EPIGENOMICS
2025, VOL. 17, NO. 10, 661–674
https://doi.org/10.1080/17501911.2025.2518919

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Methylation aberrations and genomic instability synergistically drive the evolution of intrahepatic cholangiocarcinoma

Guanghao Li[a,b,c,d], Youhuang Bai[a,b,c,d], Feng Tao[a,b,c,d,e], Tingting Hu[a,b,c,d], Ting Wang[f], Yong Zeng[g] and Deqiang Sun[a,b,c,d,e]

[a]Department of Cardiology, The Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China; [b]State Key Laboratory of Transvascular Implantation Devices, Hangzhou, China; [c]Heart Regeneration and Repair Key Laboratory of Zhejiang Province, Hangzhou, China; [d]Transvascular Implantation Devices Research Institute, Hangzhou, China; [e]Binjiang Institute of Zhejiang University, Hangzhou, China; [f]Department of Radiation Oncology, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, China; [g]Division of Liver Surgery, Department of General Surgery and Laboratory of Liver Surgery, and State Key Laboratory of Biotherapy and Collaborative Innovation Center of Biotherapy, West China Hospital, Sichuan University, Chengdu, China

ABSTRACT

Aims & methods: DNA methylation and genomic instability are critical drivers of cancer initiation and malignant progression. However, the roles of methylation aberrations and genomic instability in malignant progression have not been thoroughly investigated in intrahepatic cholangiocarcinoma (ICC). To address this, we identified differentially methylated regions (DMRs) and somatic copy number alterations (SCNAs) from 341 ICC samples across various stages.

Results: Our findings revealed that stages IAIB, II, IIIA, and IIIB exhibited comparable methylation changes, whereas stage IV ICC showed a pronounced accumulation of stage-specific methylation alterations. Leveraging these findings, we developed a classification model that effectively distinguished stage IV ICC from earlier stages with high accuracy using 15 DMRs. Furthermore, stage IV ICC exhibited slightly higher genomic instability, including an elevated aneuploidy score and a greater proportion of focal amplifications. We also observed a positive correlation between SCNA burden and DNA methylation entropy in the promoter, gene body, and CpG island regions, with the gene body of *MDM2* serving as a notable example.

Conclusions: These findings highlight the potential of DNA methylation as a biomarker for metastasis diagnosis and the interplay between local genomic instability and aberrant methylation, emphasizing their synergistic roles in driving the evolutionary trajectory of ICC.

PLAIN LANGUAGE SUMMARY

Intrahepatic cholangiocarcinoma (ICC) is a rare but aggressive type of liver cancer with a poor prognosis. In this study, we analyzed DNA methylation and genetic alterations in 341 ICC samples to understand how these molecular changes contribute to cancer progression. DNA methylation is a chemical modification to DNA that can switch genes on or off without changing the DNA sequence itself. Genomic instability, which refers to a high frequency of mutations across the genome, is a common feature of cancer. We found that abnormal DNA methylation occurs more frequently in advanced-stage ICC and can help distinguish advanced from early-stage disease, suggesting it could serve as a potential biomarker for cancer spread. We also observed greater genomic instability in advanced-stage ICC. Notably, the abnormal DNA methylation patterns were closely linked to genomic instability – especially in key cancer-related genes – potentially working together to disrupt normal gene function. Our findings provide new insights into how epigenetic and genetic alterations drive the development and progression of ICC and may support future efforts to develop new diagnostic tools or targeted treatments.

ARTICLE HISTORY
Received 6 March 2025
Accepted 9 June 2025

KEYWORDS
Intrahepatic cholangiocarcinoma; DNA methylation; methylation entropy; genomic instability; somatic copy number alterations; malignant progression; diagnosis of metastasis

1. Introduction

DNA methylation aberration and genomic instability are the hallmarks of cancer [1], playing crucial roles in tumorigenesis and malignant progression [2]. Both types of alterations are genome-wide in nature. Cancer genomes are hypomethylated at the whole-genome level and hypermethylated in regulatory regions and promoters [3]. Hypomethylation and hypermethylation are associated with activation of oncogenes and deactivation of tumor suppressor genes, respectively [4,5]. Genomic instability, a ubiquitous feature of human cancers,

CONTACT Deqiang Sun deqiangs@zju.edu.cn Department of Cardiology, The Second Affiliated Hospital, School of Medicine, Zhejiang University, 88 Jiefang Road, Hangzhou 310009, China

Supplemental data for this article can be accessed online at https://doi.org/10.1080/17501911.2025.2518919

Article highlights

- Exploring the DNA methylation aberrations and genomic stability across stages IAIB, II, IIIA, IIIB and IV of ICC samples.
- Stage IV ICC accumulated a pronounced of stage-specific methylation alterations, accompanied by an enrichment of cancer-associated transcription factors.
- Developing a robust classification model, which distinguished stage IV ICC from earlier stages with high accuracy using a panel of 15 DMRs.
- Stage IV ICC exhibited slightly higher genomic instability, including an elevated aneuploidy score and a greater proportion of focal amplifications.
- Chromosome arm aneuploidy was associated with the prognosis of ICC patients, including 6q related with better prognosis, and 17p correlated with worse outcome.
- SCNA burden and DNA methylation entropy was positively associated in the promoter, gene body, and CpG island regions, with the gene body of *MDM2* serving as a notable example.

manifests in multiple forms, including point mutations, chromosomal rearrangements, somatic copy number alterations (SCNAs), microsatellite instability and aneuploidy. These alterations contribute to oncogenesis by activating oncogenes, inactivating tumor suppressor genes, fostering the development of aggressive phenotypes, and equipping cancer cells with the adaptability to evade immune surveillance and resist therapeutic interventions. Given their critical roles in cancer evolution, profiling DNA methylation aberrations and genomic instability offers valuable insights into the underlying mechanisms driving tumor progression and adaptation.

DNA methylation disorder plays a role analogous to genomic instability, enhancing the evolutionary potentials of cancer cells [6]. Beyond tumor initiation, methylation aberrations are key drivers of tumor invasion, metastasis, and recurrence in various cancers, such as large B-cell lymphoma relapse, lung adenocarcinoma progression, and advanced prostate cancer [7–10]. Genomic instability accelerates cancer evolution and malignant progression, with aneuploidy and focal amplifications as two critical contributors. Aneuploidy exhibits context-dependent effects, promoting or hindering cancer evolution based on the genomic landscape and tumor type [11–13]. For example, 1p/19q codeletion predicts favorable prognosis in gliomas, while chromosome 7 gain and chromosome 10 loss indicate poor outcomes [14]. Focal amplifications, particularly extrachromosomal circular DNA (ecDNA), significantly increase gene copy number heterogeneity, intensifying oncogenic signaling and tumor malignancy [15]. Together, DNA methylation disorder and genomic instability drive tumor evolution, fostering adaptability and progression.

Intrahepatic cholangiocarcinoma (ICC), the second most common primary liver cancer, accounts for approximately 15% of cases. According to the 8th edition of the American Joint Committee on Cancer (AJCC) Staging Manual, ICC is classified into four stages: I (IA, IB), II, III (IIIA, IIIB), and IV. This classification is based on tumor size, number of regional lymph nodes, and occurrence of distant metastasis [16,17]. The majority of patients with ICC are diagnosed at an advanced stage [18], with limited therapeutic options and a five-year survival rate of only 7–20% [19]. Stage IV ICC, characterized by distant metastasis, represents the most aggressive form with the worst prognosis, underscoring the importance of understanding the molecular mechanisms underlying its malignant progression. Profiling DNA methylation aberrations and genomic instability across ICC stages provides crucial insights into tumor evolution and potential biomarkers of malignancy. Previous studies have characterized the genomic and epigenomic landscapes of ICC, leading to molecular subtyping based on canonical alterations and associations with prognosis, which may also serve as potential therapeutic targets [20–27]. Despite these advances, a systematic analysis of methylation aberrations, aneuploidy, and ecDNA across ICC stages is lacking. Additionally, the relationship between methylation and SCNAs remains unexplored, warranting further investigation to uncover their joint contributions to ICC progression.

In this study, we investigated DNA methylation disorder and genomic instability in ICC using whole-genome bisulfite sequencing (WGBS) data from our previous studies [24,25]. Stage-specific and shared differentially methylated aberrations were identified across different stages of ICC. Additionally, SCNA at the chromosome arm and gene levels, as well as their relationship with DNA methylation, were analyzed. The interplay between DNA methylation and SCNAs provides insights into the evolutionary dynamics of ICC.

2. Methods

2.1. Datasets

The WGBS data of 341 ICC and nine intrahepatic bile duct (IBD) samples were obtained in previous study [24,25], download from Genome Sequence Archive (GSA-Human: HRA004700). The human reference genome, hg38, was obtained from the University of California Santa Cruz (UCSC). Genome annotation tables for hg38, including genes, CpG islands (CGIs), and repetitive elements, were downloaded from the UCSC Table Browser. The promoter was defined as the region 1000 bp upstream and 1000 bp downstream of the transcription start site (TSS). Genes were defined as regions spanning from the TSS to transcription end site. The 3′ untranslated region (3′UTR), 5′UTR, exon, intron, and intergenic regions were defined according to the NCBI RefSeq gene table. The CGI shore was defined as the 2 kb flanking region of the CGI, and the CGI shelf was defined as the 2 kb flanking region of the CGI shore. Repetitive elements, including long interspersed nuclear element (LINEs), short interspersed nuclear elements (SINEs), long terminal repeat (LTRs), and satellites, were defined according to the RepeatMasker table. Assay for transposase-accessible chromatin (ATAC) peak regions were obtained from TCGA database [28].

2.2. WGBS data analysis

Low-quality sequencing reads were removed and adapters and poor-quality bases were trimmed using fastp [29]. Reads with lengths of < 36 bp were filtered. The remaining clean reads were aligned to the human reference genome (UCSC hg38) using BSMAP [30]. The methylation ratios of the CpG

sites were calculated using the MCALL module of MOABS [31], a statistical model-based analysis method for bisulfite sequencing data. CpG sites with fewer than three reads were excluded from further analysis. The mean methylation ratios of genomic elements, such as CGIs, were calculated using bedtools [32]. Figures with lines of the average methylation ratios of CGI, CGI exon, 3'UTR, 5'UTR, gene, exon, intron, and promoter were generated using deeptools [33].

2.3. Identification of differentially methylated cytosines (DMCs) and regions (DMRs)

DMCs and DMRs were identified using the MCOMP module of MOABS [31]. Using IBD samples as normal controls, we identified the DMCs/DMRs of ICC by integrating all ICC samples. The DMCs/DMRs in stages IAIB, II, IIIA, IIIB, and IV were identified using the same methods. The DMCs/DMRs of the ICC and each stage were annotated using the R package ChIPseeker [34] to obtain the located genes.

2.4. Pathway enrichment analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was performed for genes annotated with DMRs of ICC and each stage using the R package clusterProfiler [35], in which Fisher's exact test was performed, and the *p* value was adjusted using the Benjamini – Hochberg method. Genes with hyper-DMRs and hypo-DMRs were enriched separately. The top 10 most significantly enriched pathways are shown.

2.5. TF motif analysis

The de novo and known motifs in ICC and at each stage were identified using HOMER [36], which detects motifs with differential enrichment between hyper/hypo-DMRs and background sequences using a cumulative hypergeometric distribution.

2.6. Validation of DMR and TF using TCGA data

To validate the identified DMRs and TF motifs enriched in ICC, we downloaded Illumina Infinium HumanMethylation450 BeadChip (450K) array data from TCGA using the R package TCGAbiolinks [37]. In total, we obtained data from 33 ICC samples and 8 normal samples. The methylation levels of the identified ICC DMRs in the 450K array were calculated as the average beta values of CpGs located in the same DMR using bedtools [32].

To investigate the correlation between DNA methylation alterations and changes in gene expression, we downloaded RNA-seq data from TCGA-CHOL, obtaining the TPM (transcripts per million) value of each gene. To integrate with our WGBS data analysis, we examined the relationship between DNA methylation and gene expression using the DMRs identified in this study and validated in TCGA-CHOL 450K array data. We filtered these DMRs using two criteria: (1) an absolute methylation difference between tumor and normal samples greater than 0.2, and (2) an absolute log fold change of TPM value between tumor and normal samples greater than 1. In total, 910 genes met these criteria.

2.7. Comparison of DMC/DMRs at different stages of ICC

The DMRs of different ICC stages were compared using the R package ChIPpeakAnno [38] to obtain shared DMRs between two or more ICC stages and stage-specific DMRs. The methylation ratios of the DMRs were calculated as the average methylation levels of CpGs in the same DMR using bedtools. Comparisons of DMCs between different stages were performed according to their genomic coordinates, and the results were plotted using a Venn diagram [39].

2.8. Distinguishing stage IV ICC from other stages

The 341 ICC samples were divided into discovery ($n = 168$) and validation ($n = 173$) cohorts according to age and stage. Both cohorts had eight stage IV patients to minimize bias caused by the small sample size of patients with stage IV disease. We trained the classification model to distinguish stage IV from the other stages in the discovery cohort and validated it in the validation cohort.

Previously identified stage-specific hyper- and hypo-DMRs were candidate variables for the classification model. This is because of the high dimensionality of DMRs relative to the number of available ICC samples. We first filtered these DMRs by comparing the methylation ratios between stage IV and the other stages. DMRs with a *t*-test P-value <0.001 and absolute DNA methylation ratio difference > 0.15 were selected. In total, 428 DMRs were retained. We then constructed a logistic regression model using the least absolute shrinkage and selection operator (LASSO) penalty for classification. To obtain the penalty value, we employed five-fold cross-validation in the discovery cohort, where four of the five folds were used to construct the model, and the fifth held-out fold was used to evaluate model performance. This cross-validation process was repeated 10 times, and the optimized LASSO penalty was 0.0156. Finally, we obtained a classification model with 15 DMRs as the variables. The performance of the model in the discovery and validation cohorts was evaluated using the AUC values.

2.9. Identification of chromosome arm aneuploidy

To identify chromosome arm aneuploidy, genomic segments were generated using ichorCNA [40] with a large window size (500kb). For each chromosome arm, copy number alteration (CNA) was obtained through calculating the proportion of gain or loss. The proportion of copy number gain, *p*, was the total length of segments with CN > 2, divided by the whole arm length, and copy number loss, *q*, was calculated in the sample way, using the segments with CN < 2. The chromosome arm was gained if $p > 0.7$, loss if $q > 0.7$, neutral if neither. Aneuploidy score of a sample was the number of gained and loss chromosome arm. The range of aneuploidy score is from 0 to 38, as sex chromosomes and short arms of chromosome 13, 14, 15, 19, 21 and 22 were excluded.

2.10. Identification of focal amplification

AmpliconArchitect [41] is a computational tool to reconstruct the structure of focally amplified regions using whole-genome sequencing (WGS) data. WGBS data containing the sequencing information of the whole genome, thus AmpliconArchitect is also applicable. All samples were analyzed using AmpliconArchitect and AmpliconClassifier [42] to detect the focal amplifications. Specifically, copy number alterations were estimated with CNVkit, and segments with copy number ≥ 4 were selected as seed regions, which were used to search discordant read pairs to identify genomic structural rearrangement and construct breakpoint graph. Amplicons were classified into four categories: (1) linear amplification; (2) heavily rearranged amplification; (3) breakage-fusion-bridge (BFB) amplification; (4) ecDNA amplification.

2.11. Linear model between DNA methylation and SCNA

SCNAs of each ICC were obtained from the previous focal amplification analysis. DNA methylation ratio of each CpG site was obtained from the previous WGBS data analysis. DNA methylation entropy of each CpG site was calculate using mHapTk [43]. We calculated the associations between DNA methylation level/entropy and SCNA for different genomic regions, including promoter, gene, CGI. For a specific region i, we denoted the average methylation ratio as M_i, the average SCNA as CN_i. The linear regression model was: $M_i = \beta_0 + \beta_1 CN_i + \beta_2 Age_i + \varepsilon_i$, age is the covariate. We denoted the average methylation entropy for region i as E_i, then the linear model is: $E_i = \beta_0 + \beta_1 CN_i + \beta_2 Age_i + \varepsilon_i$. We calculated the p-value for the hypothesis testing of $\beta_1 = 0$ vs. $\beta_1 \neq 0$. The regions with p-value smaller than 10^{-10} were shown.

2.12. Statistics

The proportion test was used to compare the proportion differences of specific hyper- and hypo-DMRs at different ICC stages. The Kruskal – Wallis test was performed to compare the differences in methylation ratios among the five ICC stages. Survival curves were estimated using the Kaplan – Meier method, and the log-rank test was performed to assess the statistical significance between high and low scores in the classification model. All statistical analyses were performed using R software (v4.3.2).

3. Results

3.1. Landscape of DNA methylation profiles for ICC samples

To explore the landscape of DNA methylation and genomic instability profiles in ICC, we analyzed WGBS data of 341 ICC and nine IBD samples from previous studies [24,25] (Figure 1(a)). Principal component analysis revealed no significant batch effects among the 341 ICC samples (Figure S1). Compared to IBD, ICC exhibited greater variability in DNA methylation profiles. The average methylation ratio of all CpGs in the whole genome for IBDs was approximately 0.77, with minimal fluctuations. In contrast, ICC displayed broader variations, with the first and third quartiles of methylation ratios at 0.73 and 0.77, respectively, suggesting the accumulation of DNA methylation aberrations during tumorigenesis. Notably, this variation was consistent across all ICC stages, with no significant differences observed (Figure 1(b), Figure S1, Kruskal – Wallis test, $p = 0.45$). Further analysis revealed distinct patterns of methylation in ICC compared to IBD. DNA hypermethylation in ICC was enriched in gene-regulatory regions, such as promoters and CGIs, whereas hypomethylation was predominantly found in other genomic regions, such as LINE, SINE and LTR (Figure 1(c), Figure S1, S2). Hypermethylation of gene-regulatory regions is associated with the deactivation of tumor suppressor genes, and hypomethylation of LINE is related with genomic instability and poor prognosis in cancer. Aberrations of DNA methylation are the important driver for the tumorigenesis of ICC.

We identified DMC and DMRs in ICC compared to IBD samples. DMRs and DMCs were categorized into hypermethylated (hyper-DMRs/DMCs) and hypomethylated (hypo-DMRs /DMCs) groups based on their methylation status in ICC. In total 30,012 hyper-DMRs and 38,928 hypo-DMRs were detected. The distribution of hyper- and hypo-DMRs varied across genomic regions. Hyper-DMRs were enriched in CGIs, exons, and promoters, whereas hypo-DMRs were predominantly found in LINEs, SINEs, LTRs and ATAC peaks (Figure 1(d)). A similar distribution pattern was observed for DMCs across genomic regions (Figure S3). These DMRs were validated using the Illumina 450K array data of ICC from TCGA, which revealed that the methylation ratios differed between normal and ICC samples (Figure S4). By integrating TCGA RNA-seq and 450K array data, we validated a negative correlation ($r = -0.16$) between DNA methylation differences and gene expression changes across these genes, with an even stronger negative correlation observed specifically in promoter regions ($r = -0.3$) (Figure S4).

3.2. Stage IV ICC accumulated more DNA methylation aberrations

To investigate the effect of DNA methylation on the malignant progression of ICC, the ICC samples were divided into five stages according to the 8th edition of the AJCC Staging Manual: IAIB ($n = 64$), II ($n = 43$), IIIA ($n = 128$), IIIB ($n = 90$), and IV ($n = 16$). The correlation of the CpG methylation ratio between stage IV and the other stages was weaker than that within the other four stages, indicating that stage IV ICC is characterized by a particularly distinctive DNA methylation disorder (Figure 2(a)).

Furthermore, DMCs/DMRs were identified for each stage using IBD as a control. This analysis revealed that the number of hyper-DMCs/DMRs in the five stages was similar, whereas the number of hypo-DMCs/DMRs in stage IV was noticeably higher than that in the other four stages (Figure S5). TF motifs enriched in DMRs were highly consistent from stages IAIB to IIIB, including NF1, NF1-halfsite, SCL, TWIST2, and BHLHA15 in hyper-DMRs and AP-1, ATF2, BATF, FOS, and JUNB in hypo-DMRs (Figure 2(b,c), Supplementary Table S1). This is

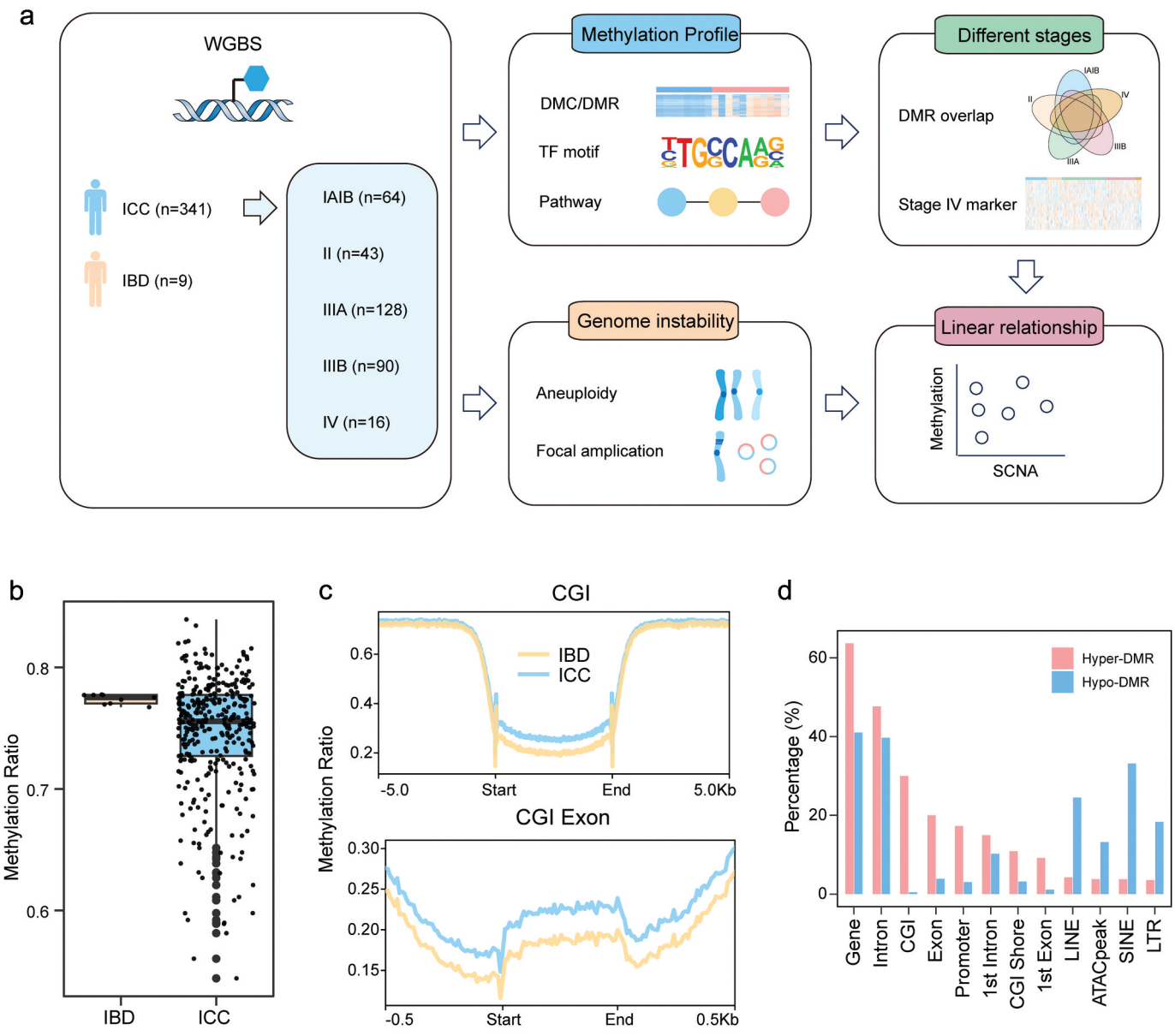

Figure 1. DNA methylation profiles of ICC and normal samples. (a) schematic representation of the study. Intrahepatic cholangiocarcinoma (ICC, $n = 341$) and intrahepatic bile duct (IBD, $n = 9$), were collected, and whole-genome bisulfite sequencing (WGBS) was performed. Differentially methylated CpGs (DMCs) and regions (DMRs) of ICC were identified. Oncogenic pathway and transcription factor (TF) motifs were enriched using DMCs/DMRs. DMCs/DMRs at different ICC stages were also detected and stage IV-specific markers were identified. Two forms of genomic instability, aneuploidy and focal amplification were identified. The relationship between somatic copy number alterations (SCNA) and methylation was explored using linear model (b) Average methylation ratios at the whole-genome level for ICC and IBD samples. (c) Line plot of methylation profiles in the CGI and CGI exon regions for ICC and IBD samples. (d) Overlap of hyper- and hypo-DMRs with functional genomic regions. CGI, CpG island; LINE, long interspersed nuclear element; SINE, short interspersed nuclear element; LTR, long terminal repeat.

consistent with the TFs shown in Figure 2(c). However, stage IV DMRs were enriched with more specific TF motifs, including TEAD1, FOXL2, and FOXF1 in hyper-DMRs, and SPIB, PU.1, ELF5, ELF4, and CEBP in hypo-DMRs. Many of these, such as TEAD1, CEBP, and ELF5, have been previously reported to be associated with tumor metastasis [44–46]. Additionally, several canonical oncogenic pathways were identified as enriched in the five stages using genes with hyper- and hypo-DMRs, indicating a common driving force for the different stages of ICC.

We further compared the DMCs/DMRs identified in the five stages and identified shared and specific hyper- and hypo-DMCs/DMRs among the different stages (Figures 2(d,e), Figure S6). For each stage, the proportion of specific hypo-DMRs was significantly higher than that of specific hyper-DMRs ($p < 0.001$), particularly at stage IV (Figure 2(f)). The numbers of stage-specific hyper-DMRs from stages IAIB to IV were 913, 1,413, 2,225, 410 and 4,438, respectively (Figure 2(d)). For hypo-DMRs, the corresponding values were 4,105, 8,913, 3,840, 1,125, and 32,120 (Figure 2(e)). Both the absolute numbers and proportions of stage-specific of hyper- and hypo-DMRs were markedly higher in stage IV than in the other stages. Notably, different from hyper-DMRs, the majority of stage IV specific hypo-DMRs were enriched in intergenic and intron regions. This pattern may reflect enhanced chromatin accessibility or activation of distal regulatory elements, contributing to increased transcriptional plasticity in advanced ICC

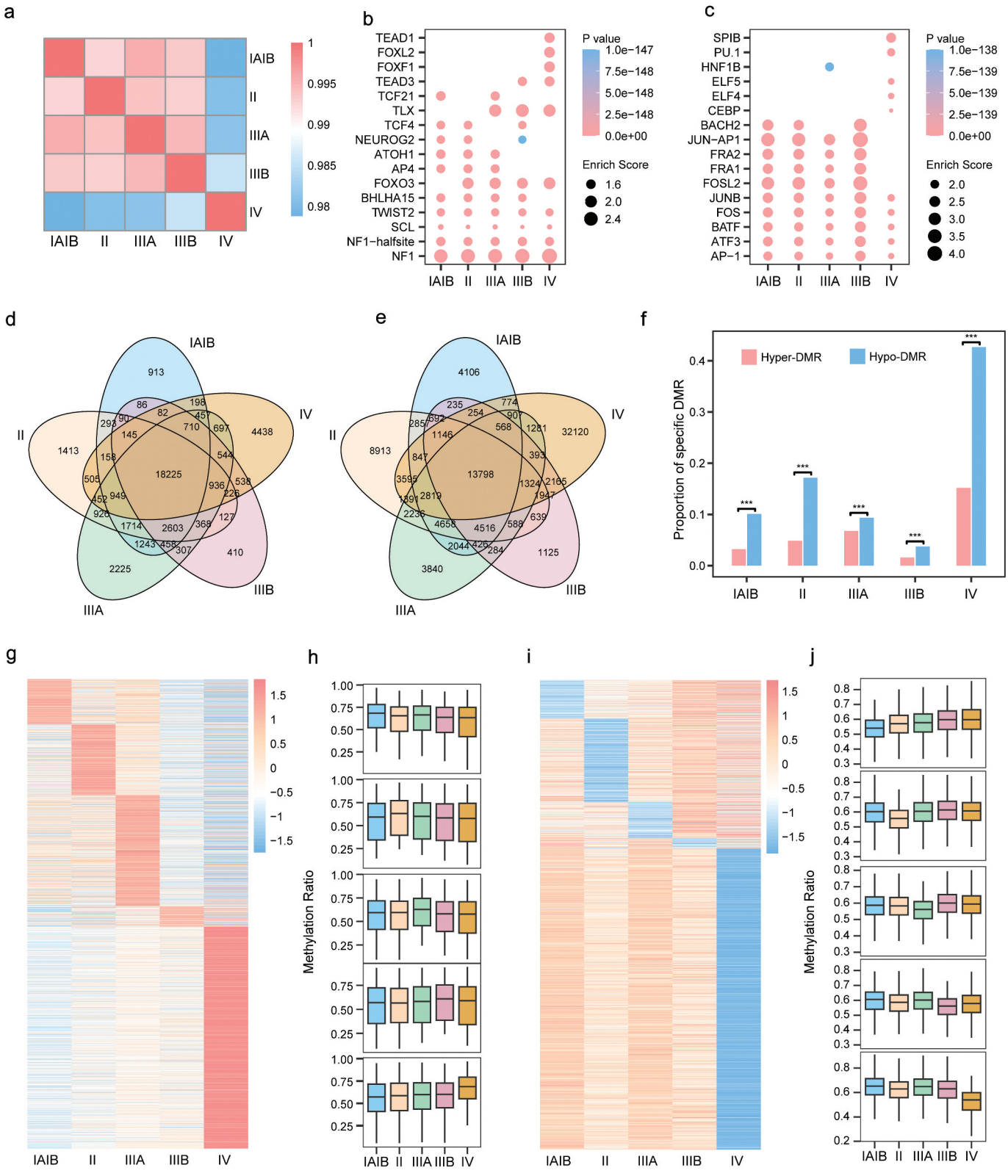

Figure 2. Stage IV ICC accumulated more DNA methylation aberrations. (a) Pearson's correlation analysis between stages IAIB, II, IIIA, IIIB, and IV using shared differentially methylated CpGs (DMCs) with a coverage of at least three reads. (b) Top 10 transcription factor (TF) motifs enriched at different ICC stages using hyper-DMRs. (c) Top 10 TF motifs enriched at different ICC stages using hypo-DMRs. P-values and enrichment scores for each motif are shown. (d) Venn diagram of hyper-differentially methylated regions (DMRs) at different intrahepatic cholangiocarcinoma (ICC) stages. (e) Venn diagram of hypo-DMRs at different ICC stages. Different colors represent different stages of ICC. Numbers in the Venn diagram represent the number of DMRs. (f) Proportions of specific hyper- and hypo-DMRs at each ICC stage. The proportion of hyper- (hypo-)DMRs in stage IAIB was calculated by dividing the number of IAIB-specific DMRs by the total number of DMRs in stage IAIB. The other stages were calculated in the same manner. (g) Heatmap of methylation ratios of stage-specific hyper-DMRs. (h) Boxplot of DNA methylation ratios of stage-specific hyper-DMRs. (i) Heatmap of methylation ratios of stage-specific hypo-DMRs. (j) Boxplot of DNA methylation ratios of stage-specific hypo-DMRs. The five boxplots from top to bottom are specific hyper- (H) and hypo- (J) DMRs for stages IAIB, II, IIIA, IIIB, and IV, respectively. For boxplots in (H) and (J), the upper and lower boundaries represent the first and third quantiles, respectively.

(Figure S7). Genes with stage IV-specific hyper- and hypo-DMRs were enriched in several pathways associated with cancer metastasis (Figure S7), including axon guidance [47], focal adhesion [48], and the cGMP-PKG signaling pathway [49]. These findings indicate that the oncogenesis and malignant progression processes of stage IV were distinct from those of the other stages.

Furthermore, we explored the cytosine methylation ratios of stage-specific hyper- and hypo-DMRs in all samples. For both hyper- and hypo-DMRs, stage IV-specific DMRs exhibited a markedly larger difference than those of the other stages (Figure 2(g–j)). We calculated the difference between the average CpG methylation ratios of stage IAIB and other stages in stage IAIB-specific DMRs, defined as MS_{IAIB}. The same methodology was used to obtain MS_{II}, MS_{IIIA}, MS_{IIIB}, and MS_{IV}. For hyper-DMRs, the values were 0.035, 0.043, 0.041, 0.032, and 0.970, for stages IAIB, II, IIIA, IIIB, and IV, respectively (Figure 2(h)). For hypo-DMRs, the values were −0.045, −0.051, −0.030, −0.031, and −0.101, respectively (Figure 2(j)). MS_{IV} was considerably larger than the other four values in hyper-DMRs and much smaller than the other four values in hypo-DMRs, suggesting that more DNA methylation aberrations accumulate in stage IV.

3.3. Distinguishing stage IV ICC from other stages using DMRs

Stage IV ICC is the most malignant tumor with distant metastases. Identifying the differences between stage IV ICC and other stages is beneficial for revealing the underlying mechanisms of metastasis and for early diagnosis. We divided all ICC cases into discovery ($n = 168$) and validation ($n = 173$) cohorts, each with eight stage IV samples. A logistic regression model with the LASSO penalty was constructed to distinguish stage IV cases from those of other stages using stage-specific DMRs identified in the Methods. In the discovery cohort, we observed excellent performance, with a receiver operating characteristic (ROC) area under the curve (AUC) of 0.991 (95% confidence interval [CI]: 0.978–1.00; Figure 3(a), Supplementary Table 2). In the validation cohort, the performance decreased slightly, with an AUC value of 0.894 (95% CI: 0.814–0.974; Figure 3(b)). The classification model included 15 DMRs (six hyper-DMRs and nine hypo-DMRs) as variables, located at the promoter ($n = 3$), exon ($n = 3$), intron ($n = 6$), or distal intergenic ($n = 3$) regions. The prognostic effects of these DMRs were explored, and it was found that the overall survival (OS) between high and low methylation ratios was significantly different in 73.3% (11/15) of the DMRs (Figure S8). The

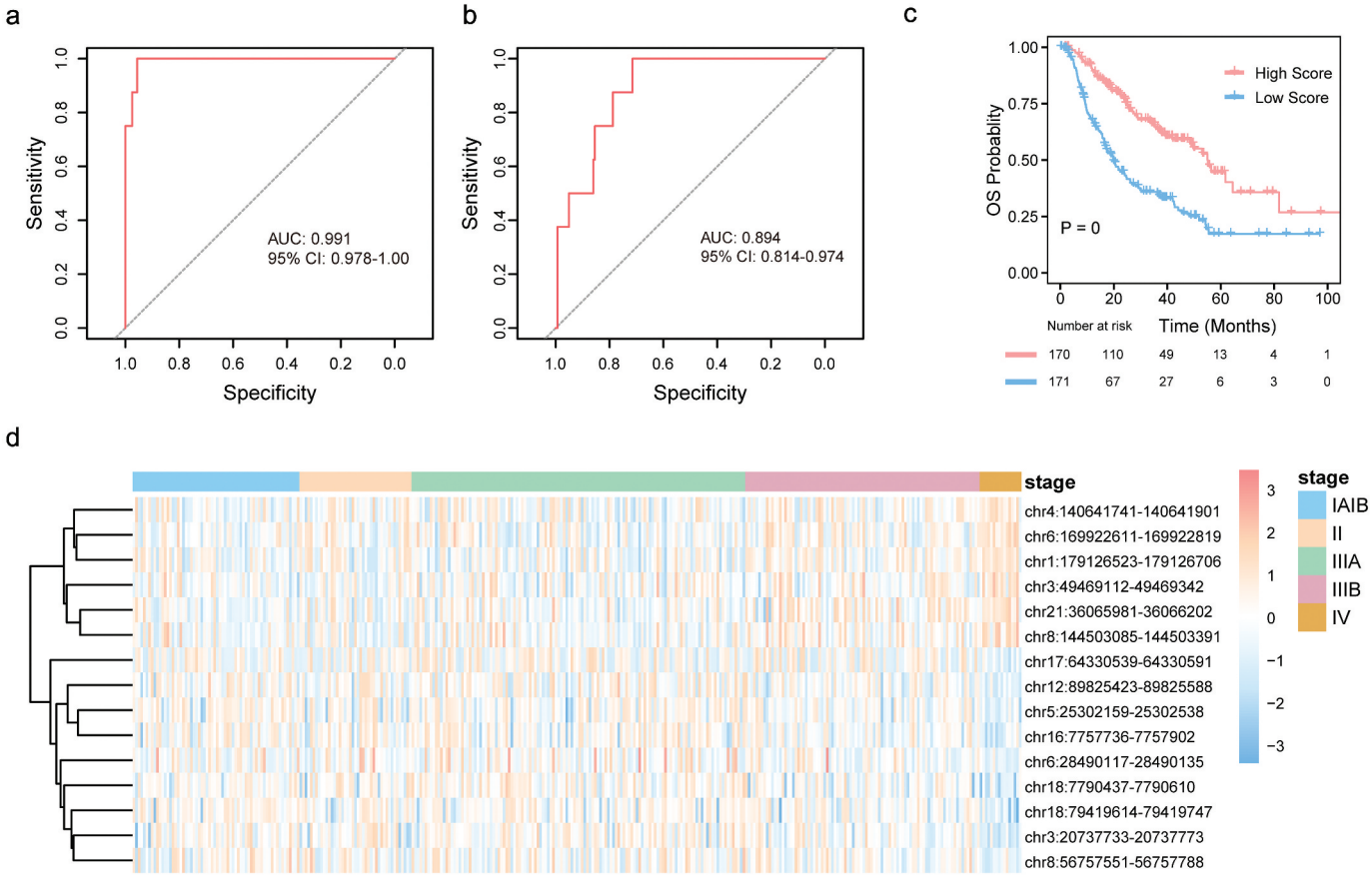

Figure 3. Distinguishing stage IV ICC from other stages using DMRs. (a) receiver operating characteristic (ROC) curve evaluating the overall performance of the model in distinguishing stage IV intrahepatic cholangiocarcinoma (ICC) from the other stages in the discovery cohort. (b) ROC curve evaluating the overall performance of the model in distinguishing stage IV ICC from the other stages in the validation cohort. The area under the ROC curve (AUC) and 95% confidence interval (CI) are shown. (c) Overall survival of patients with ICC with high or low scores calculated using the classification model. Patients with a score higher than the median value were classified into the high-score group, while the remainder were classified into the low-score group. P-values were obtained using the log-rank test. (d) Heatmap of methylation ratios of differentially methylated regions (DMRs) used in the classification model.

integrated model, which included all 15 DMRs, also demonstrated a significant prognostic effect. Individuals with a low score (median OS 55.2 months) had a significantly longer OS than those with a high score (median OS: 19.9 months, log-rank test, $p < 0.001$; Figure 3(c)). The methylation ratios of six hyper-DMRs and nine hypo-DMRs were higher and lower, respectively, in stage IV samples than in other ICC stage samples (Figure 3(d)).

3.4. Chromosome aneuploidy associated with ICC overall survival

In addition to methylation aberrations, genomic instability was prevalent in ICC, characterized by widespread chromosome arm aneuploidies. Gains (e.g., 1q) and losses (e.g., 6q) were observed across all chromosome arms, with gains significantly outnumbering losses (Figure 4(a)). The distribution of gains and losses in all arms was consistent across different stages (Figure S9), indicating conserved aneuploidy patterns in ICC. Aneuploidy scores, calculated by the number of chromosome arm aneuploidies per sample, were slightly higher in stage IV compared to earlier stages (Figure 4(b)), reflecting increased genomic instability. However, aneuploidy score showed no significant impact on patient OS (log-rank test, $p = 0.752$, Figure 4(c)). Specific arm aneuploidies, however, were associated with prognosis. Patients with 6q losses had longer OS than those with 6q copy neutral (log-rank test, $p = 0.016$, Figure 4(d)), and cases with 17p losses had worse prognosis than those without this alteration (log-rank test, $p < 0.001$, Figure 4(e)). The mechanistic basis for these associations may be related to key genes located on these chromosomal arms. For instance, *TP53* located on 17p, is a well-known tumor suppressor gene and its loss is associated with poor outcomes in multiple cancer types [50]. In contrast, deletion of oncogene *ROS1* and *MYB* located on 6q may attenuate oncogenic signaling and potentially mitigate malignant progression [51,52].

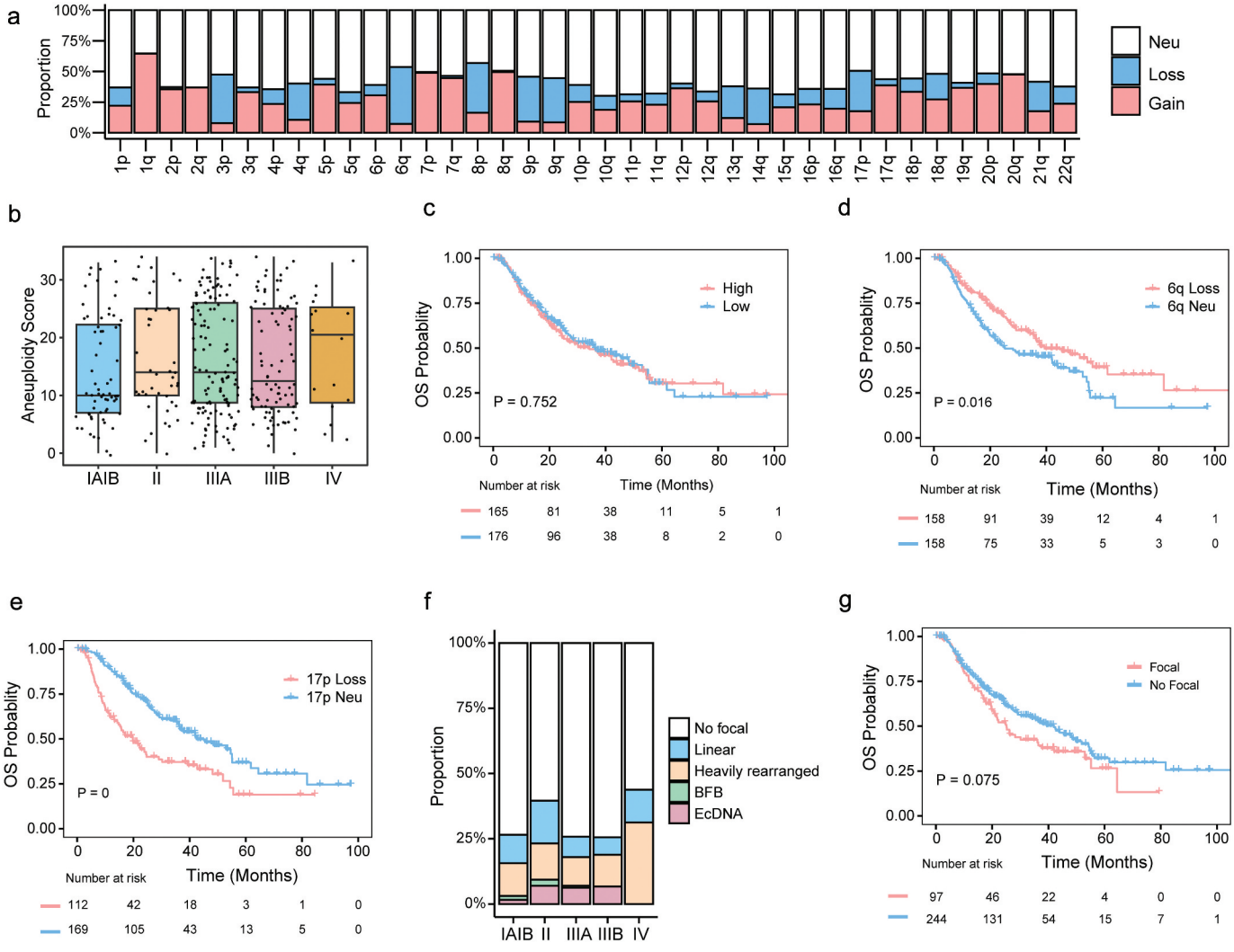

Figure 4. Chromosome arm aneuploidy of ICC and the association with survival. (a) the proportion of chromosome arm aneuploidy for 38 arms in ICC. Neu represents copy number neutral. (b) Box plot showing the aneuploidy score for each stage of ICC. (c) Overall survival of ICC patients with high or low aneuploidy scores. Patients with an aneuploidy score higher than the median value were classified into the high group, while the remainder were classified into the low group. (d) Overall survival of ICC patients with 6q loss or neutral. (E) Overall survival of ICC patients with 17p loss or neutral. (f) The proportion of four types of focal amplifications detected in different stages of ICC. Different colors represent different types of focal amplifications. No focal represents no focal amplification detected. (g) Overall survival of ICC patients with or without focal amplifications. *P*-values of (c), (d), (e) and (g) were obtained using log-rank test. BFB, breakage-fusion-bridge; EcDNA, extrachromosomal circular DNA.

These findings underscore the dual roles of chromosome arm aneuploidies in either promoting or suppressing tumorigenesis, depending on the specific alteration.

3.5. Focal amplifications in ICC

Focal amplification is another form of genomic instability, playing a vital role in cancer initiation and progression. Using AmpliconArchitect [41], we detected focal amplifications in 28% (97/341) of ICC samples. These focal amplifications were classified into four types: linear ($n = 32$), heavily rearranged ($n = 44$), breakage-fusion-bridge (BFB, $n = 3$) and ecDNA amplification ($n = 18$). The prevalence of focal amplifications varied across stages, with stage IV ICC exhibiting the highest proportion (44%), compared to stages IAIB (27%), II (40%), IIIA (26%), and IIIB (26%) (Figure 4(f)). Canonical driver genes were recurrently identified in amplified regions, with 23 genes recurring in at least two samples. Heavily rearranged amplifications accounted for the majority (76%) of driver gene amplifications, followed by ecDNA amplifications (16%). Among these, *MDM2* and *YEATS4* were the most frequently amplified ($n = 10$), all within heavily rearranged regions. Meanwhile, *CCND1*, *FGF3*, and *FGF4* were predominantly amplified in the form of ecDNA (Figure S10). Patients harboring focal amplifications exhibited a trend toward worse prognosis, although not statistically significant (log-rank test, $p = 0.075$, Figure 4(g)). These findings highlight the diversity of focal amplifications and their potential contributions to tumor aggressiveness.

3.6. Linear relationship between SCNA and DNA methylation

We were curious about the relationship between DNA methylation and SCNA, as they play similar role in tumor evolution. To dissect these interactions, we employed linear regression models to analyze the associations between SCNA and DNA methylation ratio or entropy across various genomic features (Methods). Our results revealed distinct patterns in methylation ratio. SCNA was predominantly negatively associated with methylation ratio in promoters and gene bodies but exhibited positive associations in CGIs (Figure 5(a–c), Figure S11). In contrast, methylation entropy demonstrated a consistent positive association with SCNA across promoters, gene bodies, and CGIs (Figure 5(d–f), Figure S12). These findings suggest that SCNA exerts differential effects on DNA methylation, influencing ratio and entropy in distinct ways based on genomic context. Importantly, we observed a robust positive correlation between SCNA and methylation entropy in the *MDM2* gene, a focal amplified oncogene in ICC (Figure 5(g)). Samples with *MDM2* amplification exhibited significantly higher methylation entropy compared to those without amplification (Student's *t*-test, $p = 0.0015$, Figure 5(h)). This highlights that focal amplifications, such as in *MDM2*, not only increase gene dosage but also induce greater variability and complexity in DNA methylation patterns. These findings suggest that SCNA amplifications, particularly in oncogenes like *MDM2*, exacerbate the disruption of methylation homeostasis. This disruption likely enhances the adaptability and aggressiveness of tumor cells, providing a selective advantage during tumor progression. Moreover, the distinct regional effects of SCNA on methylation ratios and the consistent amplification-driven increase in methylation entropy underscore the complex interplay between genomic and epigenomic instability in ICC evolution. These results further implicate SCNA-induced methylation alterations as a potential mechanism underlying the malignant progression of ICC.

4. Discussion

DNA methylation aberrations and genomic instability are pervasive across the cancer genome, providing the substrates for rapid evolution. These alterations have been extensively utilized in molecular subtyping, prognosis prediction, and the identification of therapeutic targets for ICC [20–27]. In this study, we systematically analyzed DNA methylation abnormalities and SCNA profiles across different stages of ICC. Stage IV ICC exhibited a higher accumulation of stage-specific methylation alterations compared to earlier stages, along with enrichment of exclusive transcription factors. Based on these stage-specific DMRs, we developed a classification model comprising 15 variables, achieving high accuracy in distinguishing stage IV ICC from other stages. Genomic instability was slightly higher in stage IV ICC, and the distribution of aneuploidies across chromosome arms was consistent across all stages. Specific chromosome arm aneuploidies were associated with OS, demonstrating both tumor-promoting and tumor-suppressing effects. Focal amplifications, another form of genomic instability, were detected in 28% of ICC samples, with stage IV showing the highest proportion. Among these, *MDM2* and *YEATS4* were the most frequently amplified oncogenes. Patients with focal amplifications exhibited slightly worse OS compared to those without. Furthermore, a linear relationship was identified between SCNA and DNA methylation. While SCNA showed varying effects on DNA methylation ratios depending on genomic regions, a predominantly positive correlation was observed between SCNA and methylation entropy across promoters, gene bodies, and CGIs, with the gene body of *MDM2* serving as a notable example. These findings highlight the pivotal roles of DNA methylation aberrations and genomic instability in ICC progression. They further suggest a synergistic interplay between these two processes, with genomic instability contributing to the disruption of DNA methylation patterns and driving malignant evolution in ICC.

This study builds upon our previous work, in which we characterized global DNA methylation alterations in ICC and classified all cases into four prognostic subtypes, identifying subtype-specific epigenetic drivers [25]. While that work focused on intertumoral heterogeneity across ICC subtypes, the present study addresses a distinct and clinically relevant question – the molecular basis of malignant progression across clinical stages of ICC. Here, we systematically dissected stage-specific DNA methylation aberrations and genomic instability across 341 ICC samples, identifying stage IV specific DMRs and CNAs and potential mechanistic link between genomic instability and epigenetic disorder during tumor progression. This work advances the understanding of how epigenetic and genetic alterations converge during late-stage ICC

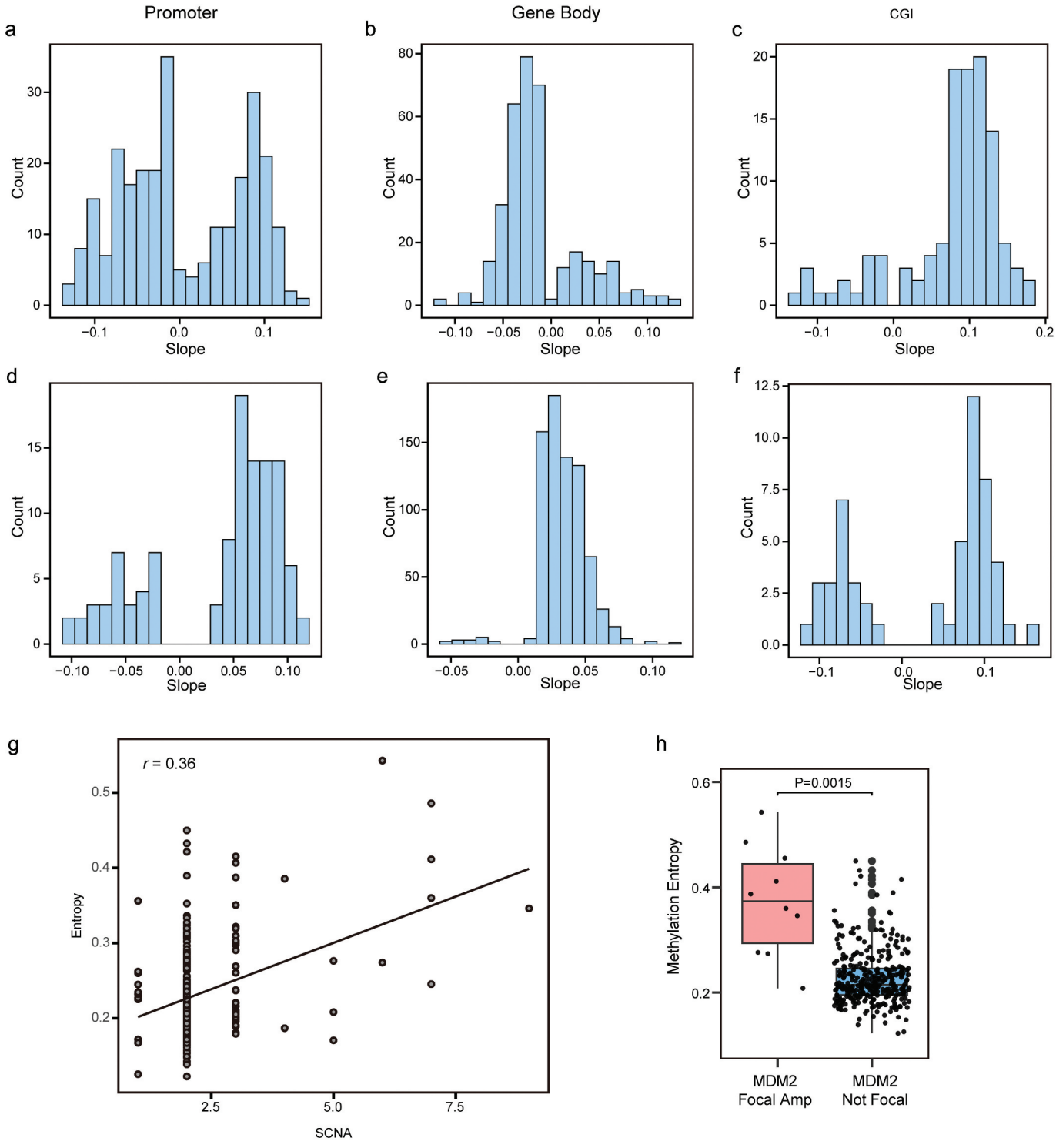

Figure 5. Linear relationship between SCNA and DNA methylation. distribution of slopes for linear model between DNA methylation ratio and SCNA in promoter (a), gene body (b), and CpG island (CGI) (c). Distribution of slopes for linear model between DNA methylation entropy and SCNA in promoter (d), gene body (e), and CpG island (CGI) (f). Slopes with p values smaller than 10^{-10} were shown. (g) Linear relationship between SCNA and methylation entropy for *MDM2*. Pearson correlation was shown ($r = 0.36$). The linear model between copy number and methylation entropy was: entropy = 0.177 + 0.025 × SCNA, R^2 value of the linear model is 0.134, $p = 3.427e\text{-}12$. (h) Box plot showing the methylation entropy of samples with (red) or without (blue) *MDM2* focal amplification. p value was obtained using Student's t-test.

evolution, offering potential biomarkers and therapeutic targets for metastatic ICCs.

Cancer evolution results from the multistep accumulation of molecular alterations [53], with longitudinal evolution of the methylome and genomic instability have been reported in various cancer types [7,9,10,54]. For instance, Methylation aberrations gradually increase from precancerous lesions to lung adenocarcinoma and evolve in parallel with genetic variations [9]. Similarly, this study revealed that stage IV ICC accumulates more methylation alterations, while earlier stages

exhibit only minor differences, with most DMRs being shared across stages. This suggests that tumors with distant metastases require more epigenetically driven forces, which was also observed in other cancer types [8,10]. In contrast, aneuploidy profiles between different stages were similar, stage IV ICC has slightly higher score. A similar trend was observed for focal amplifications of oncogenes, which were moderately more frequent in stage IV ICC. These findings imply that genomic instability likely represents an early event in ICC evolution, establishing a permissive landscape for tumor development. In later stages, DNA methylation aberrations appear to accumulate more extensively, potentially playing a critical role in driving tumor progression and metastasis. Together, these results highlight a temporal interplay between genomic instability and DNA methylation in ICC evolution, with early genomic alterations setting the stage for subsequent epigenetic changes that promote malignancy in advanced stages.

The association between DNA methylation and SCNA has been previously investigated in various cancer types [54–56]. Earlier studies revealed that the methylation m-values of CpG site are positively or negatively correlated with the copy number of adjacent genes. CpG sites with negative/positive associations were mainly located at CpG island/ocean respectively [55]. In this study, we adopted a different analytical approach by constructing a linear model where the dependent variable is the average methylation ratio or entropy, and the independent variable is the copy number of the same genomic region. This methodology allowed us to directly evaluate the localized effects of SCNA on DNA methylation within the same region, providing deeper insights into the interplay between these two types of alterations. Our results revealed that the association between SCNA and methylation entropy was primarily positive, unlike the variable associations observed for methylation ratio. DNA methylation entropy is associated with DNA-binding motifs, regulatory elements, and gene expression variability, playing critical roles in developmental plasticity and aging [57,58]. Additionally, methylation entropy serves as an important mediator of cancer cell plasticity, enabling adaptability and survival in changing environments [59]. Our findings suggest that increased copy number is associated with elevated methylation entropy, potentially enhancing the plasticity of cancer cells. This increased plasticity may provide tumor cells with greater evolutionary flexibility, promoting aggressive phenotypes and resistance to therapies. These results underscore the importance of integrating methylation entropy into studies of genomic instability to better understand its role in cancer progression.

DNA methylation aberrations are widely used as biomarkers for aging and early cancer detection, enabling opportunities for curative therapies and improved OS [60–62]. Early diagnosis of cancer metastasis also facilitates timely interventional treatment. This study revealed a notable difference in DNA methylation between stage IV and the other ICC stages. Distinguishing stage IV ICC from the other stages was achieved using only 15 markers. Previous studies have reported that metastatic tumors have a lower degree of intratumor heterogeneity and that the genomic landscape between early- and late-stage tumors is moderately different or highly consistent for most cancer types [63,64]. Several tumor types, such as breast, prostate, and kidney clear-cell carcinomas, are exceptions to this. Our findings suggest that while genomic instability plateaus across stages, methylation aberrations pronouncedly accumulate in advanced ICC. This underscores the critical role of epigenetic alterations in driving malignant progression and highlights their potential as biomarkers for detecting advanced-stage ICC.

MDM2 is a negative regulator of the tumor suppressor p53, promoting its degradation via ubiquitination. Amplification of *MDM2* was identified in many cancer types, which can impair p53-mediated cell cycle arrest and apoptosis, contributing to tumor progression even in the absence of *TP53* mutations [65,66]. In this study, *MDM2* amplification was observed in 2.9% (10/341) ICC cases, a frequency generally comparable with previous reports in biliary tract cancers, which have documented *MDM2* amplification rates ranging from ~ 4% to 8%, and ranging from 1.1% to 6% in ICC [67–69]. Several *MDM2* inhibitors, including small molecules targeting the *MDM2*-p53 interaction (e.g., Idasanutlin), are currently under clinical investigation and may offer new therapeutic strategies for patients with *MDM2* amplification [70,71]. These findings underscore the potential clinical relevance of *MDM2* amplification in a subset of ICC patients and warrant further functional and therapeutic exploration.

This study had several limitations. First, integrating genomic, epigenomic, and transcriptomic data could provide a more comprehensive understanding of malignant progression in ICC, but such multi-omics datasets are currently lacking. Second, the 15 DMRs identified to distinguish between early- and late-stage ICC require validation in independent cohorts with WGBS data and plasma samples spanning different stages of ICC to assess their potential utility in liquid biopsy. The validation using TCGA-CHOL 450K methylation array data was limited, as only two of the 15 DMRs overlapped with probes on the array, and the number of stage IV cases was only three, which is insufficient for robust evaluation. Finally, the observed relationship between DNA methylation and SCNA should be integrated with gene expression data to uncover the potential mechanisms linking these alterations to aberrant gene expression in ICC. Such analyses could facilitate the identification of novel therapeutic targets, paving the way for improved treatment strategies.

5. Conclusions

In conclusion, this study utilized WGBS data to characterize DNA methylation aberrations and genomic instability across different stages of ICC. We found that stage IV ICC exhibits significantly greater epigenetic changes and modestly increased genetic alterations compared to earlier stages, shedding light on the evolutionary trajectory of ICC progression. The positive correlation between SCNA and methylation entropy observed in our analysis suggests that local genomic instability contributes to methylation disorder. Future research should investigate the underlying mechanisms of the interplay between genetic alterations and epigenetic variations, which could provide valuable insights into ICC progression and potential therapeutic strategies.

Acknowledgments

The authors acknowledge the use of OpenAI's ChatGPT (version GPT-4) for English language editing during manuscript preparation. The tool was used to enhance grammar, clarity, and consistency in the writing. All scientific content, interpretation, and conclusions were developed independently by the authors.

Author contributions

Guanghao Li: Methodology, Formal analysis, Investigation, Data curation, Writing – original draft preparation, Writing – review and editing, Visualization. Youhuang Bai: Formal analysis, Data curation, Writing – review and editing. Feng Tao: Formal analysis. Tingting Hu: Data curation. Ting Wang: Formal analysis. Yong Zeng: Resources, Writing – review and editing. Deqiang Sun: Conceptualization, Supervision, Writing – review and editing, Funding acquisition.

Disclosure statement

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Funding

This manuscript was funded by the National Natural Science Foundation of China [Project 81773012] and the National Key Research and Development Program of China [2022YFA1105200 and 2023YFA1800700].The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data availability statement

All data analyzed in this study were obtained from publicly available datasets, as detailed in the Methods section. No new datasets were generated during this study.

References

Papers of special note have been highlighted as of interest (•) to readers.

1. Hanahan D. Hallmarks of cancer: new dimensions. Cancer Discov. 2022;12(1):31–46. doi: 10.1158/2159-8290.CD-21-1059
2. Chen X, Agustinus AS, Li J, et al. Chromosomal instability as a driver of cancer progression. Nat Rev Genet. 2025;26(1):31–46. doi: 10.1038/s41576-024-00761-7
3. Feinberg AP, Koldobskiy MA, Gondor A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. Nat Rev Genet. 2016;17(5):284–299. doi: 10.1038/nrg.2016.13
4. Nishiyama A, Nakanishi M. Navigating the DNA methylation landscape of cancer. Trends Genet. 2021;37(11):1012–1027. doi: 10.1016/j.tig.2021.05.002
5. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13(7):484–492. doi: 10.1038/nrg3230
   • **This paper summarizes the functional roles of DNA methylation across different genomic regions.**
6. Landau DA, Clement K, Ziller MJ, et al. Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. Cancer Cell. 2014;26(6):813–825. doi: 10.1016/j.ccell.2014.10.012
7. Pan H, Jiang Y, Boi M, et al. Epigenomic evolution in diffuse large B-cell lymphomas. Nat Commun. 2015;6(1):6921. doi: 10.1038/ncomms7921
8. Zhao SG, Chen WS, Li H, et al. The DNA methylation landscape of advanced prostate cancer. Nat Genet. 2020;52(8):778–789. doi: 10.1038/s41588-020-0648-8
9. Hu X, Estecio MR, Chen R, et al. Evolution of DNA methylome from precancerous lesions to invasive lung adenocarcinomas. Nat Commun. 2021;12(1):687. doi: 10.1038/s41467-021-20907-z
   • **This paper reveals the stepwise evolution of DNA methylation from precancerous lesions to invasive cancer.**
10. Zhao S, Li J, Zhang H, et al. Epigenetic alterations of repeated relapses in patient-matched childhood ependymomas. Nat Commun. 2022;13(1):6689. doi: 10.1038/s41467-022-34514-z
11. Sdeor E, Okada H, Saad R, et al. Aneuploidy as a driver of human cancer. Nat Genet. 2024;56(10):2014–2026. doi: 10.1038/s41588-024-01916-2
12. Shukla A, Nguyen THM, Moka SB, et al. Chromosome arm aneuploidies shape tumour evolution and drug response. Nat Commun. 2020;11(1):449. doi: 10.1038/s41467-020-14286-0
   • **This paper reveals the roles of aneuploidy in cancer evolution and drug response.**
13. Zhang J, Feng Y, Li G, et al. Distinct aneuploid evolution of astrocytoma and glioblastoma during recurrence. NPJ Precis Oncol. 2023;7(1):97. doi: 10.1038/s41698-023-00453-1
14. Stichel D, Ebrahimi A, Reuss D, et al. Distribution of EGFR amplification, combined chromosome 7 gain and chromosome 10 loss, and TERT promoter mutation in brain tumors and their potential for the reclassification of IDHwt astrocytoma to glioblastoma. Acta Neuropathol. 2018;136(5):793–803. doi: 10.1007/s00401-018-1905-0
15. Yan X, Mischel P, Chang H. Extrachromosomal DNA in cancer. Nat Rev Cancer. 2024;24(4):261–273. doi: 10.1038/s41568-024-00669-8
   • **This paper summarizes the role of ecDNA for cancer.**
16. Chun YS, Pawlik TM, Vauthey JN. 8th edition of the AJCC cancer Staging Manual: pancreas and hepatobiliary cancers. Ann Surg Oncol. 2018;25(4):845–847. doi: 10.1245/s10434-017-6025-x
17. Lee AJ, Chun YS. Intrahepatic cholangiocarcinoma: the AJCC/UICC 8th edition updates. Chin Clin Oncol. 2018;7(5):52. doi: 10.21037/cco.2018.07.03
18. Marin JJG, Prete MG, Lamarca A, et al. Current and novel therapeutic opportunities for systemic therapy in biliary cancer. Br J Cancer. 2020;123(7):1047–1059. doi: 10.1038/s41416-020-0987-3
19. Banales JM, Marin JJG, Lamarca A, et al. Cholangiocarcinoma 2020: the next horizon in mechanisms and management. Nat Rev Gastroenterol Hepatol. 2020;17(9):557–588. doi: 10.1038/s41575-020-0310-z
20. Goeppert B, Toth R, Singer S, et al. Integrative analysis defines distinct prognostic subgroups of Intrahepatic Cholangiocarcinoma. Hepatology. 2019;69(5):2091–2106. doi: 10.1002/hep.30493
21. Farshidfar F, Zheng S, Gingras MC, et al. Integrative genomic analysis of cholangiocarcinoma identifies distinct IDH-Mutant molecular profiles. Cell Rep. 2017;18(11):2780–2794. doi: 10.1016/j.celrep.2017.02.033
22. Jusakul A, Cutcutache I, Yong CH, et al. Whole-genome and epigenomic landscapes of etiologically distinct subtypes of cholangiocarcinoma. Cancer Discov. 2017;7(10):1116–1135. doi: 10.1158/2159-8290.CD-17-0368
23. Boerner T, Drill E, Pak LM, et al. Genetic determinants of outcome in intrahepatic cholangiocarcinoma. Hepatology. 2021;74(3):1429–1444. doi: 10.1002/hep.31829
24. Chen X, Dong L, Chen L, et al. Epigenome-wide development and validation of a prognostic methylation score in intrahepatic cholangiocarcinoma based on machine learning strategies. Hepatobiliary Surg Nutr. 2023;12(4):478–494. doi: 10.21037/hbsn-21-424
25. Liao H, Chen X, Wang H, et al. Whole-genome DNA methylation profiling of intrahepatic cholangiocarcinoma reveals prognostic subtypes with distinct biological drivers. Cancer Res. 2024;84(11):1747–1763. doi: 10.1158/0008-5472.CAN-23-3298
26. Wang XY, Zhu WW, Wang Z, et al. Driver mutations of intrahepatic cholangiocarcinoma shape clinically relevant genomic clusters with

distinct molecular features and therapeutic vulnerabilities. Theranostics. 2022;12(1):260–276. doi: 10.7150/thno.63417
27. Zhang Y, Ma Z, Li C, et al. The genomic landscape of cholangiocarcinoma reveals the disruption of post-transcriptional modifiers. Nat Commun. 2022;13(1):3061. doi: 10.1038/s41467-022-30708-7
28. Corces MR, Granja JM, Shams S, et al. The chromatin accessibility landscape of primary human cancers. Science. 2018;362(6413). doi: 10.1126/science.aav1898
29. Chen S, Zhou Y, Chen Y, et al. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–i890. doi: 10.1093/bioinformatics/bty560
30. Xi Y, Li W. BSMAP: whole genome bisulfite sequence MAPping program. BMC Bioinformatics. 2009;10(1):232. doi: 10.1186/1471-2105-10-232
31. Sun D, Xi Y, Rodriguez B, et al. MOABS: model based analysis of bisulfite sequencing data. Genome Biol. 2014;15(2):R38. doi: 10.1186/gb-2014-15-2-r38
32. Quinlan AR, Hall IM. Bedtools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841–842. doi: 10.1093/bioinformatics/btq033
33. Ramirez F, Dundar F, Diehl S, et al. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. 2014;42 (Web Server issue):W187–91. doi: 10.1093/nar/gku365
34. Yu G, Wang L-G, Q-Y H. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics. 2015;31(14):2382–2383. doi: 10.1093/bioinformatics/btv145
35. Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innov. 2021;2(3):100141. doi: 10.1016/j.xinn.2021.100141
36. Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38(4):576–589. doi: 10.1016/j.molcel.2010.05.004
37. Colaprico A, Silva TC, Olsen C, et al. Tcgabiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. 2016;44(8):e71. doi: 10.1093/nar/gkv1507
38. Zhu LJ, Gazin C, Lawson ND, et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics. 2010;11(1):237. doi: 10.1186/1471-2105-11-237
39. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics. 2011;12(1):35. doi: 10.1186/1471-2105-12-35
40. Adalsteinsson VA, Ha G, Freeman SS, et al. Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. Nat Commun. 2017;8(1):1324. doi: 10.1038/s41467-017-00965-y
41. Deshpande V, Luebeck J, Nguyen ND, et al. Exploring the landscape of focal amplifications in cancer using AmpliconArchitect. Nat Commun. 2019;10(1):392. doi: 10.1038/s41467-018-08200-y
42. Kim H, Nguyen NP, Turner K, et al. Extrachromosomal DNA is associated with oncogene amplification and poor outcome across multiple cancers. Nat Genet. 2020;52(9):891–897. doi: 10.1038/s41588-020-0678-2
43. Ding Y, Cai K, Liu L, et al. mHaptk: a comprehensive toolkit for the analysis of DNA methylation haplotypes. Bioinformatics. 2022;38 (22):5141–5143. doi: 10.1093/bioinformatics/btac650
44. Tome-Garcia J, Erfani P, Nudelman G, et al. Analysis of chromatin accessibility uncovers TEAD1 as a regulator of migration in human glioblastoma. Nat Commun. 2018;9(1):4020. doi: 10.1038/s41467-018-06258-2
45. Lourenco AR, Roukens MG, Seinstra D, et al. C/EBPa is crucial determinant of epithelial maintenance by preventing epithelial-to-mesenchymal transition. Nat Commun. 2020;11(1):785. doi: 10.1038/s41467-020-14556-x
46. Gallego-Ortega D, Ledger A, Roden DL, et al. ELF5 drives lung metastasis in luminal breast cancer through recruitment of Gr1+ CD11b+ myeloid-derived suppressor cells. PLOS Biol. 2015;13(12): e1002330. doi: 10.1371/journal.pbio.1002330
47. Jurcak NR, Rucki AA, Muth S, et al. Axon guidance molecules promote perineural invasion and metastasis of orthotopic pancreatic tumors in mice. Gastroenterology. 2019;157(3):838–850 e6. doi: 10.1053/j.gastro.2019.05.065
48. Murphy JM, Rodriguez YAR, Jeong K, et al. Targeting focal adhesion kinase in cancer cells and the tumor microenvironment. Exp Mol Med. 2020;52(6):877–886. doi: 10.1038/s12276-020-0447-4
49. Lv Y, Wang X, Li X, et al. Nucleotide de novo synthesis increases breast cancer stemness and metastasis via cGMP-PKG-MAPK signaling pathway. PLOS Biol. 2020;18(11):e3000872. doi: 10.1371/journal.pbio.3000872
50. Chen X, Zhang T, Su W, et al. Mutant p53 in cancer: from molecular mechanism to therapeutic modulation. Cell Death Dis. 2022;13 (11):974. doi: 10.1038/s41419-022-05408-1
51. Drilon A, Jenkins C, Iyer S, et al. ROS1-dependent cancers - biology, diagnostics and therapeutics. Nat Rev Clin Oncol. 2021;18(1):35–55. doi: 10.1038/s41571-020-0408-9
52. Ciciro Y, Sala A. MYB oncoproteins: emerging players and potential therapeutic targets in human cancer. Oncogenesis. 2021;10(2):19. doi: 10.1038/s41389-021-00309-y
53. Li G, Yang Z, Wu D, et al. Evolution under spatially heterogeneous selection in solid tumors. Mol Biol Evol. 2022;39(1). doi: 10.1093/molbev/msab335
54. Haga Y, Sakamoto Y, Kajiya K, et al. Whole-genome sequencing reveals the molecular implications of the stepwise progression of lung adenocarcinoma. Nat Commun. 2023;14(1):8375. doi: 10.1038/s41467-023-43732-y
55. Sun W, Bunn P, Jin C, et al. The association between copy number aberration, DNA methylation and gene expression in tumor samples. Nucleic Acids Res. 2018;46(6):3009–3018. doi: 10.1093/nar/gky131
• **This paper reveals the assocation between CNA, gene expression and DNA methylation.**
56. Shi X, Radhakrishnan S, Wen J, et al. Association of CNVs with methylation variation. NPJ Genom Med. 2020;5(1):41. doi: 10.1038/s41525-020-00145-w
57. Fang Y, Ji Z, Zhou W, et al. DNA methylation entropy is associated with DNA sequence features and developmental epigenetic divergence. Nucleic Acids Res. 2023;51(5):2046–2065. doi: 10.1093/nar/gkad050
58. Vaidya H, Jeong HS, Keith K, et al. DNA methylation entropy as a measure of stem cell replication and aging. Genome Biol. 2023;24 (1):27. doi: 10.1186/s13059-023-02866-4
59. Feinberg AP, Levchenko A. Epigenetics as a mediator of plasticity in cancer. Science. 2023;379(6632):eaaw3835. doi: 10.1126/science.aaw3835
• **This paper summarizes how epigenetic regulation contributes to cancer plasticity.**
60. Zhao D, Wang Y, Wong Nathan D, et al. Impact of aging on cardiovascular diseases. JACC Asia. 2024;4(5):345–358. doi: 10.1016/j.jacasi.2024.02.002
61. Perry RJ, Shulman GI. Mechanistic links between obesity, insulin, and cancer. Trends Cancer. 2020;6(2):75–78. doi: 10.1016/j.trecan.2019.12.003
62. Alix-Panabieres C, Pantel K. Liquid biopsy: from discovery to clinical application. Cancer Discov. 2021;11(4):858–873. doi: 10.1158/2159-8290.CD-20-1311
63. Martinez-Jimenez F, Movasati A, Brunner SR, et al. Pan-cancer whole-genome comparison of primary and metastatic solid tumours. Nature. 2023;618(7964):333–341. doi: 10.1038/s41586-023-06054-z
64. Priestley P, Baber J, Lolkema MP, et al. Pan-cancer whole-genome analyses of metastatic solid tumours. Nature. 2019;575 (7781):210–216. doi: 10.1038/s41586-019-1689-y
65. Wade M, Li YC, Wahl GM. MDM2, MDMX and p53 in oncogenesis and cancer therapy. Nat Rev Cancer. 2013;13(2):83–96. doi: 10.1038/nrc3430

66. Karni-Schmidt O, Lokshin M, Prives C. The roles of MDM2 and MDMX in cancer. Annu Rev Pathol. 2016;11(1):617–644. doi: 10.1146/annurev-pathol-012414-040349
67. Kim SJ, Akita M, Sung Y-N, et al. MDM2 amplification in intrahepatic Cholangiocarcinomas: its relationship with large-duct type morphology and uncommon KRAS mutations. Am J Surg Pathol. 2018;42(4):512–521. doi: 10.1097/PAS.0000000000001006
68. Kratz JD, Klein AB, Gray CB, et al. The epidemiology of biliary tract cancer and associated prevalence of MDM2 amplification: a targeted literature review. Target Oncol. 2024;19(6):833–844. doi: 10.1007/s11523-024-01086-5
69. Muñoz Martín AJ, Castet F, Soto Alsar J, et al. MDM2 amplification in patients with biliary tract cancer in RETUD registry. J Clin Oncol. 2025;43(4_suppl):615–615. doi: 10.1200/JCO.2025.43.4_suppl.615
70. Yao Y, Zhang Q, Li Z, et al. MDM2: current research status and prospects of tumor treatment. Cancer Cell Int. 2024;24(1):170. doi: 10.1186/s12935-024-03356-8
71. Khurana A, Shafer DA. MDM2 antagonists as a novel treatment option for acute myeloid leukemia: perspectives on the therapeutic potential of idasanutlin (RG7388). Onco Targets Ther. 2019;12:2903–2910. doi: 10.2147/OTT.S172315