Research Article

Guangmou Zhang*, Kefeng Zhang, Meng Yuan, Yichen Li, Jiahui Li and Zhiqing Yuan

Revealing distinct DNA methylation patterns in hepatic carcinoma through high-throughput sequencing

https://doi.org/10.1515/tjb-2023-0151
Received August 9, 2023; accepted January 19, 2024;
published online March 13, 2024

Abstract

Objectives: To study the relationship between DNA methylation and tumour development and provide experimental evidence for the personalized diagnosis and treatment of hepatic carcinoma.
Methods: The DNA of hepatic carcinoma tissue (Ca group) and adjacent normal tissue (T group) were extracted using the phenol-chloroform method and then treated with bisulfite. Twenty-five genes including 45 subtypes were amplified by PCR. The PCR products were sequenced via the Illumina 450k methylation array assay. The changes of methylated DNA performance were analysed through principal component analysis (PCA). Cluster analysis was used to evaluate the classification of methylated DNA regions. Haplotype abundance variation was tested for methylation differences. Statistical analysis was performed using the chi-square (χ^2) test or Fisher's exact test.
Results: Sequencing discoveries indicated CG-type methylation pervading all amplicons. However, CHG-type and CHH-type methylations were confined to only four amplicons (or nine subtypes). The methylation ratios of three specific amplicons (DAB2IP, PRDM14-1, Rab31-1) out of 45 amplicon subtypes in the Ca group significantly increased (over 10 %) compared to the T group ($p<0.05$). Nineteen amplicons demonstrated minor distinction (methylation pattern variations between 1 and 10 %), with the remaining 23 amplicons showing only minimal disparities (under 1 %). PCA and cluster analysis unveiled a marked difference in methylation levels between cancerous and healthy tissues ($p<0.05$).
Conclusions: The changes in haplotypes and methylation sites could serve as a biomarker for the clinical diagnosis of hepatic carcinoma. Methylation patterns might play an important role in the occurrence and development of hepatic carcinoma.

Keywords: DNA methylation; high-throughput sequencing; hepatic carcinoma; epigenetic; biomarkers

Introduction

Liver dysfunction can precipitate a variety of diseases and complications. Hepatic carcinoma is one of the most common malignant tumours worldwide. The incidence rate of liver cancer continues to rise, and the survival rate is low because it is almost always diagnosed in the late stage [1]. Early diagnosis is very important for the treatment and prognosis of liver cancer. Recent studies have shown that DNA methylation is deregulated during the development of hepatocellular carcinoma (HCC) by significantly affecting cell differentiation, proliferation, and function [2]. Differential methylation may drive the pathogenesis of liver diseases [3].

Among the various studied epigenetic marks, DNA methylation at CpG dinucleotide cytosines is the most prevalent, and it can be triggered by a variety of environmental factors. The ease of laboratory investigation into DNA methylation makes it suitable for use in epigenome-wide association studies and as a potential epigenetic biomarker for predicting complex diseases [4, 5]. Epigenetic changes in the blood mirror age-related DNA methylation

***Corresponding author: Guangmou Zhang**, College of Life Science and Technology, Xinxiang Medical University, No. 601 Jinsui Avenue, Hongqi District, 453003, Xinxiang, Henan, China, Phone: (+86) 0373-3029887, Fax: (+86) 0373-3029887, E-mail: 951028@xxmu.edu.cn. https://orcid.org/0000-0003-2437-3752
Kefeng Zhang, College of Stomatology, Xinxiang Medical University, Xinxiang, Henan, China. https://orcid.org/0000-0001-5982-5064
Meng Yuan, The Third Affiliated Hospital of Xinxiang Medical University, Xinxiang, Henan, China. https://orcid.org/0000-0002-4036-3709
Yichen Li, Junji College of Xinxiang Medical University, Xinxiang, Henan, China. https://orcid.org/0009-0009-6598-960X
Jiahui Li, School of Cellular & Molecular Medicine, University of Bristol Faculty of Life Sciences, Bristol, UK. https://orcid.org/0009-0003-3727-7468
Zhiqing Yuan, College of Life Science and Technology, Xinxiang Medical University, Xinxiang, Henan, China. https://orcid.org/0000-0001-8324-4841

changes are seen in various human tissues, including the brain, skeletal muscle, adipose tissue, and pancreatic islets [6, 7]. DNA methylation, which occurs in tandem with environmental exposure and other biological events, is frequently associated with an array of diseases, including cancer and autoimmune disorders. Clinical phenotypes exhibit strong correlations with DNA methylation patterns, underscoring their potential as biomarkers for disease diagnosis and treatment guidance [8].

Gaining insights into the genetic and epigenetic profiles of tumours is crucial for understanding carcinogenesis mechanisms and developing therapeutic strategies. With next-generation sequencing, recent advancements in genomic research have enabled one-base-pair resolution analysis of DNA methylation, unveiling more intricate operational mechanisms [9]. These include DNA methylation's role in gene expression suppression or activation, transcription factor binding, splicing, and nucleosome positioning. In this study, we probed the gene expression and DNA methylation status of genes in human liver cancer using The Cancer Genome Atlas database, aiming to uncover common regulatory patterns in human tumorigenesis.

Materials and methods

DNA samples preparation and bisulfite modification

Following the approval of the appropriate ethics committee, we collected a total of 33 samples of post-surgery from a certified hospital. Twenty-six samples were divided into 13 liver cancer tissues (Ca group) and 13 adjacent normal tissues (T group). All tissues were promptly frozen in liquid nitrogen and then stored at −80 °C. We utilised the phenol-chloroform method to extract genomic DNA (gDNA) from the samples. The integrity of the gDNA was confirmed through 1 % agarose gel electrophoresis, which exhibited clear and consistent electrophoresis bands devoid of noticeable degradation or RNA contamination. The concentration and quality of the gDNA were assessed by using the NanoDrop 2000 spectrophotometer (NanoDrop Technologies, DE, USA). The gDNA were treated by bisulfite modification using the QiagenEpiTect Bisulfite Kit (QIAGEN, Hilden, Germany) according to the manufacturer's guidelines. The methylated DNA was purified using EpiTect Spin Columns in accordance with the gene list (Supplemental Table S1). The untreated gDNA were used as control samples.

Methylation-specific PCR (MS-PCR)

We designed MS-PCR primers based on these methylation sites. Primers for both methylated and unmethylated states were generated by Sangon Biotech (Shanghai, China) using MethPrimer software, as detailed in Supplemental Table S2. The PCR reaction mix (total volume: 20 μL) comprised KAPA2G Robust Hot Start ReadyMix, a forward primer, a reverse primer, bisulfite-treated gDNA, and double-distilled water. The target was amplified from the PCR reaction mix after thermal cycling. Concentration detection of equal volumes of PCR product mixtures was conducted following purification with Backman Agtampure XP magnetic beads, and subsequent testing was performed.

DNA methylation sequencing

We integrated specific tag sequences compatible with the Illumina platform into the DNA library using a multi-primer panel for PCR. The library was subsequently subjected to high-throughput sequencing with the Illumina 450k methylation array (Illumina, San Diego, CA, USA), which comprised 485,577 probes that covered 99 % of the 21,231 RefSeq genes, generating FastQ data in dual-end sequencing mode, with each read measuring either 2×150 bp or 2×250 bp. BeadChip images were captured using Illumina iScan. High-quality samples were ensured by utilising control probes for staining, hybridization, extension, and specificity, along with the high bisulfite conversion efficiency (signal intensity>4,000). The raw data exported from GenomeStudio was analysed using Bioconductor and the lumi/methylumi packages, and the transformation from β-values to M-values was executed. BISMARK compared the transformed available data with the transformed target area, yielding the average methylation level of C bases. The final methylation level was calculated using the following formula: Methylation level=100×reads supporting methylation/all reads.

The combination of gDNA bisulfite modification and high-throughput sequencing allowed for the construction of a single base precision methylation map. The most significant advantage of Illumina products is the resolution of a single CpG site. While other methods locate methylation in a certain region, Illumina analysis can accurately determine the methylation level of an individual CpG site.

Statistical analysis

Statistical analysis was conducted using SPSS21 software (SPSS, Chicago, USA). The volumes of the three gene methylation states (methylated, partially methylated, and unmethylated) were compared using the Chi-square test. To investigate the relationship between gene expression/methylation and clinicopathological parameters, either the chi-square (χ^2) test or Fisher's exact test was utilized. The Chi-squared test was used to compare CpG sites with significant genomic distributions. $p<0.05$ indicated a statistically significant difference.

Results

Coverage depth across genes

DNA from the Ca and T groups, post-bisulfite conversion, was utilized to generate PCR amplicons for each gene under investigation. The average length of the amplicons was approximately 279.4 base pairs, with an average coverage of 36.8 CpG dinucleotides per amplicon.

A total of 11,368,632 sequencing reads were accurately aligned. The T group consistently exhibited a higher quantity of reads compared to the Ca group (Table 1). Despite the higher read count for the T samples, discrepancies in the

Table 1: PCR products, number of Illumina reads aligned for each CpG island and the position of genes on chromosomes.

PCR product	Length (bp)	#CpG	T reads	Ca reads	Chr	Start	End
STK11	443	59	24,050	36,279	chr19	1,206,818	1,207,260
PTEN	438	57	5,704	6,147	chr10	87,864,158	87,864,595
MET-1	258	40	23,909	22,260	chr7	116,671,674	116,671,931
MET-2	282	24	775,256	738,751	chr7	116,672,484	116,672,765
NFE2L2-1	237	20	337	187	chr2	177,265,269	177,265,505
NFE2L2-2	233	36	98,322	69,172	chr2	177,264,896	177,265,128
KEAP1-1	215	28	1,533	1,526	chr19	10,504,485	10,504,699
KEAP1-2	181	18	809,359	744,181	chr19	10,504,225	10,504,405
CTNNB1-1	214	22	195,809	175,086	chr3	41,198,649	41,198,862
CTNNB1-2	172	36	56,113	47,358	chr3	41,199,507	41,199,678
FGFR3-1	198	10	61,834	58,389	chr4	1,792,657	1,792,854
FGFR3-2	247	52	28,233	44,704	chr4	1,794,039	1,794,285
FBXW7	203	20	2,958	2,772	chr4	152,536,426	152,536,628
PIK3CA-1	175	30	141,784	131,981	chr3	179,148,226	179,148,400
PIK3CA-2	285	60	32,522	9,634	chr3	179,148,857	179,149,141
MAP2K1-1	214	6	31,852	22,975	chr15	66,386,256	66,386,469
MAP2K1-2	199	17	45,349	22,565	chr15	66,386,793	66,386,991
EGFR-1	208	17	135,042	80,566	chr7	55,019,845	55,020,052
EGFR-2	241	46	50,172	37,268	chr7	55,020,223	55,020,463
ERBB2-1	268	46	39,532	17,938	chr17	39,687,837	39,688,104
ERBB2-2	246	24	25,729	16,240	chr17	39,700,559	39,700,804
ARID1A-1	358	46	37,157	36,329	chr1	26,695,126	26,695,483
ARID1A-2	120	22	14,002	6,387	chr1	26,695,616	26,695,735
CBL-1	650	145	1,173	3,745	chr11	119,205,858	119,206,507
CBL-2	274	56	10,773	9,594	chr11	119,206,596	119,206,869
ARID2-1	346	54	216,472	58,561	chr12	45,727,415	45,727,760
ARID2-2	130	12	262,008	254,750	chr12	45729196	45729325
ERCC2-1	247	22	48,719	31,427	chr19	45,370,780	45,371,026
ERCC2-2	201	28	5,279	1,181	chr19	45,370,342	45,370,542
EP300-1	187	24	34,433	21,379	chr22	4,1,091,758	4,1,091,944
EP300-2	180	44	12,620	8,345	chr22	41,092,590	41,092,769
NOTCH3	259	26	320,072	262,190	chr19	15,201,140	15,201,398
NOTCH2	466	94	7,209	6,077	chr1	120,069,001	120,069,466
ERCC1-1	259	22	536,945	477,879	chr19	45,468,570	45,468,828
ERCC1-2	187	22	103,036	93,346	chr19	45,468,186	45,468,372
RB1	441	89	45,692	18,105	chr13	48,303,789	48,304,229
ABCA3-1	406	2	34,039	27,829	chr16	2,341,829	2,342,234
ABCA3-2	313	40	16,364	24,719	chr16	2,340,796	2,341,108
ABCA3-3	367	43	13,424	16,205	chr16	2,339,941	2,340,307
PRDM14-1	343	26	198,822	41,791	chr8	70,072,522	70,072,864
PRDM14-2	401	39	41,885	31,167	chr8	70,071,595	7,0,071,995
PRDM14-3	442	69	14,448	17,588	chr8	70,070,229	70,070,670
DAB2IP	277	28	1,531	2,272	chr9	121,736,208	121,736,484
RAB31-1	264	16	308,870	235,449	chr18	9,707,714	9,707,977
RAB31-2	298	19	1,399,534	1,126,432	chr18	9,709,128	9,709,425
Total			6,269,906	5,098,726			

The phrase “PCR product” is indicative of tumour-associated genes and subtype. The term “Length” refers to the anticipated length of the PCR product measured in base pairs (bp). The symbol “#CpG” represents the count of potential methylation sites accessible for analysis within each PCR product. The abbreviations “Ca” and “T” are shorthand for hepatic carcinoma tissue and neighbouring normal tissues, respectively. The data outlined in the “Ca Reads” and “T Reads” columns represent the quantities of short Illumina sequencing reads that were aligned to each CpG island’s reference sequences, and these were employed in subsequent examinations. All of these CpG islands were harnessed in the ensuing analyses. The “chr” represent the localization of genes on chromosomes.

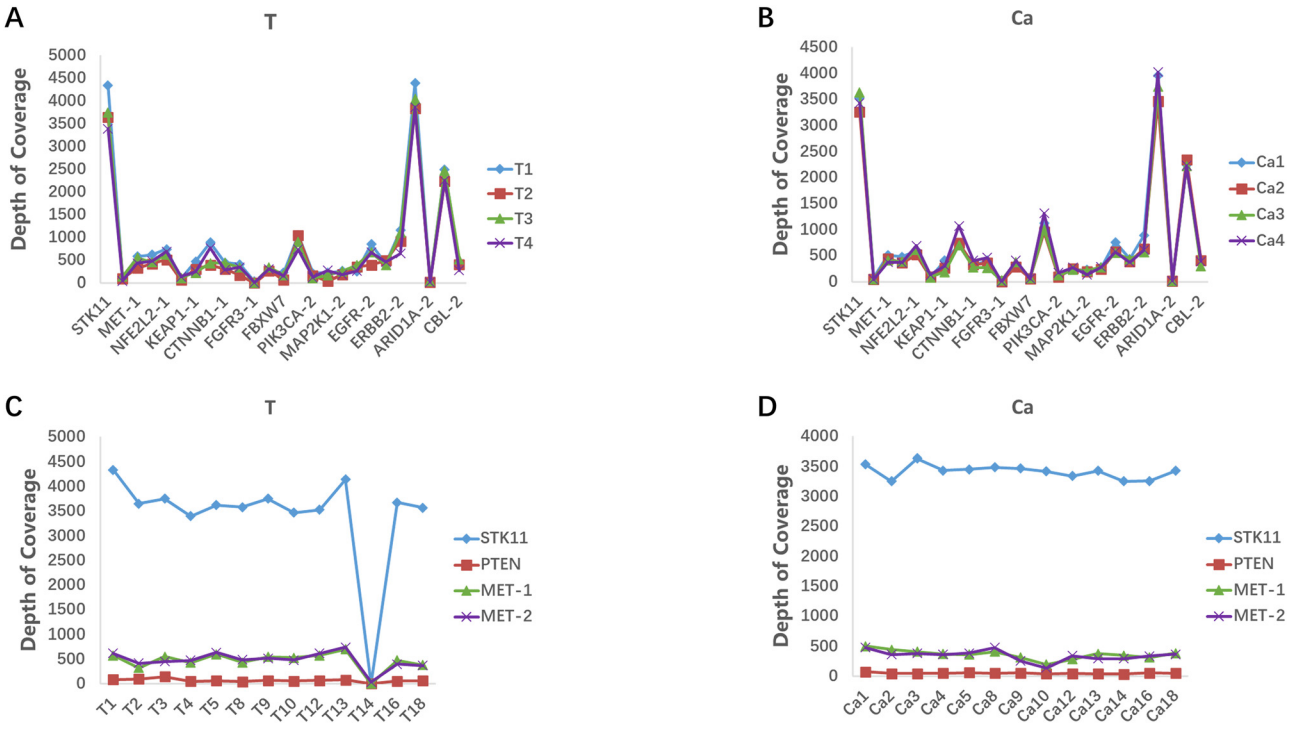

Figure 1: Depth of coverage yielded through high-throughput bisulfite sequencing. (A, B) Exemplary line charts illustrating absolute depth of coverage (Y-axis) vs. genes (X-axis) across the complete amplicon length for eight PCR product samples. These samples include 4 T and 4 Ca samples. The trace patterns are nearly identical for all samples, with minor alterations in coverage depth at each position. (C, D) Exemplary line charts demonstrating absolute depth of coverage (Y-axis) vs. samples (X-axis) along the entire amplicon length for four of the 25 sequenced and aligned PCR products.

read counts for each gene within the T samples were negligible, following a similar pattern, with the sole exception of sample T14 (Figure 1). The distribution of each gene within the Ca group was similarly uniform.

Figure 1 displays the depth of coverage for eight PCR products representing genes in samples. These samples include four T and four Ca samples. The trace patterns were nearly identical for all samples, with minor alterations in coverage depth at each position (Figure 1A and B). There were differences in the absolute coverage depth along the entire length of the amplicon for four out of 25 sequencing and alignment PCR products. One sample (T14) displayed a notable deviation in its read distribution, with prominent gaps in coverage. Consequently, this genomic locus was excluded from further investigation. While there were differences in the absolute depth of coverage among the samples, the contours of the read distribution traces for specific PCR products were nearly identical across both the T and Ca samples (Figure 1C and D). The average coverage depth across all genes within amplicons was identified as 30,548 for the T group and 29,346 for the Ca group, with a maximum of 3,675 at the ARID1A amplicon and a minimum of 65 at the PTEN amplicon.

Methylation levels in hepatic carcinoma

Through the meticulous application of a single nucleotide polymorphism identification technique, we pinpointed each methylation site, quantified the sequencing depth, and assessed the methylation ratio spanning 25 amplicons (or 45 subtypes) for a single sample. Data analysis outcomes revealed variations in methylation ratios among the 25 amplicons between two groups of samples.

Our analysis included three distinct methylation patterns (specifically CG, CGH, and CHH), for which we calculated their respective methylation ratios within the amplicons of both the Ca and T categories. Sequencing discoveries showed CG-type methylation pervading all amplicons. However, CHG and CHH methylation were confined to only four amplicons (or nine subtypes). We observed substantial differences (>10 %) within the methylation ratios of three specific amplicons (DAB2IP, Prdm14-1, Rab31-1) when contrasting Ca and T across the 45 amplicon subtypes (Figure 2A). Compared to the T group, the methylation ratios of DAB2IP, Prdm14-1, and Rab31-1 in the Ca group increased notably ($p<0.05$). Conversely, 19 amplicons demonstrated minor distinction (methylation pattern

Figure 2: Examination of methylation ratio differences. (A) Comparative study of DNA methylation patterns (CG) in Ca group and T group tissues. The chart displays the percentages of methylated cytosines at each potential methylation site for each of the 25 unique PCR products. (B) Comparative study of DNA methylation patterns (CHG) in Ca group and T group tissues. The chart portrays the percentages of methylated cytosines at each potential methylation site for each of the 9 unique PCR products. (C) Comparative study of DNA methylation patterns (CHH) in Ca group and T group tissues. The chart demonstrates the percentages of methylated cytosines at each potential methylation site for each of the 9 unique PCR products. Deep red signifies high methylation levels; yellow indicates low methylation levels.

variations between 1 and 10 %), with the remaining 23 amplicons showing only trivial disparities (<1 %).

The methylation ratios of 20 amplicons in the T group increased relative to the Ca group, in contrast to a down-regulation in 25 amplicons. Notably, among all 25 down-regulated amplicons, the difference in methylation rates between the Ca and T groups was less than 4 %, with most being less than 1 %. Three amplicons were detected with no

significant changes in methylation ratios, while the other six amplicons showed minimal differences. Among these nine amplicons, three had an increased methylation ratio in the Ca group, while six saw a reduction (Figure 2B). In the CHH methylation pattern, two amplicons showed an insignificant difference in methylation ratios between the Ca and T groups, with seven amplicons showing slight variations. Only a single amplicon exhibited an up-regulated methylation ratio in the Ca group, with the rest being down-regulated (Figure 2C).

PCA and cluster examination

PCA was performed on all samples based on their DNA methylation profile to gain deeper insights into the variations in their methylated DNA performance. It was observed that all Ca samples, in contrast to the T samples, exhibited a similar pattern of methylation status alteration, indicating potential physiological transformations in the T group (Figure 3). Further investigation was conducted using cluster analysis grounded on the methylation levels of CpG loci in the amplicon to appraise the classification potential of the identified differentially methylated DNA regions. The samples were bifurcated into two clusters: one consisting of T tissues and the other of Ca tissues (Figure 4). Intriguingly, the methylation level in Ca samples was considerably elevated compared to the T samples ($p<0.05$), suggesting that Ca samples were more prone to methylation than the T samples.

Haplotype and the variation in haplotype abundance

In this study, amplicon was employed as a unit to scrutinize CpG haplotype with high-throughput sequencing facilitating multidimensional methylation analysis. There were 10 specific haplotypes in two Ca samples (Table 2). The abundance of minority haplotypes demonstrated significant variations in the two groups using diverse statistical models. For instance, only one out of the eight haplotypes of the ARID1A_1 sequence manifested a significant disparity by the t-test evaluation ($p<0.05$) (Table 3).

Discussion

Assessment of DNA methylation

DNA methylation analysis emerges as a reliable technique for facilitating early diagnosis, given its ability to detect modifications before the onset of mRNA and protein alterations [10]. The repertoire of techniques for DNA methylation analysis includes sodium bisulfite modification, enzyme digestion, and affinity-based methods, supplemented by next-generation sequencing, locus-specific, gel-based, and array-based approaches [11]. Sodium bisulfite modification is often employed to assess the CpG methylation status, as it transforms unmethylated cytosine to uracil [12, 13]. With its potential for quantitative evaluation, exceptional sensitivity, and compatibility with diverse samples, this method has established itself as a gold standard for DNA methylation analysis [14]. However, the bisulfite treatment process poses restrictions due to DNA denaturation and degradation, thereby limiting the quantity of available DNA. Although this technology comes with certain drawbacks, such as the need for selecting a large number of clones during cloning, complex operations, and high costs, it still provides a more precise detection of genome-wide methylation modifications and offers superior sensitivity and accuracy compared to traditional methylation sequencing techniques. This has led to its widespread application in exploring biological development and disease mechanisms.

With advancements in sequencing technologies, next-generation sequencing facilitates high-throughput analysis and can swiftly generate copious sequence data [15]. Currently, most high-throughput sequencing studies focus on comprehensive epigenome analysis across the entire genome [16, 17], which can lead to substantial costs. However, obtaining accurate data on the proportion of methylated to non-methylated cytosine is a complex endeavour due to the vast size of the mammalian genome, and securing sufficient coverage depth for each CpG site can present a significant challenge. Absolute DNA methylation assays can provide quantitative measurements of DNA methylation levels at the resolution of individual CpG sites.

Epigenetic deviations in hepatocellular carcinoma

The precise molecular pathways leading to liver cancer remain elusive, yet the accumulation of genetic and epigenetic aberrations is widely acknowledged as a key factor in the initiation of the disease [18]. Gene expression alterations, independent of genetic sequence changes, can manifest through various epigenetic mechanisms, such as DNA methylation, microRNAs, non-coding RNAs, histone modifications, and shifts in nucleosome positioning [19–23]. Specifically, anomalous DNA methylation has garnered

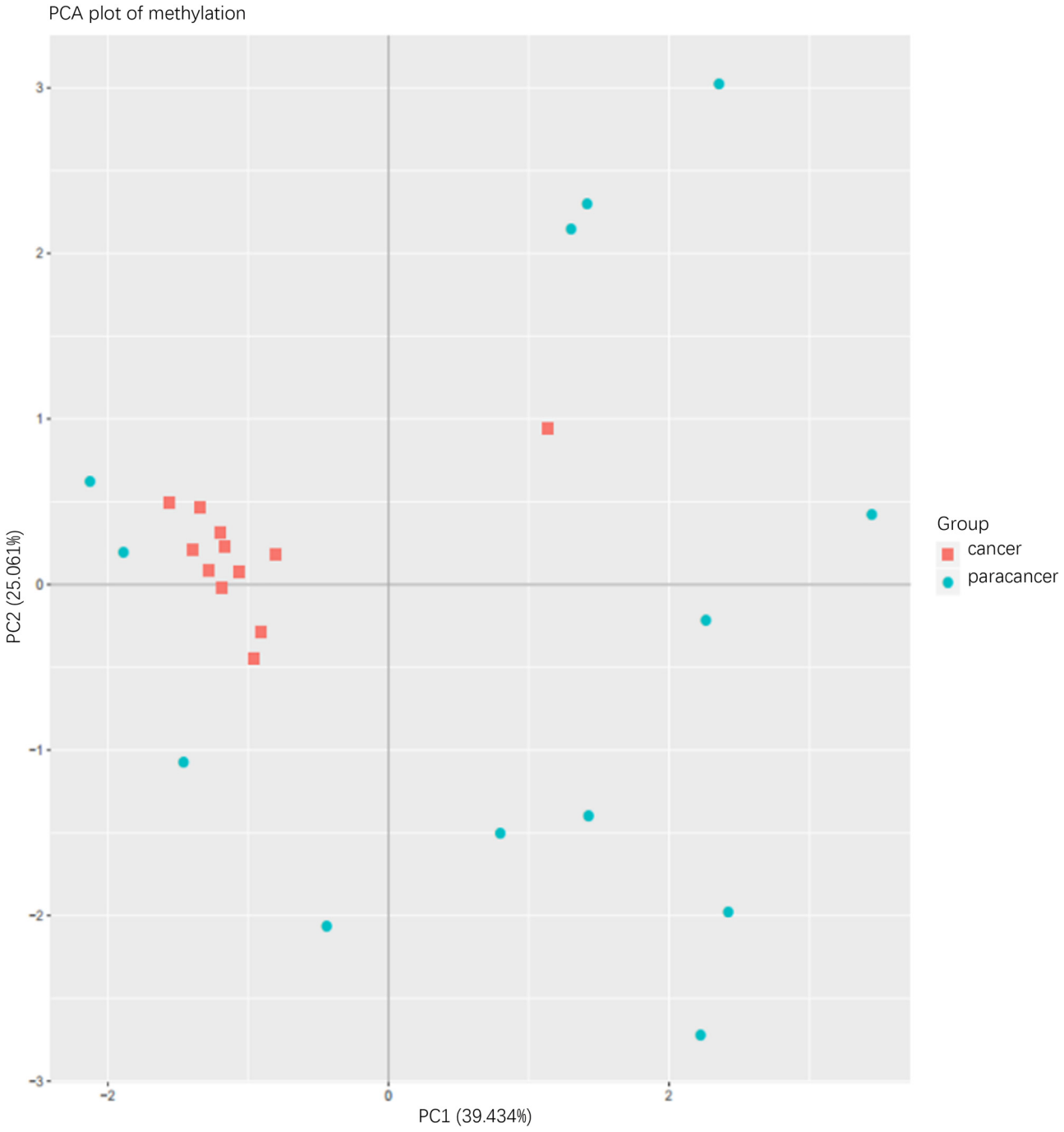

Figure 3: Principle component analysis. PCA visually categorizes sample groups using colour codes. The X and Y axes denote the indicators that most accurately represent the true composition of the samples.

substantial attention in research related to HCC and its precancerous lesions [22].

Studies have drawn connections between several abnormally methylated genes (p16, RASSF1A, GSTP1, p14, CDH1, APC, RUNX3, SOCS1, and WT1) and HCC. These aberrant DNA methylation patterns could potentially serve as valuable biomarkers for predicting and diagnosing HCC [24]. In our research, through detailed research literature, it was found that the expression of 25 genes is closely related to the occurrence and development of liver cancer, but the mechanism is not fully elucidated [25–29]. The sequence specifics and chromosomal locations of the 25 cancer-associated genes are detailed in Table 1. Therefore, we investigated these genes to reveal the potential role of promoter region

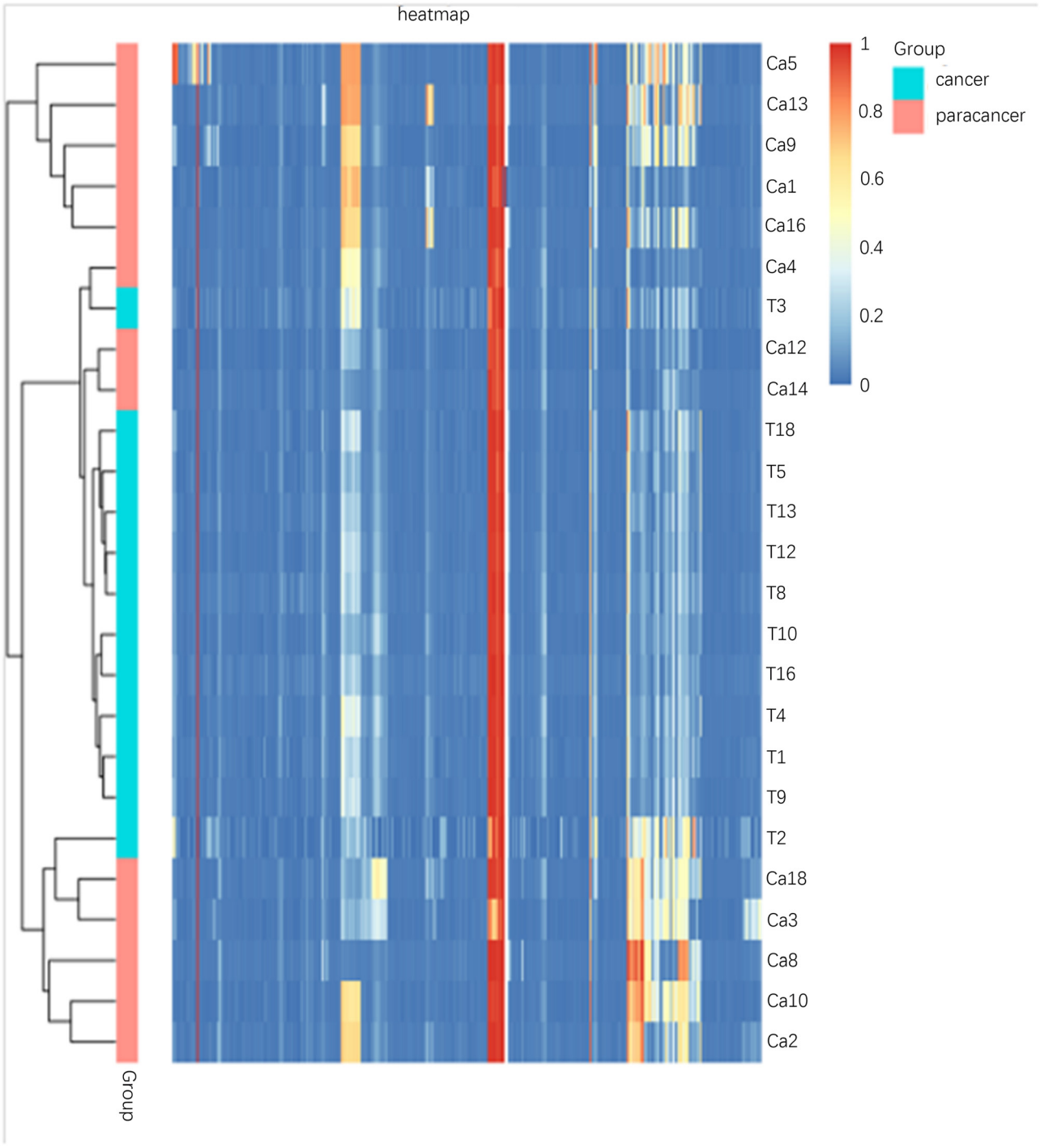

Figure 4: Cluster analysis. Each cell in the heatmap represents the relative methylation level of CpG sites for the corresponding row of samples, with the colour gradient depicting the relative methylation levels of the sites. Blue and red denote low and high methylation levels, respectively. The arrangement order of rows suggests methylation level similarity. Neighbouring rows indicate greater overall similarity in sample methylation levels. The tree on the left systematically exhibits the degree of similarity.

methylation in cancer occurrence and development. The application of next-generation sequencing enables the detection of minute quantities of methylated DNA, suggesting that DNA methylation could possibly act as a molecular marker for HCC across a range of samples, including tissue, plasma, and serum.

Table 2: Haplotype type information in Ca group.

Target	Haplotype	Depth	Ca10	Ca12
ARID1A_1_	ttttttttttt	287,238	0.571429	0.608743
ARID1A_1_	cttttttttttt	37,373	0.099065	0.111211
ARID1A_1_	ttcttttttttt	17,717	0.063523	0.046122
ARID1A_1_	tttcttttttt	11,133	0.021312	0.021255
ARID1A_1_	tttttttcttt	10,476	0.013474	0.016317
ARID1A_1_	ttttttcttttt	8,603	0.020624	0.017702
ARID1A_1_	tttttctttttt	6,722	0.014368	0.009995
ARID1A_1_	ctcttttctttt	1,179	0.000894	0.000000
ARID1A_2_	ttttttttttt	149,929	0.773132	0.822343
ARID1A_2_	ttctttttttt	5,491	0.025589	0.018645

In this scenario, “Target” refers to the amplicon name, whereas “Haplotype” symbolizes the haploid type, assuming an amplification subsequence ATCATXGATCXGCTAXGCTTTAXGCCTATAT, wherein X can be either c (implying methylation) or t (indicating unmethylated status). If a sequence reads: ATCATCGATCTGCTACGCTTTAT, the corresponding methylated haplotype of an amplicon configured to align with the read would be: CTCT. “Depth” is the term used to denote the quantity of sequences that substantiate this haplotype. This chart selectively features 10 haplotypes and 2 samples.

Table 3: Abundance differences in haplotype.

Target	Haplotype	p-Value (T-test)	p-Value (U-test)	p-Value (logistic)
ARID1A_1_	ttttttttttt	0.77824	0.9361618	0.7717829
ARID1A_1_	cttttttttttt	0.94024	0.9786974	0.9370238
ARID1A_1_	ttcttttttttt	0.033058*	0.0768398	0.05901794
ARID1A_1_	tttttttcttt	0.35586	0.8937907	0.4012169
ARID1A_1_	ttttttctttt	0.74136	0.8517193	0.7187338
ARID1A_1_	tttttctttttt	0.90396	0.7283391	0.8952096
ARID1A_1_	ctcttttctttt	0.35548	0.6046477	0.5704589
ARID1A_1_	ttttttttttt	0.22912	0.503319	0.2271915
ARID1A_2_	ttctttttttt	0.23392	0.3203219	0.2982954
ARID1A_2_	tttttttttttc	0.4623	0.8517193	0.4533269

In this context, “Target” alludes to the amplicon name, whereas “Haplotype” connotes the haploid type, taking into account the amplification subsequence ATCATXGATCXGCTAXGCTTTAXGCCTATAT. Here, X can either be ‘c’ (suggesting a methylation modification) or ‘t’ (signifying an absence of methylation modification). Should one encounter the sequence: ATCATCGATCTGCTACGCTTTAT, the corresponding methylated haplotype of an amplicon that aligns with this read would be: CTCT. The parameters “p-Value (T-test)” and “p-Value (U-test)” are derived through the implementation of a T-test and a U-test respectively for model assessment. The logistic regression model was harnessed to compute the “p-Value (Logistic)”. “Depth” reflects the tally of sequences that bolster a specific haplotype. This illustration includes only 10 selectively chosen haplotypes and 2 sample instances. *p<0.05 vs. the T groups.

Analysis of differences in methylation ratio

DNA methylation lacks a definitive threshold, raising questions about the physiological implications of these minor deviations. Nevertheless, even slight changes in DNA methylation patterns can potentially impact both cell and organism function [24, 30]. Therefore, shifts in methylation patterns are noteworthy, as those surpassing 10 % are likely to induce physiological changes. Research has shown that when methylation increases by 10 %, it triggers phenotypic changes in living yellow Aguti mice [31]. Some research indicated that CDKN2A promoter methylation was associated with an enhanced HCC risk and played a crucial role in the process of HCC, with a potential value to being a triage marker for HCC [32].

Given the evident changes in the methylation ratio of DAB2IP, PRDM14-1, and Rab31-1, it was speculated that the methylation abnormalities of these three amplicons may be associated with the growth and development of HCC. The hDAB2IP gene is a potential tumour suppressor gene that has two variants: hDAB2IPA and hDAB2IPB. The hDAB2IPA was the predominant isoform, being expressed in most of the examined tissues. The expression of hDAB2IPA was silenced or down-regulated in liver cancer cell lines. The methylation of the hDAB2IPA promoter in HCC was confirmed in an additional 53 pairs of patient samples. More than 80 % of HCC samples showed hDAB2IPA promoter methylation, compared to 11.5 % in the corresponding adjacent normal tissue [33]. Our results were congruent with previous reports.

PRDM14 is a member of the PR domain-containing (PRDM) family. In adults, PRDM14 has low expression in human tissues. Overexpression of PRDM14 enhances cancer cells growth and reduced cancer cells sensitivity to chemotherapeutic agents. PRDM14 might be a potential therapeutic molecular target for tumour treatment [34]. The study of PRDM14 methylation in liver cancer has not been previously reported.

Rab31 is part of the Ras superfamily of small GTP-binding proteins, which regulates vesicle transport from the Golgi apparatus to early and late endosomes. Rab31 was aberrantly upregulated in HCC tissues compared to adjacent liver tissues. These data suggest that Rab31 promotes HCC cell growth by suppressing cell apoptosis [35]. The study of Rab31 methylation in liver cancer has not been previously reported. Further research is required on the relationship between these genes’ methylation and liver cancer.

Haplotype analysis

A spectrum of single nucleotide polymorphisms has been recognized as potential drivers of unique phenotypic differences and disease predisposition; an observation substantiated by genome-wide association studies (GWASs).

Additionally, it was inferred that haplotype-dependent allele-specific methylation (hap-ASM) could function as a potential determinant of disease vulnerability, although the criteria defining tissue correlation in this context remain ambiguous. Six single nucleotide polymorphisms of TLR4 in the 5′-untranslated region and intron were associated with risk of HCC. In haplotype analysis, one haplotype (GCCCTTAG) of TLR4 was significantly associated with a decrease in the occurrence of HCC [36]. Multivariate logistic regression indicated that with no G-C-T haplotype as a reference, the haplotype of G-C-T from these loci was associated with a lower risk for HCC under the recessive model [37]. Two major susceptible haplotypes were identified in familial HBV-related HCC. The results indicate different genetic susceptibility between familial and sporadic HBV-related HCC [38].

The allele-specific methylation decides the methylation state of haplotypes in genetically associated regions and can establish epigenetic distinctions with potential functional consequences. To contrast the risk and non-risk haplotypes as determined by associations, DNA methylation data can be assessed via haplotype-specific methylation. Interestingly, a discrepancy in prevalence between the Ca and T samples could potentially sway tumour growth. Due to the fact that the polymorphism of a single locus in a gene often cannot reveal its true association with diseases, haplotype analysis for multiple loci has become an effective means of searching for complex disease genes. Our results indicate that haplotype analysis targeting multiple loci is an effective means of identifying the genetic mechanisms underlying the occurrence of liver cancer. The difference in abundance of a few haplotypes between the Ca and T samples might affect tumour growth.

Conclusions

High-throughput sequencing has emerged as an instrumental tool for examining specific DNA methylation patterns in liver cancer tissues and adjacent healthy tissues. Notably, the expense of high-throughput sequencing is on par with standard bisulfite sequencing. In our current investigation, this technique offered an average coverage depth of 30,548 for the healthy tissues (T samples) and 29,346 for the liver cancer tissues (Ca samples). This enabled us to scrutinize the methylation levels at each distinct site in genetically identical samples and ensured a sufficient data quantity for haplotype, PCA, and cluster analysis.

By contrasting the methylation patterns and conducting haplotype analysis, we pinpointed significant methylation sites and differing potential regulatory units between the two groups. The disparities in haplotype and methylation sites between liver cancer and neighbouring healthy tissues could potentially serve as useful references for biomarker research and development in clinical diagnostics.

Furthermore, PCA and cluster analysis revealed a marked difference in methylation levels between cancerous and healthy tissues, indicating distinct characteristics. This observed distinction in methylation patterns between liver cancer and adjacent healthy tissues could have implications for comprehending the progression and development of liver cancer.

Research ethics: Not applicable.
Informed consent: Not applicable.
Author contributions: The authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: The authors state no conflict of interest.
Research funding: This study was supported by the Major science and technology innovation projects in Xinxiang City (ZG1403).
Data availability: The raw data can be obtained on request from the corresponding author.

References

1. Cheishvili D, Wong C, Karim MM, Kibria MG, Jahan N, Das PC, et al. A high-throughput test enables specific detection of hepatocellular carcinoma. Nat Commun 2023;14:3306.
2. Fu S, Debes JD, Boonstra A. DNA methylation markers in the detection of hepatocellular carcinoma. Eur J Cancer 2023;191:112960.
3. Cheung AC, Juran BD, Schlicht EM, McCauley BM, Atkinson EJ, Moore R, et al. DNA methylation profile of liver tissue in end-stage cholestatic liver disease. Epigenomics 2022;14:481–97.
4. Pan Y, Liu G, Zhou F, Su B, Li Y. DNA methylation profiles in cancer diagnosis and therapeutics. Clin Exp Med 2018;18:1–14.
5. Unnikrishnan A, Freeman WM, Jackson J, Wren JD, Porter H, Richardson A. The role of DNA methylation in epigenetics of aging. Pharmacol Ther 2019;195:172–85.
6. Yin Z, Guo X, Qi Y, Li P, Liang S, Xu X, et al. Dietary restriction and rapamycin affect brain aging in mice by attenuating age-related DNA methylation changes. Genes 2022;13:699.
7. Hadad N, Masser DR, Blanco-Berdugo L, Stanford DR, Freeman WM. Early-life DNA methylation profiles are indicative of age-related transcriptome changes. Epigenet Chromatin 2019;12:58.
8. Nishiyama A, Nakanishi M. Navigating the DNA methylation landscape of cancer. Trends Genet 2021;37:1012–27.
9. Law PP, Holland ML. DNA methylation at the crossroads of gene and environment interactions. Essays Biochem 2019;63:717–26.
10. Pu W, Qian F, Liu J, Shao K, Xiao F, Jin Q, et al. Targeted bisulfite sequencing reveals DNA methylation changes in zinc finger family genes associated with KRAS mutated colorectal cancer. Front Cell Dev Biol 2021;9:759813.

11. Sadikovic B, Levy MA, Kerkhof J, Aref-Eshghi E, Schenkel L, Stuart A, et al. Clinical epigenomics: genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders. Genet Med 2021;23:1065–74.
12. Gagliano T, Brancolini C. Epigenetic mechanisms beyond tumour-stroma crosstalk. Cancers 2021;13:914.
13. Ilango S, Paital B, Jayachandran P, Padma PR, Nirmaladevi R. Epigenetic alterations in cancer. Front Biosci 2020;25:1058–109.
14. Kashyap MP, Sinha R, Mukhtar MS, Athar M. Epigenetic regulation in the pathogenesis of non-melanoma skin cancer. Semin Cancer Biol 2022;83:36–56.
15. Benson HM, Lindsay C, Ward AHeywood SPrather RSIsom SC. Targeted DNA methylation analysis by high throughput sequencing in porcine peri-attachment embryos. J Reprod Develop 2013;59:314–20.
16. Nair SV, Madhulaxmi, Thomas G, Ankathil R. Next-generation sequencing in cancer. J Maxillofac Oral Surg 2021;20:340–4.
17. Zhu T, Liu J, Beck S, Pan S, Capper D, Lechner M, et al. A pan-tissue DNA methylation atlas enables in silico decomposition of human tissue at cell-type resolution. Nat Methods 2022;19:296–306.
18. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr., Kinzler KW. Cancer genome and scapes. Science 2013;339:1546–58.
19. Pan J, Li D, Fan X, Cheng J, Jin S, Chen P, et al. Aberrant DNA methylation patterns of eleted in liver cancer 1 isoforms in hepatocellular carcinoma. DNA Cell Biol 2023;42:140–50.
20. Mohr R, Özdirik B, Lambrecht J, Demir M, Eschrich J, Geisler L, et al. From liver cirrhosis to cancer: the role of micro-RNAs in hepatocarcinogenesis. Int J Mol Sci 2021;22:1492.
21. Shen H, Liu B, Xu J, Zhang B, Wang Y, Shi L, et al. Circular RNAs: characteristics, biogenesis, mechanisms and functions in liver cancer. J Hematol Oncol 2021;14:134.
22. Luo B, Ma F, Liu H, Hu J, Rao L, Liu C, et al. Cell-free DNA methylation markers for differential diagnosis of hepatocellular carcinoma. BMC Med 2022;20:8.
23. Yu F, Shen X, Fan L, Yu Z. Analysis of histone modifications at human ribosomal DNA in liver cancer cell. Sci Rep 2015;5:18100.
24. Zhang C, Li J, Huang T, Duan S, Dai D, Jiang D, et al. Meta-analysis of DNA methylation biomarkers in hepatocellular carcinoma. Oncotarget 2016;7:81255–67.
25. Morishita A, Iwama H, Fujihara S, Watanabe M, Fujita K, Tadokoro T, et al. Targeted sequencing of cancer-associated genes in hepatocellular carcinoma using next-generation sequencing. Oncol Lett 2018;15:528–32.
26. Liu L, Long H, Wu Y, Li H, Dong L, Zhong JL, et al. HRD1-mediated PTEN degradation promotes cell proliferation and hepatocellular carcinoma progression. Cell Signal 2018;50:90–9.
27. Zhang R, Gao X, Zuo J, Hu B, Yang J, Zhao J, et al. STMN1 upregulation mediates hepatocellular carcinoma and hepatic stellate cell crosstalk to aggravate cancer by triggering the MET pathway. Cancer Sci 2020;111:406–17.
28. Vilarinho S, Erson-Omay EZ, Harmanci AS, Morotti R, Carrion-Grant G, Baranoski J, et al. Paediatric hepatocellular carcinoma due to somatic CTNNB1 and NFE2L2 mutations in the setting of inherited bi-allelic ABCB11 mutations. J Hepatol 2014;61:1178–83.
29. Liu Y, Tao S, Liao L, Li Y, Li H, Li Z, et al. TRIM25 promotes the cell survival and growth of hepatocellular carcinoma through targeting Keap1-Nrf2 pathway. Nat Commun 2020;11:348.
30. Cao P, Li H, Zuo Y, Nashun B. Characterization of DNA methylation patterns and mining of epigenetic markers during genomic reprogramming in SCNT embryos. Front Cell Dev Biol 2020;8:570107.
31. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol 2003;23:5293–300.
32. Zhou Y, Wang X-B, Qiu X-P, Zhang S, Wang C, Zheng F. CDKN2A promoter methylation and hepatocellular carcinoma risk: a meta-analysis. Clin Res Hepatol Gastroenterol 2018;42:529–41.
33. Qiu GH, Xie H, Wheelhouse N, Harrison D, Chen GG, Salto-Tellez M, et al. Differential expression of hDAB2IPA and hDAB2IPB in normal tissues and promoter methylation of hDAB2IPA in hepatocellular carcinoma. J Hepatol 2007;46:655–63.
34. Ou M, Li S, Tang L. PRDM14: a potential target for cancer therapy. Curr Cancer Drug Targets 2018;18:945–56.
35. Sui Y, Zheng X, Zhao D. Rab31 promoted hepatocellular carcinoma (HCC) progression via inhibition of cell apoptosis induced by PI3K/AKT/Bcl-2/BAX pathway. Tumour Biol 2015;36:8661–70.
36. Shi M, Xu X, Chen H, Shen B, Deng X, Xie J, et al. Single nucleotide polymorphisms of toll-like receptor 4 decrease the risk of development of hepatocellular carcinoma. PLoS One 2011;6:e19466.
37. Li Y, Wang J, Jiang F, Lin W, Meng W. Association of polymorphisms in survivin gene with the risk of hepatocellular carcinoma in Chinese Han population: a case control study. BMC Med Genet 2012;13:1.
38. Lin Y-Y, Yu M-W, Lin S-M, Lee S-D, Chen C-L, Chen D-S, et al. Genome-wide association analysis identifies a GLUL haplotype for familial hepatitis B virus-related hepatocellular carcinoma. Cancer 2017;123:3966–76.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/tjb-2023-0151).