 may be mediated by c-Met activation. Activated LINE-1 not only induces HCC, but is also significantly associated with poor prognosis, and may therefore be used as a potential biomarker to evaluate the clinical progress of HCC patients (112,116-118). Similarly, hypomethylation has been found in imprint gene regions, Alu sequences also exhibit active DNA demethylation in early steps of HCC (111). In HCC patients the hypermethylated H-cadherin gene becomes inactivated, and the hypomethylated vimentin gene becomes upregulated, leading to the assumption that the interaction of H-cadherin and vimentin prompts the onset and development of HCC (113).

At present there are several reports in the literature of atypical methylation patterns, global hypomethylation, or hypermethylation on specific gene promoter regions in HCC tumor cells. The hypermethylation of CpG islands in the promoter regions of tumor suppressor genes leads to transcription silencing and genomic instability (119-126). Calvisi et al. evaluated the methylation levels of HCC samples by assessing the incorporation of [3H]dCTP into DNA, and found that the extent of genome-wide hypomethylation and CpG hypermethylation correlates with the biological features and clinical outcomes of HCC patients (76).

High-throughput sequencing techniques

Advances in chip-based examination techniques

Along with the rapid development of high-throughput biology techniques, high-density DNA methylation array has become more and more widely used in HCC studies. Most previous studies have focused on only one or a few genes, the most comprehensive of which tested as many as 105 genes (69,76,104,127-132). Aided by the new generation sequencing techniques, however, HCC methylation studies transitioned from single-gene/site methylation analysis to global/genomic methylation analysis. Moreover, associated statistical analysis methods—hierarchical cluster analysis, principal components analysis, etc.—are capable of deriving further useful information.

Recently, Illumina chip have been used to study the DNA methylation patterns in various HCC models using diverse sample sizes, examination techniques, and etiological factors. A study using Illumina 1500 Golden Gate arrays and methylation-specific PCR found abnormal methylation on CpG promoter regions of 24 new genes as potential methylation markers for HCC; 4 of these genes (TNFRSF10C, HOXA9, NPY, and IRF5) were hypermethylated and subsequently down-regulated in the tumor tissues, and thus can be used as HCC biomarkers (90).

Another study using similar techniques discovered DNA methylation signatures associated with pathogenic risk factors, tumor progression, diagnosis, and prognosis of HCC. Abnormal modes of methylation were defined as hypermethylation of promoter regions and hypomethylation of non-promoter regions. DNA methylation of 124 CpG sites from 94 genes showed significant differences between tumor tissues and para-tumor tissues, with 34 genes being hypermethylated (including RASSF1, APC, and CDKN2A) and 90 being hypomethylated (including GABRA5,
<table>
<thead>
<tr>
<th>Gene functions</th>
<th>Hypermethylated genes</th>
<th>Hypomethylated genes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cell cycles, cell proliferation, cell growth, cell apoptosis</strong></td>
<td>ALOX12 (79), APC (28,68,80-84), CACNA1G (68), CASP8 (68,85,86), CCND2 (28,81,82), CDKL2 (36,87), CDKN2A (28,36,68,81,82,88), DAB2IP (36), DCC (68), DKK1 (7,72), DRC-1 (74,84,89), DNMT1 (28), Fzd7 (28), Hic1 (68), HIST1H3G (36), IRAK3 (28), IRF5 (90), MYOD1 (28), NFATC1 (91), P15INK4B (92,93), P16INK4A (74,80,93,94), P53 (74,95), PTGS2 (68), RASSF1A (28,68,80,87,96), RASSF2 (68), RPRM (68), RASSF3 (68), RASSF4 (68), RASSF5 (68), RASSF6 (68), RUNX3 (68,82,84), SFRP2 (68,72,97), SMDP3 (98), SOCS1 (68,75,80,99,100), TNFRSF10C (90), ZNF154 (36)</td>
<td>AKT3 (36), c-myc (74), NENF (101), NOTCH4 (28), PGR (28), SFN (87), TNFSF15 (79)</td>
</tr>
<tr>
<td><strong>Cell migration and movement</strong></td>
<td>ALOX12 (79), APC (28,68,80-84), CDH1 (28,68,84), CRMP1 (79), DAB2IP (36), Fzd7 (28), PTGS2 (68), RASSF1A (28,68,80,87,96), ST8SIA6 (102), TIMP-3 (76,95,103), VIM (102)</td>
<td>MMP1 (28), MMP3 (28), MMP9 (28)</td>
</tr>
<tr>
<td><strong>Cell differentiation, cell potency</strong></td>
<td>DAB2 (76), HOXA9 (90), IRF5 (90), NFATC1 (91), RUNX2 (102), SFRP2 (68,72,97), ZNF154 (36)</td>
<td>NENF (101), NOTCH4 (28), PAX4 (36), PGR (28)</td>
</tr>
<tr>
<td><strong>Immune and inflammatory responses</strong></td>
<td>DAB2IP (36), GSTM2 (28), IRF5 (90), STEAP4 (36), VIM (102)</td>
<td>CCL20 (36), HLA-DQA2 (28), IFNG (28), SPP1 (87)</td>
</tr>
<tr>
<td><strong>Metabolism related</strong></td>
<td>GSTP1 (68,82,84,91,92,104-107), STEAP4 (36)</td>
<td>AKT3 (36)</td>
</tr>
<tr>
<td><strong>Tumor angiogenesis</strong></td>
<td>AKT3 (36), DAB2IP (3), PTGS2 (68), SFRY1 (76), SFRY2 (76)</td>
<td>TNFSF15 (79)</td>
</tr>
<tr>
<td><strong>Signal transduction</strong></td>
<td>BMP4 (36,91), CELSR1 (79), DAB2 (76), DAB2IP (36), IGFR1 (28), MINT1 (68), MINT2 (68), NFKB (28), SOCS4 (76), SOCS5 (76), ST8SIA6 (102), VIM (102)</td>
<td>EGF (28), FGFR6 (28), RASAL2 (101)</td>
</tr>
</tbody>
</table>
NOTCH4, and PGR). These abnormally methylated genes are involved in multiple critical cellular activities: apoptosis (IRAK3), immunoresponse (GSTM2), cell cycle regulation (CCND2), etc. Hypomethylation of imprint genes (GABRA5, GABRG3, ZIM3) was also reported in this study. Some of the CpG sites found were associated with pathogenic risk factors of HCC, as specifically indicated in this study: HBV infection (NAT2, VSPG2, DCC, NTRK3, etc.), HCV infection (RIK, CHGA), and alcohol consumption (DIO3, STAT5A) (28).

Hernandez-Vargas et al. used Illumina Golden Gate BeadArray technology to examine 30 HCC samples with HBV or HCV infection, and found that HCC tumor cells exhibited special DNA methylation patterns. Meanwhile, DNA methylation states were closely associated with major infection factors and tumor progression. Of all 1,505 CpG islands that were examined, those in the promoter regions of the APC, RASSF1A, CDH12A and FZD7R3 genes showed significantly different methylation states between tumor tissue and para-tumor (normal) tissue. In addition, methylation of the DNMT1 promoter was linked to the low differentiation level of tumor (28).

One of the representative methylation arrays, the Infinium Human Methylation 27K beadchip, is widely used to quantitatively measure the methylation levels at different sites in the genome, and can cover 27,578 CpG sites from 14,495 genes. These genes include RefSeq genes from the NCBI CCDS database, oncogenes with high methylation frequency, tumor-associated target genes, and promoters of miRNAs (133). During Infinium Human Methylation 27K beadchip analysis, DNA samples undergo bisulfite salt treatment, whole genome amplification, hybridization and single-base amplification, fluorescent staining, and chip scanning before acquisition of DNA methylation data. Several studies have affirmed the reliability and comparability of chip-based methylation analysis and traditional pyrosequencing-based methylation analysis (36,134-136). Compared to classical single-gene/site methylation analysis, chip-based methylation analysis exhibits great advantages due to its high methylation detection capacity, absence of PCR steps, and its integration of both gene expression and microRNA profiling. As such, examining methylation status in HCC is not limited to researching methylation abnormality on specific well-known sites, but extends to massive exploration of the methylation distribution of CpG sites. The chip-based methylation analysis is therefore of great importance to our understanding of the whole picture of epigenetics in HCC and the disclosure of additional critical methylation sites.

Ammerpohl et al. used Illumina Human Methylation 27K array to analyze 27,578 CpG sites and found unique methylation patterns in promoter regions, CpG islands, and PRC2-targets in cases of cirrhosis and HCC. Compared to normal liver tissues, HCC cells had 998 CpG sites with hypomethylation and 278 CpG sites with hypermethylation. The genes corresponding with abnormal methylation were involved in cellular immunity, communication, signaling transduction, and cell movements. Compared to cirrhotic tissues, HCC had 1,050 CpG sites with hypomethylation and 378 CpG sites with hypermethylation, and corresponding genes were involved in maintenance of cellular structure, signaling pathways, cell movements, and immunity. The hypermethylated signaling pathway components in the study included heterogeneous trimeric G protein pathways (Gia and Gsa regulatory pathways), PI3K pathway, TNF-β pathway, etc.; the hypomethylated pathways included angiogenesis signaling pathway, acetylcholinergic M1&3 pathway, TNF-β and EGF receptor pathway. Specifically, eight sites located within eight genes (SPRR3, TNFSF15, ALOX12, ANGPTL7, CELSR1, CRMP1, GNRH2 and LOC55908) showed significantly different methylation levels between HCC, cirrhosis, and normal liver tissues (79).

One large-scale research project using the Illumina Human Methylation 27K array discovered that a total of 2,324 CpG sites significantly differed in methylation level between HCC tumor and non-tumor tissues, with 684 CpG sites (covering 548 genes) significantly hypermethylated and 1,640 (covering 1,290 genes) hypomethylated in tumor tissue. DAB2IP, BMP4, ZFP41, SPDY1, and CDKN2A were the most hypermethylated genes, while CCL20, AKT3, SCGB1D1, WDFC6, and P4X4 were the most hypomethylated genes. Within the tumor tissues, the majority of CpG islands showed significant hypermethylation, and most non-CpG sites were significantly hypomethylated, regardless of their proximity to promoter regions. Importantly, this study identified abnormal methylation in two novel pathways (Wnt and 5-HT4-type receptor-mediated signaling), as well as finding ten cellular pathways that overlapped with those from a previous study by Ammerpohl et al. (79).

Designed as an improved chip-based methylation sequencing technique, the Infinium Human Methylation 450K Beadchip (abbreviated as 450K beadchip) can cover more than 480,000 CpG sites, including 99% of RefSeq genes, 96% of CpG islands in the UCSC database, 92%
of CpG shores (2 kb areas around CpG islands) and 86% of CpG shelves (2–4 kb away from CpG islands), HMM islands, FANTOM 4 activators, differentially methylated regions (DMRs), informatically predicted enhancers, DNase-sensitive sites, Ensemble regulatory features, loci in MHC regions, non-CpG methylation sites in human stem cells, as well as 90% sites from 27K beadchips (137). This 450K chip can be used to perform quantitative methylation analysis for single CpG sites, and also to provide a suitable platform for genome-wide association study (GWAS). The genomic regions with low density of CpG islands may be easily missed by traditional methylation sequencing, but can be detected with the 450K chip.

Compared with the 27K beadchip, using the 450K chip to examine global DNA methylation can help to uncover more sites with abnormal methylation, and simultaneously oversee the methylation distribution of whole genome.

DNA methylation changes in HCC have been confirmed by a large number of researchers, especially on the level of the whole genome (20,36,68,97,105). One large-scale study analyzed the global DNA methylation data from 800 HCC patients using the 450K beadchip and verified the presence of global DNA hypomethylation (GDH) or CpG island methylator phenotypes (CIMP) in around 90% of HCC samples (35). Another study reported the finding of 28,017 CpG sites with hypermethylation and 102,495 CpG sites with hypomethylation by comparing the DNA methylation pattern between HCC tissue and para-tumor tissues from 66 patients. None of the genes exhibited simultaneous abnormal hypermethylation and hypomethylation, implying that these two epigenetic events are independent biological mechanisms (13). Due to the extensive genome coverage available with the 450K beadchip, this study also explored the lesser-known genomic regions referred to as CpG shores and shelves. While most of the significantly hypermethylated CpG sites were located in CpG islands, the majority of the significantly hypomethylated CpG sites sit in open sea regions, shores (2 kb area around CpG island), and shelves (2–4 kb away from CpG island) (13). Similar observations have been made for other tumors (e.g., colorectal cancer), indicating that the location of abnormal methylation in CpG islands, shores, and shelves may be relevant to the onset of tumorigenesis (138-140).

Methylation beadchip analysis can also reveal p53 induction statistical analysis about etiological or risk factors in HCC patients. According to a study by Shen et al. that tested liver tissue samples with 27K DNA beadchips, one CpG site of the VPREBI gene in non-tumor tissue showed significantly different methylation levels between alcohol users and non-users. Similarly, in HCC tumor tissues, methylation levels of CRISPLD1, PCDHB2, PCSK1, LXH1, KCTD8, TSHD3, and CXCL12 could distinguish alcohol users from non-users (36). In another study from Shen et al using 450K beadchips to analyze adjacent non-tumor tissues, 228 CpG sites were found to be significantly associated with HCV infection, and 17,207 CpG sites were associated with cirrhosis. Methylation levels on one CpG site of the OAZ3(MRPL9) gene exhibited significant difference between smokers and non-smokers. In the future, abnormal DNA methylation profiling maybe used to further distinguish HCC induced by different etiological and risk factors in HCC (e.g., HCV) (13).

Using 450K beadchips, Song also measured the DNA methylation profiles of 27 HCC samples and 20 para-tumor liver tissues. This study analyzed 10,775 differentially methylated sites distributed in or around gene promoter regions, 4% of which are known methylation sites. An additional 10,325 sites (covering 4,106 genes) may be differentially methylated in HCC (105). These DNA methylation sites are distributed abundantly in the CpG islands of promoter regions, which play important roles in regulatory signaling networks involving cell growth, gene expression regulation, apoptosis, and carcinogenesis. CpG sites of promoter regions of BMP4, CDKN2A, GSTP1 and NFATC1 genes showed the highest methylation levels. Similar studies have suggested that hypermethylation of promoter regions of tumor associated genes (i.e., suppressor genes and cell cycle genes) can inhibit their expression and thus increase the chance of HCC onset (141).

Many recent HCC-related epigenetic studies not only focused on the biological meaning of abnormal methylation, but also attempted to link biological discoveries with clinical applications. Samples taken from a cohort of 304 HCC patients were applied to a450K beadchip and subjected to random forest processing method to obtain a diagnostic methylation signature composed of 36 methylation probes. This methylation signature, which is supported by data from HCV-associated HCC and alcohol-associated HCC, accurately predicted survival rates independent of clinical and pathological risk factors (e.g., multiple hepatic nodes, platelet count), and independently predicted the HCC recurrence rate. In addition, patients displaying a particular methylation signature and mRNA signature were more inclined to express the molecular subclass of proliferation with progenitor cell features (37,38). Zheng et al. analyzed global DNA sequencing data from 800 HCC samples, selected ten sites out of 80 hypermethylated
promoter sites as candidate biomarkers to distinguish between tumor and non-tumor tissues, and reported that the sensitivity and specificity were both close to 100%. The selected sites included the CDKL2, TBX15, and NKX6-2 genes which have been previously reported. By integrative analysis of DNA methylation and gene expression data, Zheng also identified 222 candidate epigenetic drivers involved in inflammatory response and metabolic pathways, several of which—SFN, SPP1, and TKT were significantly associated with overall survival rates of patients (35).

**Advances in next generation sequencing based examination techniques**

As of now, bisulfite sequencing remains the gold standard of DNA methylation profiling analysis (142). Recently, however, new examination techniques based on next generation sequencing have been developed and utilized in DNA methylation study. When using sequencing techniques to acquire methylation data, designing probes for known methylation levels is unnecessary; thus there are no constraints from previous experiments. However, the replicability of sequencing techniques is incomparable to that of chip array, and the processing of sequencing data is much more complicated. Sequencing based methylation analysis needs to be improved and optimized before it is fit for clinical application.

It has recently been discovered that the formation of 5-methylcytosine (from methylation of cytosine) can be modified by ten-eleven translocation methylcytosine dioxygenase, which oxidizes 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) (143,144). Hydroxymethylation, as one of several DNA demethylation mechanisms, was found to be related to gene expression regulation and tumorigenesis (88,145). Traditional bisulfite sequencing can scarcely distinguish 5hmC from 5-Mc, thus restricting the study of hydroxymethylation to some extent. Oxidative bisulfite sequencing (oxBS-seq) is the first method to quantitatively measure 5hmC levels in genome DNA at single-base resolution, and is distinguished by its simplicity, efficiency, and simultaneous collection both methylation and demethylation data (146). The oxBS-seq technique uses highly selective oxidative potassium perruthenate to convert 5-hmc into 5-formylcytosine (5fC), which is then converted into uracil by bisulfite treatment before finally being sequenced. Genome DNA 5hmC patterns can be obtained by comparing and mapping 5fC sequencing results with traditional bisulfite sequencing on the same sample (146).

Meanwhile, another new sequencing technique, TET-assisted bisulfite sequencing (TAB-seq), uses β-glucosyltransferase mediated glycosylation and TET1 mediated oxidation to convert 5hmC to 5-carboxycytosine (5caC). 5caC and unmodified cytosine can further be converted into uracil by bisulfite treatment, thus achieving the goal of quantitatively measuring 5hmC levels at single-nucleotide resolution (147).

Li et al., using the oxBS-seq technique, identified the global DNA methylation and hydroxymethylation signatures of normal livers and liver tumors at single-base resolution. 5hmC was enriched in tissue-specific DMRs and tumor-specific DMRs, and was inversely correlated with methylation modification (148). Expanding the findings from previous experiments showing that 5hmC is abundant in gene promoter regions, gene coding regions, and distal cis-regulatory elements-like enhancers (144), recent studies have reported that 5hmC is also enriched in CpG island shores (148).

Another widely used technique based on bisulfite sequencing is reduced representation bisulfite sequencing (RRBS). It uses MspI restrictive enzyme to cleave the genome, followed by adapter modification, bisulfite treatment, PCR amplification and sequencing, to determine the methylation levels in samples (149). RRBS covers most of CpG island and promoter sites, but this coverage only accounts for 10% of all CpG sites of human genome, making it difficult to monitor the genomic region’s lack of CpG sites (150).

Whole genome bisulfite sequencing (WGBS) is another technique based on bisulfite sequencing. The sequencing is performed following the fragmentation of the whole genome and sodium bisulfite treatment, which covers 5-methylcytosine within the range of whole genome (except overlapping regions). The WGBS method resolves the issues of low resolution, biased sequencing specificity, and partial genome analysis that are characteristic of traditional techniques (142,151,152). Lister et al. performed WGBS and obtained human methylation signatures at single-base resolution (153), but this technique has not yet been used for analyzing clinical samples (151). Considering the vast amount of data generated by WGBS, appropriate data processing approaches remain to be explored, and consistency between experimental and clinical results needs to be further verified (see Table 2).

Methylated DNA immunoprecipitation sequencing (mDIP-seq) and methylated DNA binding domain sequencing (MBD-seq) are global methylation analysis
Table 2 DNA methylation examine techniques/platforms

<table>
<thead>
<tr>
<th>Techniques</th>
<th>DNA methylation examine techniques/platforms</th>
<th>Features</th>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chip-based techniques</td>
<td>Illumina Infinium Human Methylation27 Beadchip</td>
<td>Covers 1505 CpG sites in 807 tumor related genes</td>
<td>Examined and compared DNA methylation levels in tumor and para-tumor tissues in 30 HCC samples. Found significantly different methylation levels in 94 genes of tumor tissues, including 34 hypermethylated CpG sites (RASSF1, APC, CDKN2A, etc.), and 90 hypomethylated CpG sites (GABRA5, NOTCH4, PGR, etc.)</td>
<td>(74,154)</td>
</tr>
<tr>
<td></td>
<td>Illumina Infinium Human Methylation27 Beadchip</td>
<td>Covers 14,495 genes (27,578 CpG sites), including RefSeq genes from NCBI CCDS database, oncogenes with frequent methylation, tumor related genes, promoters of microRNA</td>
<td>Examined and compared DNA methylation levels in tumor and para-tumor tissues in 62 HCC samples. Found significantly different DNA methylation levels in 2,324 CpG sites. 684 CpG sites from 548 genes exhibited significant hypermethylation in tumors (DAB2IP, BMP4, ZFP41, SPDY1, CDKN2A, etc.), while 1,640 CpG sites from 1,290 genes exhibited significant hypomethylation in tumors (CCL20, ATK3, SCGB1D1, WFDC6, PAX4, etc.)</td>
<td>(36,133)</td>
</tr>
<tr>
<td></td>
<td>Illumina Infinium Human Methylation 450K Beadchip</td>
<td>Covers more than 480,000 CpG sites, including 99% RefSeq genes, 96% CpG islands in UCSC database, 92% CpG island shores, 86%; island shelves, HMM islands, FANTOM4 activators, differentially methylated sites in normal tissues compared to tumor tissues, DNase hypersensitive sites, Ensemble regulatory features, loci in MHC region, identified non-CpG methylation sites in human stem cells. Also covers 90% of methylations sites in Infinium Human Methylation 27K BeadChip</td>
<td>Examined and compared DNA methylation levels in tumor and para-tumor tissues in 66 HCC samples. Identified 28,017 hypermethylation sites and 102,495 hypomethylation sites in tumors</td>
<td>(13,118,137)</td>
</tr>
<tr>
<td>New-generation-sequencing based techniques</td>
<td>Oxidative bisulfite sequencing (oxBS-seq)</td>
<td>5hmC was catalyze by KRU04 and converted to 5Fc, subsequently treated by bisulfite and converted to uracil to be sequenced. DNA 5hmC profile with single-base resolution can be obtained by comparing oxBS-seq results with traditional bisulfite sequencing results</td>
<td>Examined and identified global DNA methylation and hydroxymethylation profiles in normal liver and liver tumors. 5hmC is enriched in tissue-specific differentially methylated regions and tumor-specific differentially methylated regions. 5fmC is negatively associated with methylation modification</td>
<td>(146,155)</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Techniques</th>
<th>DNA methylation examine techniques/platforms</th>
<th>Features</th>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>New-generation-sequencing based techniques</td>
<td>TET-assisted bisulfite sequencing (TAB-seq)</td>
<td>Uses β-glucosyltransferase mediated glycosylation and TET1 mediated oxidation to convert 5hmC to 5-carboxylcytosine (5caC). 5caC and unmodified cytosine can further be converted into uracil by bisulfite treatment, thus achieving the goal of quantitatively measuring 5hmC levels at single-nucleotide resolution</td>
<td>–</td>
<td>(156)</td>
</tr>
<tr>
<td>Reduced representation bisulfite sequencing (RRBS)</td>
<td></td>
<td>uses MspI restrictive enzyme to cleave the genome, followed by adapter modification, bisulfite treatment, PCR amplification and sequencing, to determine the methylation levels in samples. Covers around 10% of CpG sites of total human genome</td>
<td>–</td>
<td>(157)</td>
</tr>
<tr>
<td>whole genome bisulfite sequencing (WGBS)</td>
<td></td>
<td>Sequencing performed following the fragmentation of the whole genome and sodium bisulfite treatment, covers 5-methylcytosine within the range of whole genome (except overlapping regions). Can obtain human DNA methylation pattern at single-base resolution</td>
<td>–</td>
<td>(16,152)</td>
</tr>
<tr>
<td>Methylated DNA Immunoprecipitation-Sequencing (MeDIP-Seq)</td>
<td></td>
<td>Recognize 5Mc with specific antibodies to enrich for methylated DNA sequences. In combination with next-generation high throughput sequencing, can be used to measure methylation sites within whole genome</td>
<td>–</td>
<td>(158)</td>
</tr>
<tr>
<td>Methylated DNA Binding Domain-Sequencing (MBD-Seq/ MBD-isolated Genomes Sequencing (MiGS))</td>
<td></td>
<td>Specifically bind CpG with MeCP2 to enrich for hypermethylated NDA segments. In combination with next-generation high throughput sequencing, can be used to measure methylation sites within whole genome</td>
<td>–</td>
<td>(159)</td>
</tr>
</tbody>
</table>
technique based on DNA methylation enrichment (142). In previous studies MeDIP-chip and MBD-chip have been widely used (156,160). In combination with next generation sequencing, MeDIP-Seq and MBD-Seq can also be suitable technique based on DNA methylation enrichment (142). In previous studies MeDIP-chip and MBD-chip have been widely used (156,160). In combination with next generation sequencing, MeDIP-Seq and MBD-Seq can also be suitable for global DNA methylation analysis. Down et al. obtained the methylome data from mammals using the MeDIP technique (158). Serre et al. examined DNA methylation in three homogenous tumor cell lines with MBD-seq, and reported methylation regions consistent with previous results as well as hundreds of new methylation regions, indicating the accuracy and sensitivity of this technique (159).

**Prospective goals and future plan**

**Infinium Methylation EPIC Kit**

Improvement in DNA methylation research in HCC relies on the advancement of methylation sequencing techniques. Following the gradual acceptance and usage of the 450k DNA methylation array, a new generation of DNA methylation array, the Illumina Infinium Methylation EPIC BeadChip, has been developed. As confirmed by the ENCODE and FANOM5 projects, this EPICchip can measure the methylation data of around 850,000 CpG sites from whole human genome, which covers >90% sites of 450K beadchip, in addition with 333,265 CpG sites in enhancer regions (155,157,161,162). A better genome wide association study can be performed with the EPIC chip owing to its sufficient probe coverage of CpG islands, gene coding regions, and promoters and enhancers, which provide an advanced technical platform for HCC epigenetic research.

**scTrio-seq**

Current single-cell methylation sequencing techniques include single-cell reduced-representation bisulfite sequencing (scRRBS) (163) and single-cell bisulfite sequencing (scBS) (164). scRRBS has been used in dynamic pattern analysis of DNA methylation during early embryo development (161). The recently introduced single-cell triple omics sequencing (scTrio-seq), first described by Hou et al., can simultaneously analyze the genomic copy number variation (CNV), DNA methylene, and transcriptome in a single mammalian cell, and thus can differentiate the contribution of genomic and epigenomic heterogeneities to the transcriptome heterogeneity. In Hou’s work, the single-cell CNV, RNA expression, and methylation levels measure by scTrio-seq were proved to be consistent with results from scRRBS. Interestingly, though large-scale CNV caused proportional changes of RNA expression levels, the DNA methylation levels were rarely affected (165).

As elucidated by multiple previous studies, gene expression levels are negatively correlated with DNA methylation of gene promoter regions, but positively correlated with DNA methylation of gene bodies (166). Similar results were confirmed by Hou et al.: overall hypomethylation in HCC cells, negative correlation between promoter methylation level and RNA expression, and positive correlation between gene body methylation and RNA expression (165). Moreover, HCC cells were categorized into two subpopulations using scTrio-seq and unsupervised hierarchical clustering analysis. One subpopulation only accounts for a small portion of a given tumor, but may express high levels of cell invasion biomarkers. These subpopulations tend to escape immune surveillance, making them difficult to detect when systemically monitoring for tumors. Filtering and capturing different subpopulations of tumor cells will be an important topic for future cancer research (165).

**Cell-free DNA from plasma**

The majority of circulating cell-free DNA (ccfDNA) originates from apoptotic cells, and very few are from nucleated blood cells. It is therefore reasonable to study circulating tumor DNA via examining plasma (167,168). Several recent studies have explored the potential focusing cell-free nucleic acid in plasma as a biomarker to screen for tumors and evaluate prognoses and therapeutic effects (87,169-171). The use of whole genome methylation analysis to measure methylation of ccfDNA has begun, and may prove to be a novel way for the early detection of HCC (164,172). However, ccfDNA is usually highly fragmented, and its levels are affected by both tumor and non-tumor tissues (167). Moreover, during the early phases of cancer, non-tumor liver tissues may release more ccfDNA than tumor tissues. Therefore, the widespread usage of ccfDNA methylation analysis is greatly limited at this time (102,173).

MCTA-seq developed by Wen et al. achieved the goal of examining genome-scale hypermethylated CpG island in ccfDNA of HCC cells. The target sequence detection rate for the human white blood cell genome can be 93.3% using
MCTA-seq analysis. Equipped with MCTA-seq, Wen et al. analyzed tissue and plasma sample from 151 HCC patients, and found four biomarkers with high specificity—RGS10, ST8SIA6, RUNX2, VIM, the methylation levels of which were positively correlated with the tumor size. Another type of biomarker revealed by MCTA-seq has higher sensitivity than ALT in detecting HCC-caused hepatocyte death. Joint application of these two biomarkers can be used to distinguish samples from HCC patients, cirrhosis patients, or healthy individuals, further improving the diagnosis rate of HCC when combined with current diagnosis criteria (i.e., AFP) (102).

As a non-invasive tumor test, hypermethylated CpG island analysis of ccfDNA has been drawing increasing attention. The hypermethylation of CDKL2, CDKN2A, HIST1H3G, STEAP4, and ZNF154 in ccfDNA can be detected in 37–63% of HCC patients (36). Based on several previous studies, gene methylation levels of CDKN2A/INK4(p16), RASSF1A, GSTP1, and MGMT in HCC cells were 30–70%, 85%, 50–90%, and 40%, respectively (87,92,96,106-108,174-176). Another study recruiting 25,000 participants found that the methylation levels of RASSF1A, CDKN2A, and INK4B (p15) in ccfDNA could predict the progression of HCC in later stages (161). The feasibility of examining DNA methylation in ccfDNA was confirmed by this study, and similar conclusions were also reached in other studies on solid tumors, although a larger-scale cohort study on HCC remains warranted (177-179).

In summary, despite a number of current genome-scale epigenetics studies in HCC, methylation data analysis is limited by various factors: huge gaps between the biological meaning and clinical meaning of the data, high heterogeneity of tumor tissues, and changeable cellular components. Therefore, future whole genome studies should include laser microdissection and high-speed flow cytometry cell sorting to isolate tumor cells from adjacent cells (17,180,181).

Furthermore, the correlation between abnormal DNA methylation and consequent dysfunction should be strictly evaluated. Even though an abundance of data may suggest the important role of candidate DNA methylation or candidate tumor suppressor genes in the pathogenesis and progression of HCC, further evidence should be obtained by manipulating in vivo DNA methylation of specific sites is required, which can more directly disclose the relationship between abnormal methylation and altered gene expression.

Last but not least, much of the current research was focused on or limited on the levels of cell or molecular biology, lacking any connection to existing clinical data, especially DNA methylation marker and clinical application. Therefore, future preclinical studies of tissues, cell, and molecules require more evidence from clinical samples and in vivo experiments. At the same time, the potential biomarkers of abnormal DNA methylation raised from HCC cells should be gradually introduced into clinical application for the purpose of evaluating cancer progression, predicting prognosis, monitoring recurrence, and early diagnosis.

Acknowledgements

We would like to thank Professor Jin Gu of Tsinghua University for his constructive suggestion.

Funding: This work was supported by Beijing Municipal Science & Technology Commission research fund (No. Z171100000417004).

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

determination of major and modified deoxyribonucleosides in DNA. Nucleic Acids Res 1980;8:4763-76.
70. Nagasaka T, Koi M, Kloor M, et al. Mutations in both...
KRAS and BRAF may contribute to the methylator phenotype in colon cancer. Gastroenterology 2008;134:1950-60.


