Whole-Genome DNA Methylation Profiling of Intrahepatic Cholangiocarcinoma Reveals Prognostic Subtypes with Distinct Biological Drivers

Haotian Liao[1], Xing Chen[2], Haichuan Wang[1], Youpei Lin[3], Lu Chen[4], Kefei Yuan[1], Mingheng Liao[1], Hanyu Jiang[5], Jiajie Peng[6], Zhenru Wu[7], Jiwei Huang[1], Jiaxin Li[1], and Yong Zeng[1]

ABSTRACT

Intrahepatic cholangiocarcinoma (iCCA) is the second most prevalent primary liver cancer. Although the genetic characterization of iCCA has led to targeted therapies for treating tumors with FGFR2 alterations and IDH1/2 mutations, only a limited number of patients can benefit from these strategies. Epigenomic profiles have emerged as potential diagnostic and prognostic biomarkers for improving the treatment of cancers. In this study, we conducted whole-genome bisulfite sequencing on 331 iCCAs integrated with genetic, transcriptomic, and proteomic analyses, demonstrating the existence of four DNA methylation subtypes of iCCAs (S1–S4) that exhibited unique postoperative clinical outcomes. The S1 group was an IDH1/2 mutation–specific subtype with moderate survival. The S2 subtype was characterized by the lowest methylation level and the highest mutational burden among the four subtypes and displayed upregulation of a gene-expression pattern associated with cell cycle/DNA replication. The S3 group was distinguished by high interpatient heterogeneity of tumor immunity, a gene-expression pattern associated with carbohydrate metabolism, and an enrichment of KRAS alterations. Patients with the S2 and S3 subtypes had the shortest survival among the four subtypes. Tumors in the S4 subtype, which had the best prognosis, showed global methylation levels comparable to normal controls, increased FGFR2 fusions/BAP1 mutations, and the highest copy-number variant burdens. Further integrative and functional analyses identified GBP4 demethylation, which is highly prevalent in the S2 and S3 groups, as an epigenetic oncogenic factor that regulates iCCA proliferation, migration, and invasion. Together, this study identifies prognostic methylome alterations and epigenetic drivers in iCCA.

Significance: Characterization of the DNA methylome of intrahepatic cholangiocarcinoma integrated with genomic, transcriptomic, and proteomic analyses uncovers molecular mechanisms affected by genome-wide DNA methylation alterations, providing a resource for identifying potential therapeutic targets.

[1]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, Sichuan, China. [2]Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, Zhejiang, China. [3]Department of Liver Surgery and Transplantation, Liver Cancer Institute, Zhongshan Hospital, and Key Laboratory of Carcinogenesis and Cancer Invasion, (Ministry of Education), Fudan University, Shanghai, China. [4]Department of Hepatobiliary Cancer, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, China. [5]Department of Radiology, West China Hospital, Sichuan University, Chengdu, Sichuan, China. [6]School of Computer Science, Northwestern Polytechnical University, Xi'an, Shanxi, China. [7]Institute of Clinical Pathology, Key Laboratory of Transplant Engineering and Immunology, NHC, West China Hospital, Sichuan University, Chengdu, Sichuan, China.

H. Liao, X. Chen, and H. Wang contributed equally to this article.

Corresponding Authors: Yong Zeng, 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, Sichuan, China. E-mail: zeng_y@scu.edu.cn; and Jiaxin Li, 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, Sichuan, China, E-mail: lijiaxin@scu.edu.cn

Cancer Res 2024;84:1747–63

doi: 10.1158/0008-5472.CAN-23-3298

Introduction

DNA methylation is one of the key epigenetic mechanisms regulating gene expression and maintaining genome stability (1), which is characterized by adding methyl groups to the 5′ carbon of cytosine nucleotides adjacent to CpG dinucleotides, resulting in 5-methylcytosine nucleotides (2). In various cancer types, the hypermethylation of CpG islands (CGI) in promoter regions decreases the expression of tumor suppressor genes, whereas hypomethylated regions are positively correlated with the activation of oncogenes and genome instability (3). Owing to the implication of DNA methylation aberrations in oncogenesis, recent studies based on large scales of patients with cancer, including those with esophageal adenocarcinoma (4), acute lymphoblastic leukemia (5), glioblastoma (6), prostate cancer (7), and meningioma (8), have depicted the DNA methylation heterogeneity among patients and identified methylation subtypes with potential therapeutic targets.

As the second most common primary liver cancer with globally increasing incidence (9), intrahepatic cholangiocarcinoma (iCCA) is characterized by high invasiveness and, while surgical resection remains the main treatment option with curative intent, high frequency of postoperative recurrence. Despite great progress in systemic therapies such as FGFR2-fusion– and IDH-mutation–targeted therapies, only a limited number of iCCA patients can benefit from them (10). Thus, deeper mechanistic insights into the pathogenesis of iCCA are still in urgent need for the identification of novel and effective therapeutics. Currently, research on iCCA is still focused on the genomic level (11–13). As the DNA methylation pattern has emerged as a diagnostic and prognostic assay for molecular

classification of various cancers, characterization of iCCA based on DNA methylation pattern may therefore provide additional information into the molecular heterogeneity from an epigenetics perspective, which cannot be captured by genomics analysis alone. Previously, some groups have performed DNA methylation studies on cholangiocarcinoma, which included extrahepatic cholangiocarcinoma (eCCA) samples, perihilar cholangiocarcinoma samples, and iCCA samples. Considering the genetic variations among different subtypes of cholangiocarcinoma, a comprehensive understanding of the DNA methylation landscape on iCCA-only samples will be more beneficial for the patient-tailored therapies (11, 12, 14). Meanwhile, because many significant regulatory regions are outside the array-targeted areas, whole-genome bisulfite sequencing (WGBS) is required to systematically study the DNA methylation of the entire genome at single-base resolution. However, at this moment, no sequencing-based methods have been conducted on the genome-wide DNA methylation pattern of iCCAs, especially on a large cohort of patients. Besides, there is only a limited number of studies reporting the application of WGBS on large cancer cohorts. Moreover, multiomics studies integrating WGBS with other high-throughput approaches to comprehensively evaluate the effect of DNA methylation on the inner biological features of cancer have rarely been reported.

Herein, we presented a WGBS study comprising over 300 iCCA cases from a multicenter cohort, in which multiple complementary approaches were integrated, including whole-exome sequencing (WES), RNA sequencing (RNA-seq), proteomics and IHC. The clinical information in detail facilitates the identification of prognostic significance of methylome alterations, and further functional analyses were performed to identify epigenetic drivers with prognostic potential.

Materials and Methods

Acquisition of clinical specimens

The iCCA samples used in this study were harvested from a multicenter and retrospective cohort of iCCA patients undergoing primary resection without any anticancer treatments before surgery from May 2010 to July 2019, including patients from Zhongshan Hospital of Fudan University (FU-iCCA, $n = 121$), West China Hospital (WCH-iCCA, $n = 154$), and Tianjin Medical University Cancer Institute and Hospital (TMUCIH-iCCA, $n = 56$). Nine non-neoplastic samples of cholangiocytes originating from the common bile duct of healthy liver-transplantation donors were enrolled as the normal controls. Tissues were obtained and stored in liquid nitrogen for less than 30 minutes after surgical resection. The WES, RNA-seq, and proteomics data from FU-iCCA cohort were applied for methylation-based integration analysis (Supplementary Fig. S1A and S1B; ref. 15). The use of cancer and normal-control specimens was approved by the Research Ethics Committee of each center, and we obtained written informed consent from each patient before surgical resection. The study was conducted according to the Declaration of Helsinki Principles. The analysis pipelines for multiomics data of iCCA samples included in this study can be found in Supplementary Methods.

Public data set

Publicly available chromatin immunoprecipitation sequencing (ChIP-seq) data of HuCC-T1 for transcription factors (YAP, TEAD1, TEAD4, and PRH) and H3K27ac were downloaded from Gene Expression Omnibus (16, 17). The annotation of CGIs for hg19 was downloaded from the UCSC Genome Browser. CGI shores were defined as the 2 kb flanking a CGI on each side, whereas shelves were defined as the 2 kb flanking the shores. CGIs with methylation levels ≥ 0.2 were defined as methylated CGIs. Promoters were defined as the 2 kb flanking annotated TSS in GENCODE (release 19). The annotated gene coordinates in GENECODE annotation were defined as gene bodies. Publicly available 18-state chromatin segmentation data of adult human liver were obtained from ChromHMM (http://compbio.mit.edu/ChromHMM/) and grouped into the following categories: Active TSS (1_TssA, 2_TssFlnk, 4_TssFlnkD, 3_TssFlnkU), Bivalent TSS (14_TssBiv), Transcript (6_TxWk, 5_Tx), Enhancer (11_EnhWk, 9_EnhA1, 7_EnhG1, 15_EnhBiv, 10_EnhA2, 8_EnhG2), Heterochromatin (12_ZNF/Rpts, 13_Het), Repressive (17_ReprPCWk, 16_ReprPC), and Quiescent (18_Quies). Annotations of consensus PMDs, HMDs, and solo-WCGW CpGs across various malignant tissues were obtained at https://zwdzwd.github.io/pmd.

Cell lines and cell culture

HuCC-T1 and HCCC-9810 human iCCA cell lines authenticated and tested negative for *Mycoplasma* were purchased from the Cell Bank of the Shanghai Institute for Biological Sciences (Chinese Academy of Sciences). These two cell lines were maintained in RPMI-1640 medium with 10% fetal bovine serum (HyClone) and cultured in a humidified incubator at 37°C with 5% CO_2. The methods for RNA interference and Western blot analysis can be found in Supplementary Methods.

DNA isolation and WGBS

Genomic DNA from tissues was extracted using the Tiangen DP304 kit according to the manufacturer's protocol. DNA purity was measured by the NanoPhotometer spectrophotometer (IMPLEN). DNA was quantified using Qubit DNA Assay Kit in Qubit 2.0 Fluorometer (Life Technologies). A total amount of 5.2-μg genomic DNA spiked with 26 ng lambda DNA per sample was randomly fragmented into 200 to 300 bp with Covaris S220, followed by end repair and adenylation. Cytosine-methylated barcodes were ligated to sonicated DNA following the manufacturer's instructions, then the barcode-ligated DNA fragments were treated twice with bisulfite using EZ DNA Methylation-GoldTM Kit (Zymo Research), which was followed by PCR amplification of bisulfite-treated single-strand DNA fragments using KAPA HiFi HotStart Uracil + ReadyMix (2×), quantification of library concentration by Qubit 2.0 Fluorometer (Life Technologies), and assessment of insert size on Agilent Bioanalyzer 2100 system. Finally, the constructed library was sequenced on an Illumina NovaSeq platform at Novogene Co., Ltd., in which 150 bp paired-end reads were generated.

Consensus clustering for variably methylated blocks

To characterize the DNA methylation heterogeneity among iCCA candidates at the whole-genome level, we first divided the genome into continuous genomic regions (blocks) showing homogeneous methylation levels across multiple CpGs for each iCCA sample by wgbstools (https://github.com/nloyfer/wgbs_tools; ref. 18) with parameters "–min_cpg 3 –max_bp 5000." Blocks with sufficient coverage of at least 10 observations (calculated as sequenced CpGs) over 2 of 3 of the iCCAs included were further retained. K-nearest neighbor (k-NN) imputation by R package "pamr" (v 1.56.1) was applied to impute the missing values. Principal component analysis (PCA) was then performed on the retained blocks ($n = 2{,}150{,}292$) in R using the base command "prcomp" (parameters: center = TRUE; scale = FALSE). The elbow method showed that the optimal number of PCs is 3 or 4, but PC4 was excluded, as it contributed less than 5% of the variance

(Supplementary Fig. S2A). In contrast, both PC1 and PC3 contributed to > 5% variance. Thus, variable blocks from the first three PCs were chosen for analysis. Then, we selected the top 666 blocks from PC1 to PC3 (1,998 blocks) based on the absolute gene loading score values within each PC. The top 1,998 blocks were used for *K*-means consensus clustering to generate iCCA subtypes using R package "ConsensusClusterPlus" (v1.54.0) with the following parameters: number of repetitions: 1,000 bootstraps; pItem: 0.8 (resampling 80% of any samples); pFeature: 0.8 (resampling 80% of any blocks); clusterAlg: "km"; innerLinkage: average, finalLinkage: average, distance: Euclidean; lower limit of cluster number: 2; upper limit of cluster number: 6. Selection of cluster number was based on the average pairwise consensus matrix within consensus clusters, the delta plot depicting the relative change in the area under the cumulative distribution function curve, and the averaged silhouette distance for consensus clusters. A 4-cluster was selected as the best solution because the consensus matrix with $k = 4$ exhibited the clearest separations among clusters (Supplementary Fig. S2B–S2D). Moreover, the average silhouette distance reached the highest when $k = 4$ (Supplementary Fig. S2E). Based on the evidence above, the iCCA WGBS data were clustered into 4 groups. Using the top 333, 1,000, 1,333, or 1,666 blocks from each PC only regrouped 1% to 6% of iCCAs, suggesting the number of selected blocks across the top three PCs did not affect the results of consensus clustering.

Global DNA methylation analysis

To avoid the bias caused by CGI, the global methylation level of each sample was calculated using the arithmetic mean of the methylation levels of all CpGs outside CGIs, as the high CpG densities within CGIs yield unbalanced mean methylation values, not representative of global methylation. The CGI methylation level of each sample was determined using the arithmetic mean of the methylation levels of all CpGs within CGIs. The methylation of each consensus partially methylated domain (PMD) or highly methylated domain (HMD) was determined by the arithmetic mean across solo-WCGW CpGs within the PMD or HMD. Overlap of CpGs with features was determined using bedtools (19) "intersectBed." Only features covered by at least 3 CpGs were considered for further analyses.

Clustering of CGIs

To identify iCCA-specific CGI clusters, we selected commonly covered CGIs between iCCAs and normal controls, followed by CGI clusters generated by R package "ConsensusClusterPlus" (v1.54.0) with the following parameters: number of repetitions: 100 bootstraps; pItem: 0.8; pFeature: 1; clusterAlg: "pam"; innerLinkage: average, finalLinkage: average, distance: Euclidean; lower limit of cluster number: 2; upper limit of cluster number: 12. CGIs with an arithmetic mean methylation level over 0.8 or less than 0.2 were excluded for clustering. Three clusters were selected as the best solution based on the consensus matrix. CGIs were annotated to overlap specific features (gene body, promoter, transcription factor binding site, H3K27ac, consensus PMDs, and HMDs) if either 20% of the CGI or 20% of the features were overlapping. The chromatin state of each CGI was determined by the largest overlap. CGIs with an arithmetic mean methylation level less than 0.2 or over 0.8 were designated as "Cluster low" and "Cluster high" clusters, respectively.

Analysis of Infinium HumanMethylation450 BeadChip

To verify our WGBS-based subtyping on global iCCA populations, we recruited 450K data of iCCAs from the studies by Goeppert and colleagues (11) and Jusakul and colleagues (12) in IDAT format. Raw data in IDAT format were preprocessed using the R package "ChAmp" (v 2.21.1). To validate the accuracy of WGBS-based subtyping, we performed consensus clustering on blocks with probes covering the top 1,998 variable blocks across iCCAs in this study by R package "ConsensusClusterPlus" (v1.54.0) with the same parameters described above, in which 269 of the top 1,998 variable blocks were involved. For each of these blocks, the arithmetic mean β value from probes within this block was used to represent the methylation level.

Tissue microarray experiment

iCCA samples from patients in the WCH-iCCAs ($n = 59$) were recruited for tissue microarray (TMA) construction, for which the tissue core punctured from a representative area of the formalin-fixed, paraffin-embedded (FFPE) slide of each iCCA sample was selected. For immunostaining, TMA slides were first deparaffinized in xylene and rehydrated by sequential incubation in EtOH/water solutions, followed by microwave antigen retrieval. After blocking endogenous peroxidase and nonspecific binding sites, the slides were incubated overnight at 4°C for anti-GBP4 primary antibody (Abcam, No. ab232693) according to the manufacturers' recommendation, and the corresponding secondary antibody and 3,3′-diaminobenzidine chromogens were applied to perform the staining. Slides were then digitally scanned in the Aperio AT system (Aperio, Leica Microsystems) and reviewed by two experienced pathologists who were blind to the clinical parameters. The expression of GBP4 was recorded from 5 representative areas at 20× magnification in each stained iCCA section by using a modified histologic score (H-score; ref. 20) based on the percentage of positively stained cells and the intensity of staining (with a maximum score of 300). The arithmetic mean value of the 5 areas was used for further analyses.

Cell viability and migration/invasion assays

Cell viability was examined by Cell Counting Kit-8 (CCK-8) assay (Beyotime Biotechnology). Detailed methodology initiated with seeding 1.5×10^3 cells suspended in 100 μL of complete culture media into 96-well plates, which was followed by 10 μL of CCK-8 solution added into each well at the indicated time point and incubated at 37°C for 2 hours. Absorbance detection of each well was performed using the EonTM Microplate Reader (BioTek) at a wavelength of 450 nm.

Migration/invasion assays were executed utilizing a transwell chamber (Corning Costar) coated with (invasion) or without (migration) Matrigel. For the migration assay setup, a cell population of 2×10^4 cells, commandeered in a volume of 500 μL of serum-depleted medium, was planted in the upper chamber in the insert of a 24-well plate. Correspondingly, for the invasion assays, 4×10^4 cells, suspended in the same serum-free medium volume, were placed in equivalent topographical conditions. To the lower chamber, a medium complemented with a 10% FBS concentration was introduced. After an incubation period of 24 hours, cells revealed to have migrated to the nethermost surface of the chamber were immobilized with paraformaldehyde and stained utilizing a crystal violet reagent.

Statistical analyses

All the statistical analyses were performed using R (v4.2.0). Mann–Whitney or Kruskal–Wallis test was used to compare quantitative data between groups. The chi-square or Fisher exact test was used to compare categorical data. Log-ranked tests were used to determine statistical significance between groups in all Kaplan–Meier (KM) curves. A multivariable Cox regression model was performed by integrating variables with statistical significance in univariable Cox

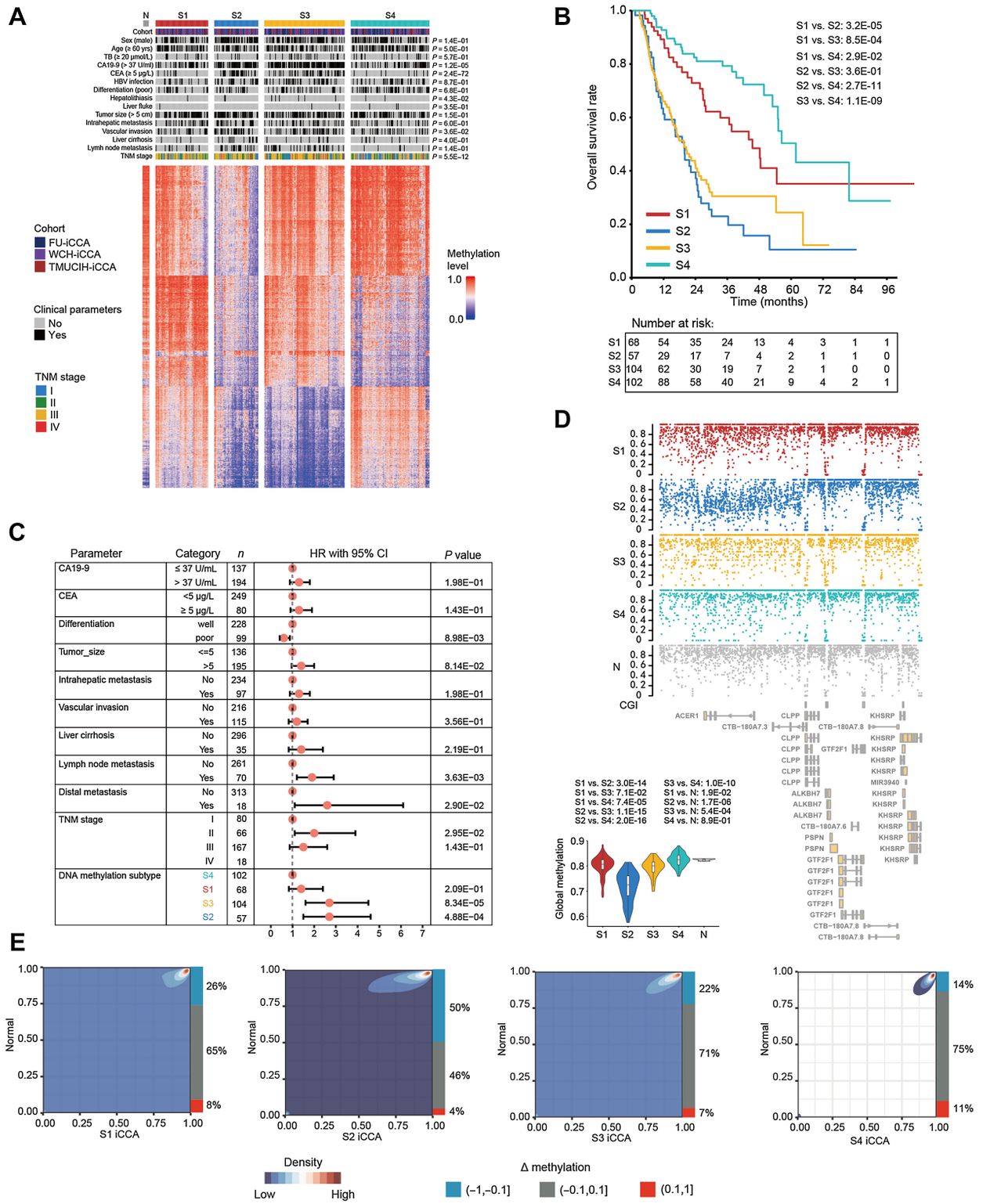


Figure 1.

WGBS-based multiomics profiling of iCCAs. **A,** Methylation subtype–based supervised hierarchical clustering of iCCAs ($n = 331$) and normal controls (N, $n = 9$) using the most variable methylation blocks (1,897 of 1,998) across iCCAs, along with the clinicopathologic features of each methylation subtype. It should be noted that 101 blocks identified in iCCA samples were not covered by WGBS data in normal controls. Thus, only commonly covered blocks ($n = 1,897$) by both iCCAs and normal controls were visualized in this figure. **B,** KM curve for the OS of each methylation subtype. (*Continued on the following page.*)

regression analysis. The threshold for statistical significance was defined as $P < 0.05$ or BH-adjusted $P < 0.05$.

Data availability

The previously published WES, RNA-seq, and proteome data analyzed in this study can be viewed in the biosino National Omics Data Encyclopedia (NODE) database at OEP001105. The 450K BeadChip-array data of iCCAs analyzed in this study were obtained from Gene Expression Omnibus at GSE201241 and GSE89803. The ChIP-seq data of HuCC-T1 analyzed in this study were obtained from Gene Expression Omnibus at GSE68388. The annotation of CGIs for hg19 was obtained from the UCSC Genome Browser at http://hgdownload.cse.ucsc.edu/goldenpath/hg19/database/. The BED file of the annotated TSSs was obtained from GENCODE (release 19) at https://www.gencodegenes.org/human/release_19.html. The 18-state chromatin segmentation data of adult human liver were obtained from ChromHMM at http://compbio.mit.edu/ChromHMM/. Annotations of consensus PMDs, HMDs, and solo-WCGW CpGs across various malignant tissues were obtained from Zhou and colleagues (21) at https://zwdzwd.github.io/pmd. The WGBS data of iCCAs generated in this study have been deposited in the National Genomics Data Center database at HRA004700. Access to these data will be granted upon completion of the following forms from applicants: (i) Ethical approval of the research project; (ii) a data sharing agreement for the project between the data owner (corresponding authors) and the PI/collaborator of the project. All other raw data generated in this study are available upon request from the corresponding author.

Results

iCCA cohorts and data resources

To characterize the DNA methylation landscape of iCCA at single-base resolution, WGBS was performed on 331 primary iCCA tumors from three independent centers, including Zhongshan Hospital of Fudan University (FU-iCCA, $n = 121$), West China Hospital (WCH-iCCA, $n = 154$), and Tianjin Medical University Cancer Institute and Hospital (TMUCIH-iCCA, $n = 56$). Meanwhile, nine healthy samples of cholangiocytes originating from the common bile duct were analyzed as the normal controls (Supplementary Fig. S1A). The clinicopathologic features of the iCCA patients are shown in Supplementary Table S1.

Paired-end sequencing generated a mean aligned sequencing depth of 21× in each sample after removing duplicate reads, which resulted in ∼22 million CpGs at 10× per sample across autosomes (Supplementary Table S2). Compared with the previously reported array-based methods (11, 12, 14), WGBS identified far more covered CpGs across samples. Of note, WGBS retained over 3 million CpGs, which were commonly captured at 10× by all the tumor and normal samples included. Therefore, the WGBS results might potentially provide more methylation information than the array-based methods. In addition, WES, RNA-seq and proteomics data on the same tumors in the FU-iCCA cohort (15) and previously published ChIP-seq data of iCCA cell lines (16, 17) were also incorporated in this study (Supplementary Fig. S1A and S1B).

WGBS profiling of iCCAs revealed methylation subtypes with loss of global methylation

To investigate the variation in DNA methylation patterns across iCCA individuals, we coalesced CpG methylations across the entire genome of each iCCA into blocks of homogeneously methylated CpGs. Unsupervised PCA based on WGBS blocks showed no obvious batch effect was observed among the three centers (Supplementary Fig. S1C). By *K*-means consensus clustering (Supplementary Fig. S2A–S2E) on the top variable blocks, four DNA methylation subtypes (S1–S4) with apparent intergroup heterogeneity were revealed, in which most of the blocks were hypomethylated in iCCAs when compared with normal controls (**Fig. 1A**; Supplementary Fig. S3 and Supplementary Table S3). This clustering output was different from the array-based clustering ($n = 3$) on 97 iCCAs by Jusakul and colleagues (12), in which even one cluster consisted almost entirely of patients with liver fluke, highlighting the necessity of DNA methylation subtyping at the genome-wide level. To validate our WGBS-based subtyping for the global iCCA populations, we reanalyzed the array-based (Infinium HumanMethylation450 BeadChip, 450K) DNA methylation data [from Jusakul and colleagues (12) and Goeppert and colleagues (11)] from 148 iCCA samples. It was found that 98.8% of 450K probes could be detected by WGBS even at a minimum coverage of 10, whereas most of the CpGs found by WGBS (98.0%) could not be targeted by 450K array (Supplementary Fig. S4A). Consensus clustering using 450K probes covering 269 of the top 1,998 variable blocks across iCCAs in this study also revealed four DNA methylation subtypes with apparent intergroup heterogeneity (Supplementary Fig. S5A–S5C), each of which could be classified into one of the WGBS-based methylation subtypes (Supplementary Fig. S5D–S5E), suggested the robustness of our subtyping methods. Remarkably, geographic analysis (Supplementary Fig. S5E) on the 450K iCCA cohort showed that most of the patients from East Asia (54 of 63) fell into the S2 subtype, whereas patients from Europe were equally distributed among S1, S3, and S4 subtypes. These findings indicated S2 as a specific subtype for East Asians and proved the sufficiency of our WGBS-based subtyping in covering global iCCA incidences.

Regarding clinical parameters, both S2 and S3 were featured by tumors with higher carbohydrate antigen 19-9 (CA19-9) level, carcinoembryonic antigen (CEA) level, and frequency of vascular invasion. Of note, a higher frequency of iCCA patients with hepatolithiasis was observed in S3 (5.77%) compared with the other three subtypes,

(*Continued.*) **C,** Multivariable regression hazard ratio (HR) forest plots for OS using clinical variables and methylation subtype based on Cox proportional hazards model (Wald test, two-sided, no adjustment for multiple comparisons). Pink dots, means; error bars, 95% confidence intervals. Only variables reaching statistical significance by univariable analysis ($P < 0.05$) were included in this analysis. **D,** Genome browser tracks for WGBS data of the representative iCCA of each methylation subtype and a representative normal control for an exemplary region (chr19: 6,282,123–6,425,048). The violin plot shows the global methylation levels averaged across all covered CpGs outside of CGIs per sample for iCCAs in each methylation subtype and normal controls. The top and bottom hinges of boxplots inside violins represent the 75th and 25th percentiles, respectively. The center line represents the median. The whiskers extend to the largest and smallest values within 1.5 times the interquartile range. The pairwise Mann–Whitney test *P* values are listed above the violin plot. **E,** Density plots showing comparisons of the mean methylation levels of CpGs commonly identified by iCCAs in each subtype and normal controls. The horizontal and vertical axes represent the mean CpG methylation levels in iCCAs and normal controls, respectively. Initially, these plots were dot plots showing the methylation of each CpG, in which each dot represented a CpG commonly identified. For the convenience of visualization, we converted the dot plots into gradient density plots, in which the density of CpGs (dots) in a region is represented by colors (blue, low density; red, high density). The bar plots next to density plots show the percentage of Δmethylation (iCCAs vs. normal controls) in each range.

despite the limited individuals with hepatolithiasis in this study. KM analysis showed that patients from S2 and S3 had significantly poorer overall survival (OS) than those from S1 and S4, with no statistical difference found between each other. Notably, patients in S4 had the best OS among the four subtypes, with a median OS of ∼60 months (**Fig. 1B**). Multivariable Cox analysis confirmed that our DNA methylation-based subtyping was a significant prognosticator independent of clinicopathologic features (**Fig. 1C**).

Global DNA methylation loss (excluding CpGs inside CGIs) has long been regarded as a characteristic feature of oncogenesis (22, 23), which has been highlighted by a previous pan-cancer study based on WGBS (5). Consistently, a representatively global methylation loss (excluding CpGs inside CGIs) was readily seen in S1–S3, with S2 showing the lowest methylation level among the four subtypes, whereas S4 demonstrated a methylation level comparable to that of normal controls (**Fig. 1D**; Supplementary Fig. S6A and S6B). Correlation analysis on normal-control samples and each iCCA subtype also showed a higher similarity in the global methylation level (excluding CpGs inside CGIs) between S4 and normal controls when compared with S1–S3, and only mild gain in CpG methylation (Δ methylation > 0.1) was observed across iCCA subtypes (**Fig. 1E**). These findings turned out to be in sharp contrast to the results from the array-based works claiming the symbolic hypermethylation states of iCCA (11, 14). One possible reason may be due to the variations as to the information of CpG sites captured by array-based methods and WGBS. Specifically, previous studies only analyzed the methylation states of ∼10,000 CpG sites (11, 14). In contrast, our analysis of global methylation states was based on information from about 20 million CpGs per sample (Supplementary Fig. S4B), which potentially provided more information and resulted in inconsistencies with the previous studies. To verify this hypothesis, we recalculated the global methylation level of each WGBS individual using CpGs overlapping 450K probes (Supplementary Fig. S4C), which inversely showed a significant trend toward hypermethylation in the majority of iCCAs when compared with normal controls, further proving the necessity of WGBS for an accurate evaluation of iCCA DNA methylation. Altogether, the WGBS revealed four methylation subtypes of iCCAs with clinical significance and that loss of global methylation (excluding CpGs inside CGIs) is a common feature among most of the iCCAs.

Characterizing the methylation pattern of each iCCA subtype

It has been reported that global loss of DNA methylation (excluding CpGs inside CGIs) in cancers tends to accumulate in PMDs, the genomic regions with incomplete loss of methylation, which are prone to CpGs lacking flanking CpGs, known as solo-WCGW CpGs (21). Contradictory findings indicate that hypomethylation of PMDs may confer a growth advantage to neoplastic cells, suggesting a potential mechanistic role in tumor progression, or conversely could merely be an epiphenomenon of oncogenesis, lacking causative significance in the pathophysiology of cancer (24, 25). Contrasted with PMD, HMD make up the remainder of the genome and help maintain the DNA methylation of the cell genome by "neighbor-guided correction" (26). In addition, aberrantly hypermethylated CGI has long been regarded as a symbolic epigenetic reprogramming in cancer (27). Typically, promoter CGIs act as an epigenetic control responsible for aberrant gene inactivation in cancer. Besides, the differentially methylated regions (DMR) between tumor and normal samples can provide information on possible functional regions involved in gene transcriptional regulation. Based on these facts, to comprehensively evaluate the DNA methylation pattern of iCCAs, 6 methylation types were summarized for each iCCA subtype (**Fig. 2A**) in this section, including methylation patterns in PMDs, solo-WCGW CpGs, HMDs, CGIs, hypermethylated DMRs (hyperDMRs, tumor vs. normal) and hypomethylated DMRs (hypoDMRs, tumor vs. normal).

First, to further characterize the global hypomethylation observed among the iCCA candidates, measurements were taken to assess the DNA methylation level in solo-WCGW CpGs, CGIs, and consensus PMDs/HMDs (only solo-WCGW CpGs were involved) obtained from the pan-cancer analysis (**Fig. 2B**; ref. 21). Overall, most iCCAs showed a decreased methylation level of solo-WCGW CpGs (**Fig. 2B**) but did not deviate much (Δ methylation ≤ 0.1) from the average of normal controls regarding solo-WCGW CpGs in PMDs or HMDs (**Fig. 2B**; Supplementary Fig. S7A). However, a pronounced trend of hypomethylation (Δ methylation > 0.1) in PMDs and a minor shift toward hypomethylation (Δ methylation ≤ 0.1) was found in HMDs for S2 iCCAs, respectively (Supplementary Fig. S7A). In concordance with these findings, S2 showed significantly lower methylation levels of solo-WCGW CpGs in both PMDs and HMDs when compared with other subtypes (**Fig. 2B**). When additionally analyzing solo-WCGW CpG methylation in PMDs across chromosome 16p, for example, we did observe that these CpGs were strongly hypomethylated in iCCAs compared with normal controls, with a sharp contrast in the hypomethylation level between S2 and other subtypes, and solo-WCGW CpGs of S2 in HMD regions were also prone to loss of methylation (Supplementary Fig. S7B). Therefore, these results suggested that methylation patterns of PMDs and HMDs are distinct on a genome-wide scale in each iCCA methylation subtype. Regarding the methylation level in CGIs, a focal hypermethylation state in CGIs was observed across all the iCCA subtypes (**Fig. 2B**; Supplementary Fig. S8). Meanwhile, we found that S1 and S4 exhibited a higher CGI methylation level than S2 or S3, and the normal tissues showed the lowest CGI methylation level (**Fig. 2B**). This was in line with the findings in most solid tumor types (5).

To globally assess the variability of PMDs across iCCAs, we defined PMDs for each sample using MethylseekR (see Supplementary Methods). Compared with normal controls, most of the iCCA samples showed a higher fraction of the genome harboring PMDs, especially for samples of the S2 subtype, with a median value over 50% (Supplementary Fig. S9A). Regarding the methylation level of PMDs, however, only S2 demonstrated significantly lower ($P < 0.001$) PMD methylation compared with normal controls (Supplementary Fig. S9B). Indeed, evaluation of the methylation of recurrent PMDs (rPMD; see Supplementary Methods) of each iCCA subtype also showed a globally lower methylation in rPMDs of S2 when compared with other iCCA subtypes (Supplementary Fig. S9C), which was mainly attributed to the lowest global methylation of S2 among these four subtypes.

Previous WGBS-RNA-seq integration analysis revealed the downregulation of genes when inside PMDs (28). With available gene-expression data at both mRNA and protein levels, we were able to determine the mean expression of genes as a function of PMD frequency through multiomics integration. As a consequence, there were numerous genes located in PMDs of iCCAs, with a median percentage of over 30%, and S3 iCCAs tended to have the most considerable number of genes within PMDs among the four methylation subtypes (Supplementary Fig. S10A). As expected, across all the iCCA subtypes, genes inside PMDs are expressed at consistently lower levels with higher variations than genes outside PMDs at both mRNA and protein levels (Supplementary Fig. S10B and S10C). Moreover, genes within PMDs showed a tendency toward lower expression in highly frequent PMDs (Supplementary Fig. S10D and S10E). These were consistent with the findings in breast cancer (28). Gene set enrichment analysis showed that genes downregulated when inside

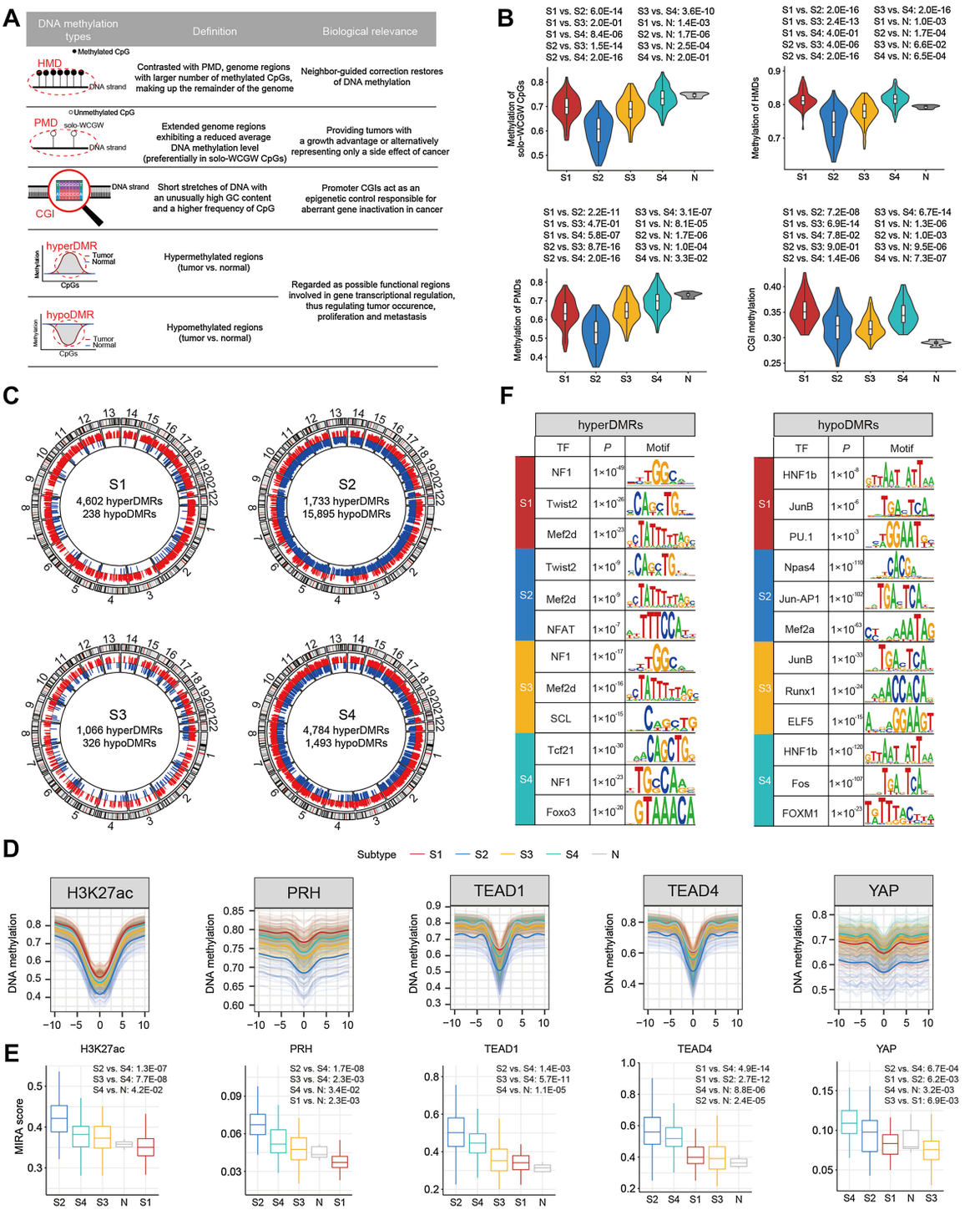


Figure 2.

Characterizing the methylation pattern of each iCCA subtype. **A,** An upfront list showing the methylation types that will be analyzed in this section. **B,** The methylation level of solo-WCGW CpGs, PMDs, HMDs, and CGIs in each iCCA subtype and normal controls. The top and bottom hinges of boxplots inside violins represent the 75th and 25th percentiles, respectively. The center line represents the median. The whiskers extend to the largest and smallest values within 1.5 times the interquartile range. The pairwise Mann–Whitney test *P* values are listed above the violin plot. **C,** Circo plots show the genome location of DMRs in each iCCA subtype. Red, hyperDMRs; blue, hypoDMRs. **D,** DNA methylation profiles for ChIP-seq (16, 17)–based transcription factor binding and H3K27ac sites from iCCA cell lines. N, normal controls. **E,** MIRA scores for transcription factor binding and H3K27ac sites in **D**. The top and bottom hinges represent the 75th and 25th percentiles, respectively. The center line represents the median. The whiskers extend to the largest and smallest values within 1.5 times the interquartile range. **F,** Enrichment of DMRs of each iCCA subtype for regulatory motifs. Shown are the top transcription factor binding site motifs using HOMER motif analysis (binomial *P* values). Motifs similar to previous (more significant) hits are not included.

PMDs were involved in the initial processes of cell development (Supplementary Fig. S10F and S10G), suggesting the participation of PMDs in downregulating genes related to iCCA carcinogenesis.

Strong methylation dip in putative regulatory regions of iCCAs

Next, local DNA methylation alterations for each iCCA subtype were evaluated by calling subtype-specific DMRs to identify the different methylation patterns between each iCCA subtype and normal controls (Supplementary Table S4A–S4D). We found that although most of the iCCAs were globally hypomethylated, hyperDMRs comprised the majority of DMRs in S1, S3, and S4, whereas S2 showed significantly more hypoDMRs (**Fig. 2C**). Enrichment analysis (Supplementary Fig. S11) showed the highest enrichment of hyperDMRs in CGIs and chromatin states predicted for adult liver [transcription start sites (TSS) in bivalent genes] in each iCCA subtype, further supporting the focal hypermethylation of iCCA in CGIs. Of note, hypoDMRs tended to enrich in H3K27ac-associated regions and transcriptional binding sites (TFBS) across the genome. Overall, these results confirmed the absence of strong hypomethylation in a major proportion of iCCAs (S1, S3, and S4) at the local level and indicated that DNA methylation changes across iCCAs affect similar types of regions.

A negative correlation of the DNA methylation level with transcription factor occupancy and regulatory activity at the binding sites has been demonstrated (29–32). To further investigate the local depletion of DNA methylation at putative regulatory regions identified in iCCA cell lines (16, 17), we calculated the DNA methylation inferred regulatory activity (MIRA) score (33) for each iCCA and normal control, with a high MIRA score reflecting a high regulatory activity of a given region set. A total of five markers (H3K27ac, PRH, TEAD1, TEAD4, and YAP) were enrolled for analysis as they were identified as putative regulatory regions by ChIP-seq (16, 17). For each iCCA, a strong methylation dip was observed at these sites when compared with the surrounding genome, suggesting the association of regulatory activities with hypomethylation at these sites (**Fig. 2D**). Although statistical significance was not always achieved, putative regulatory regions showed an inclination toward increased MIRA scores in iCCAs compared with normal controls. Moreover, comparison among iCCA subtypes revealed a subtype-specific transcription factor regulatory pattern, in which S2 and S4 generally showed higher MIRA scores than S1 and S3 (**Fig. 2E**). To further calculate the enrichment of DMRs on known transcription factor binding motifs, we performed motif analysis using HOMER (34). We found that the top shared motifs for hyperDMRs in each subtype included bHLH (such as Twist2/Tcf21) and CTF (such as NF1) transcription factors acting as tumor suppressors (35–38), whereas the shared top motifs for hypoDMRs in each subtype included bZIP transcription factors (such as JunB/Fos/Jun-AP1) acting as oncogenes (**Fig. 2F**; Supplementary Table S5A and S5B; refs. 39–41). The findings suggested hypomethylation of bZIP-binding sites as a symbolic biological feature of iCCA.

Clustering of CGIs revealing different CGI subgroups with biological relevance

Although iCCA is characterized by global methylation loss (excluding CpGs in CGIs), it was also notable that a positive association was revealed between averaged global and CGI methylation levels across iCCAs (**Fig. 3A**). In addition, the averaged CpG methylation level within CGIs was much lower than that of global methylation in each sample, which ranged from < 0.25, corresponding to the lowest global methylation, to about 0.5 for samples with the global methylation level reaching 0.8 (**Fig. 3A**). To depict the biological relevance of CGI methylation accumulation in iCCAs, we performed PCA on commonly covered CGIs ($n = 3{,}103$) across all tumor and normal-control samples, in which the averaged methylation value of CpGs within each CGI was used, showed a distribution of samples by CGI methylation and a continuous range from levels close to normal controls to more extreme hypermethylation, rather than a clear separation of normal controls from tumor samples or a clear separation of each methylation subtype from the others (**Fig. 3B**). PCA based on the methylation states of CGIs in each sample (the percentage of methylated CGIs, which was defined as > 0.2) also showed a similar distribution pattern (**Fig. 3C**), suggesting the accumulation of CGI hypermethylation across iCCA subtypes. To inspect these findings further, we then performed a consensus clustering based on the commonly covered, variable CGIs across tumors and normal controls ($n = 805$). CGIs with an arithmetic mean methylation over 0.8 or less than 0.2 were excluded for clustering, as these CGIs contributed little to the variations. Finally, five CGI clusters were identified (**Fig. 3D**; Supplementary Fig. S12A), including three clusters identified by consensus clustering (Clusters 1–3) and two clusters not recruited for consensus clustering, which were designated as Cluster low (CGIs with an arithmetic mean methylation < 0.2) and Cluster high (CGIs with an arithmetic mean methylation > 0.8). Cluster low mostly consisted of unmethylated CGIs (methylation level < 0.2) across iCCAs and normal controls, whereas both Clusters 1 and 2 mostly consisted of unmethylated CGIs in normal controls, and heterogeneous and sample-specific methylation states in the iCCAs. In contrast, CGIs of Cluster 3 and Cluster high showed a relatively homogeneous gain of methylation across iCCAs and normal controls, tending to be fully methylated in almost all samples (**Fig. 3D**; Supplementary Fig. S12B).

To elucidate the outcomes of CGI consensus clustering, we analyzed the overlap of CGIs with genomic features and annotated functional elements in each cluster (Supplementary Table S6), specifically for those exhibiting sustained low and high levels of methylation (Cluster low and Cluster high). We found that the fraction of CGIs in gene bodies increased along with the elevated methylation across clusters (from Cluster low to Cluster high) and constituted the highest fraction of CGIs in each cluster, which is presumably owing to the hypermethylation of intragenic CGIs as a common mechanism regulating transcriptional activity in healthy and tumor cells (42–45). On the other hand, the fraction of CGIs overlapping promoters, enhancer sites (marked by H3K27ac ChIP-seq peaks; refs. 16, 17), and transcription factor binding sites decreased along with the elevated methylation across clusters (from Cluster low to Cluster high; **Fig. 3E**), suggesting the stable maintenance of hypomethylation at these transcription-regulating sites in normal and iCCA samples. Although the fractions of CGIs overlapping consensus HMDs remain stable across clusters, the fraction of CGIs overlapping consensus PMDs was significantly higher in Clusters 1–3 and Cluster high than in Cluster low, which was consistent with the finding that CGI methylations preferentially accumulate in PMDs (28). At the same time, the relatively stable frequency of HMD CGIs across clusters also suggested no preference for CGI hypermethylation in HMD regions.

As for the chromatin states of CGIs in each cluster (**Fig. 3F**), we found that the fraction of active and bivalent TSSs descended from 97.5% (Cluster low) to less than 2% (Cluster high). On the contrary, the fraction of CGIs overlapping repressive, heterochromatin, and quiescent regions increased along with the elevated methylation across clusters (despite a decreased fraction in Cluster high when compared with Cluster 3), suggesting a trend of CGI hypermethylation in regions with low transcriptional activity. Indeed, genes associated with promoter CGIs in Cluster high were mostly silenced at both RNA and

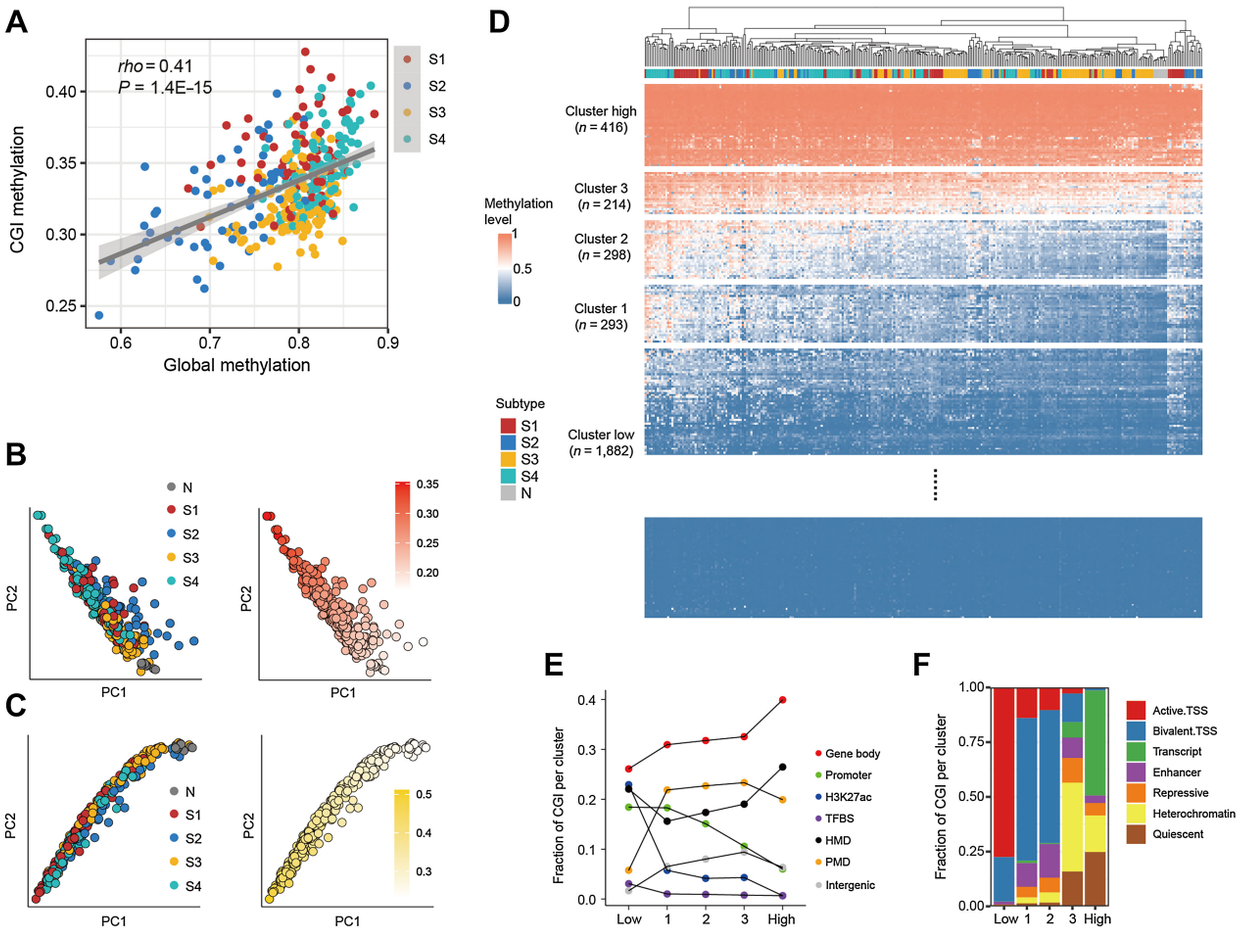


Figure 3.

CGI methylation patterns across iCCAs. **A,** Correlation between global methylation (excluding CpGs in CGIs) and CGI methylation levels across iCCAs. **B,** PCA based on the methylation of the commonly covered CGIs ($n = 3{,}103$) of iCCAs included in this study. **C,** PCA based on the percentage of methylated CGIs (defined as > 0.2) in the commonly covered CGIs ($n = 3{,}103$) of iCCAs included in this study. **D,** Hierarchical clustering of the five clusters of CGIs identified across iCCAs. **E,** Fraction of CGIs per cluster overlapping annotated functional elements/genome regions. **F,** Fraction of CGIs per cluster overlapping chromatin states.

protein levels, whereas genes in Cluster low remained actively expressed (Supplementary Fig. S12C and S12D). However, the inconsistency between the RNA and protein levels of gene-expression transition across CGI clusters was found, suggesting the influence of posttranscriptional regulations. Moreover, genes associated with promoter CGIs in Cluster low and Cluster 1, which were mostly methylated in S1 and S4, were generally enriched in pathways related to cell maintenance, such as cell division, mRNA splicing, differentiation, and translation, which clarified the reason for their unmethylation states (Supplementary Fig. S12E and S12F).

Taken together, these findings indicated that the unmethylation states of CGI are essential for the transcription activity of genes responsible for cell maintenance in both iCCA and normal samples.

The relationship of iCCA methylation subtyping with covariates

To explore the association of covariates with different iCCA subtypes, we performed multiomics integration to identify the genome and molecular features underlying methylation subtyping. In our previous work based on a Chinese iCCA cohort consisting of 262 patients (15), we have found somatic copy-number variations (SCNV) and 16 significantly altered genes with driver potential in iCCAs (15). The landscape of genetic alterations in this cohort was generally consistent with previous reports based on White, Asian, and Black or African populations (12, 14, 46, 47). Of note, this cohort also showed statistically higher KRAS mutation frequency and lower IDH1, ARID1A, and TERT mutation frequencies than the cohort of Farshidfar and colleagues (14, 15), which is composed of Western iCCA candidates, suggesting the geographical difference in genomic alterations between Western and Eastern populations. Then, we mapped these events to this study, the WES data of which were available in 93.4% (113 of 121) of FU-iCCAs. It was shown that S2 demonstrated the highest tumor mutation burdens and tumor neoantigen burdens among the four methylation subtypes, whereas S4 was characterized by the highest SCNV burdens (**Fig. 4A**). Furthermore, genetic alterations closely relating to iCCAs carcinogenesis (48) appeared to be subtype-specific (Supplementary Tables S7–S9). Specifically, consistent with the clustering results by multiomics integration (11), we also found a methylation subtype (S1) that nearly all the samples with IDH (IDH1 or IDH2) mutations (17/18 samples with mutation) were classified into, and these samples comprised almost all the S1 samples (17/18 S1

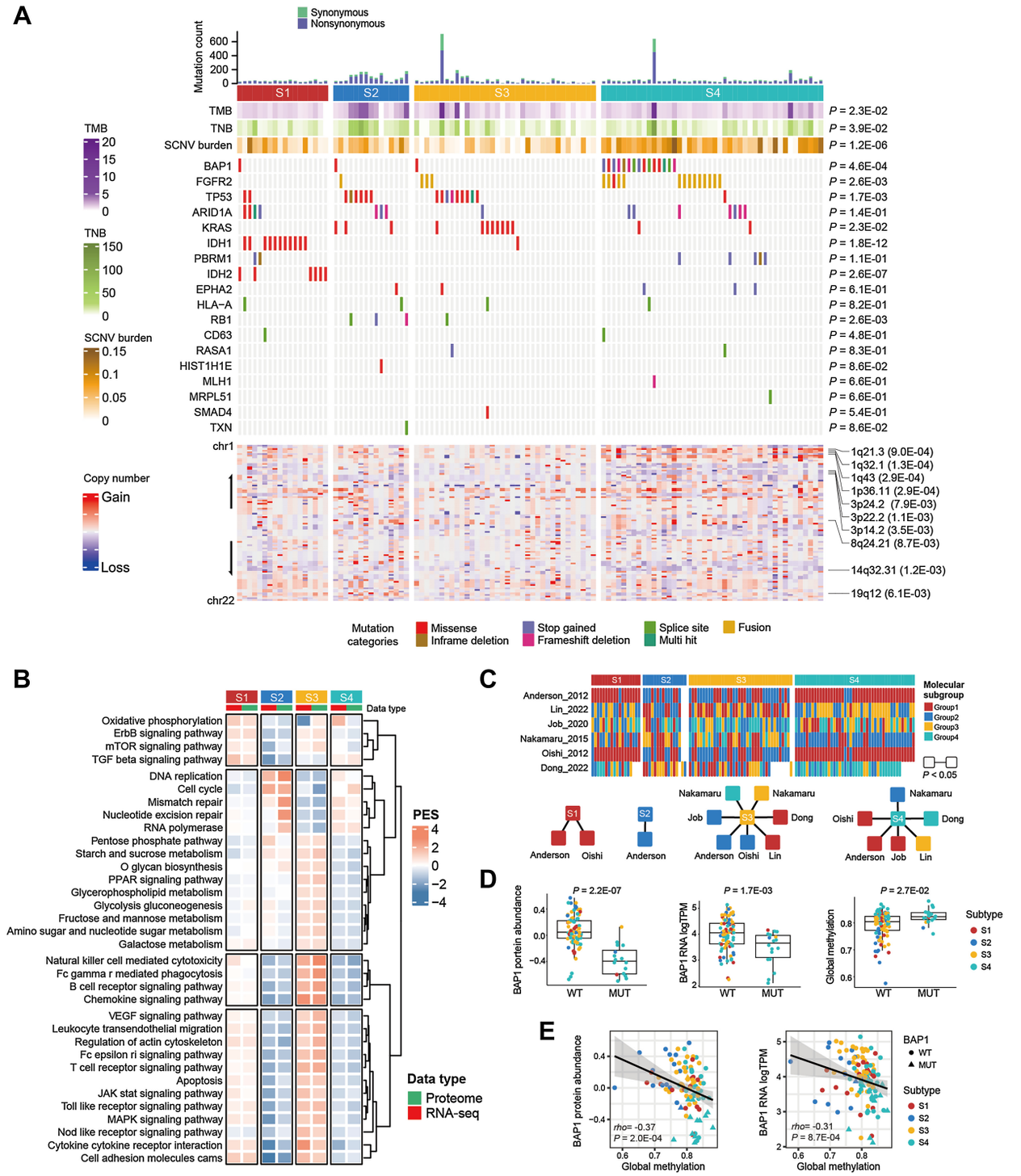


Figure 4.

The relationship of iCCA methylation subtyping with covariates. **A,** Genetic profile of different iCCA subtypes in the multiomics cohort (FU-iCCA). Only SCNVs with chi-square or Fisher exact *P* values < 0.01 are shown. **B,** Integrated analysis of altered pathways at transcriptomics and proteome levels among the four methylation subtypes in the multiomics cohort. **C,** Comparisons of our methylation subtyping with previous biliary cancer-related molecular (mRNA or protein) signatures. **D,** Comparisons of BAP1 protein abundance (left), mRNA level (middle), and global methylation level (right) between BAP1 mutation and wild-type samples in the multiomics cohort. The top and bottom hinges represent the 75th and 25th percentiles, respectively. The center line represents the median. The whiskers extend to the largest and smallest values within 1.5 times the interquartile range. **E,** Correlation between BAP1 expression (protein and mRNA) and global methylation in the multiomics cohort.

samples). Interestingly, in the work by Goeppert and colleagues (11), the IDH-mutation subgroup had a modest OS outcome, which mirrors the OS of S1 among our iCCA methylation subgroups. S2 was characterized by the highest proportion of samples with TP53 alterations (5/15 S2 samples), which showed the worst OS and the lowest global methylation level. In contrast, the mutations of KRAS, which were mutually exclusive with IDH1/2, BAP1, and FGFR2 alterations (15), preferentially occurred in S3 (8/13 samples with mutation), the prognosis of which was also poor. In addition, most of the samples with BAP1 alterations (15/18 samples with mutation) or FGFR2 fusions (13/17 samples with fusions) were classified into subtype S4, which had the most favorable prognosis and the highest global methylation level. Meanwhile, amplification of known oncogenes (such as MCL1 and MDM4, located in 1q21.3 and 1q32.1, respectively) and deletions of tumor suppressors (such as ARID1A and MLH1, located in 1p36.11 and 3p22.2, respectively) were dominantly presented by S4 samples.

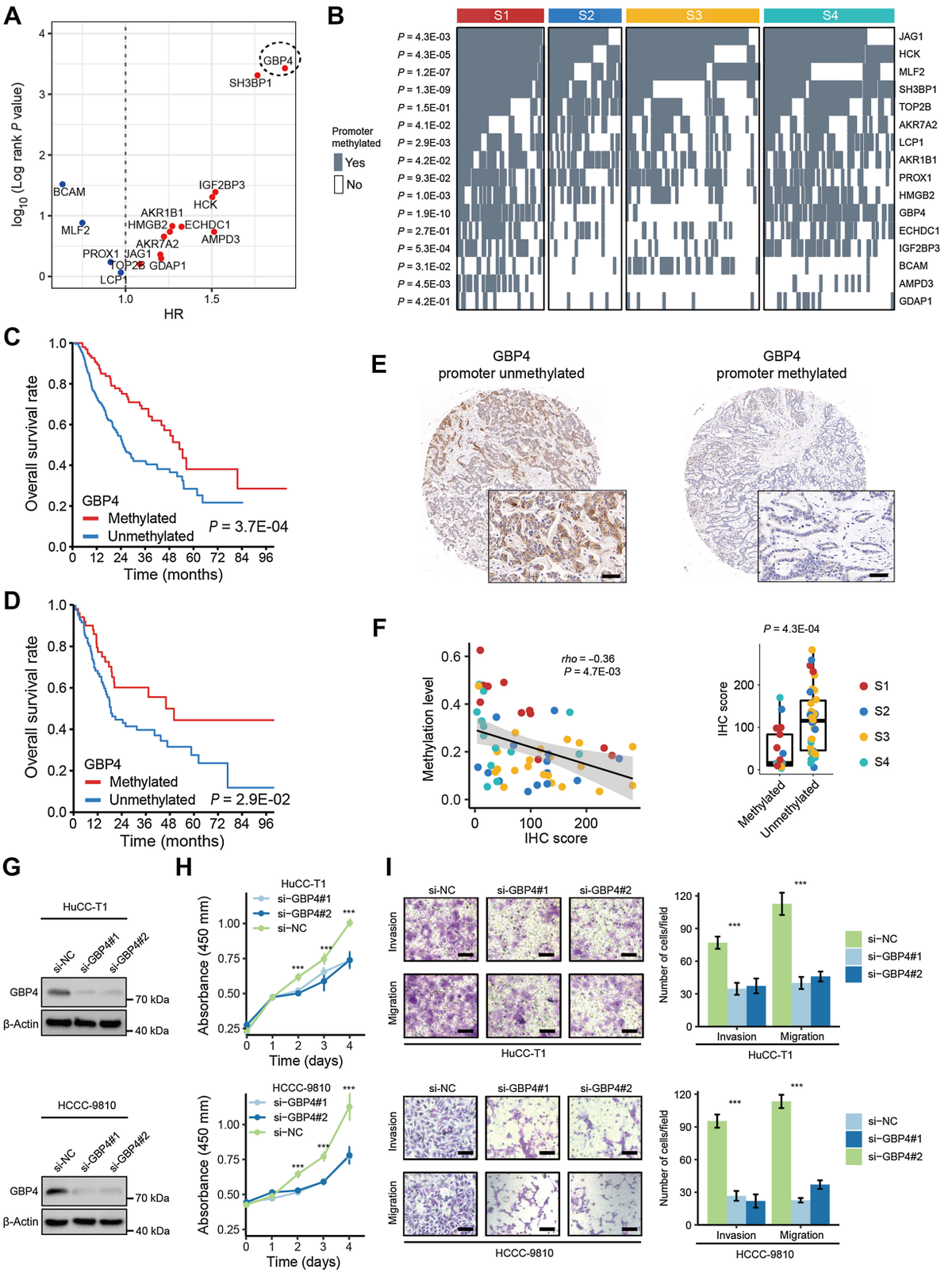
Next, to investigate the molecular aberrations related to different methylation subtypes, we conducted a pairwise analysis comparing the gene-expression profiles between each methylation subtype with the others using both RNA-seq and proteomic approaches (**Fig. 4B**). S1 reached the highest enrichment of TGFβ signaling, ErbB signaling, and mTOR signaling, whereas S2 had the highest enrichment of cell cycle, DNA replication, and transcription. S3 was characterized by the highest enrichment of various innate/adaptive immunities, carbohydrate metabolism, and cytoskeleton regulation/ECM interaction. As expected, subtype S1, which contained a higher frequency of IDH mutation, showed enhanced dependence on oxidative phosphorylation (49–51). Although S1 showed a higher enrichment of oxidative phosphorylation than S2 and S3 at both mRNA and protein levels. However, this pathway was also found to be enriched in S4 at the mRNA level, which was potentially attributed to the highly prevalent BAP1-mutant samples. This was supported by existing evidence indicating the contributory role of BAP1 mutations in the activation of oxidative phosphorylation (52, 53).

Furthermore, we compared our DNA methylation subtyping with previously described molecular subtypes of iCCAs (**Fig. 4C**; 15,46,54–57). Besides the signature for survival outcomes (55), no statistically significant overlap was found between molecular signatures and S2, and only the signature for mature hepatocyte (MH) was found to be significantly overlapped with S1. On the contrary, both S3 and S4 had significant overlap with various molecular signatures. S3 was characterized by hepatic-stem-cell (HpSC) iCCAs with poor survivals, whereas S4 was characterized by MH iCCAs with good survivals and cellular differentiation signatures. Moreover, both S3 and S4 demonstrated a high intersample heterogeneity in tumor immunities, characterized by the large proportion of immunogenic/immune-actiavted (Job_2020, Group2 and Lin_2022, Group 3, respectively) and immune-suppressive/immune desert (Job_2020, Group1 Lin_2022, Group 1; refs. 54, 56) samples presented simultaneously. Further deconvolution analysis also exhibited that, despite a stronger enrichment of various innate and adaptive immune cell types compared with other subtypes, a greater intersample heterogeneity of immune infiltration was found among S3 samples (Supplementary Fig. S13A–S13C). Concomitantly, the analysis of immune checkpoints (IC) showed that most ICs reached statistical significance in the comparisons of methylation subtypes (Supplementary Fig. S13D), suggesting the potential options of IC inhibitor therapies based on DNA methylation subtyping. It was also notable that both S2 and S3 subtypes have frequent TP53 mutations. Accordingly, we also performed DMR analysis (the same methodology as used for iCCA subtype-normal control comparison) between the S2 and S3 subtypes. Despite the identification of limited DMRs (Supplementary Table S10), we observed a DMR located in the 15th exon of OBSCN (Supplementary Fig. S14A), a well-known tumor suppressor commonly mutated in various cancer types (58). Although no significant difference in OBSCN mRNA was found between S2 and S3 (Supplementary Fig. S14B), there was a positive correlation between OBSCN expression and methylation level in the DMR region (Supplementary Fig. S14C), suggesting the *cis* association of exon methylation and OBSCN gene expression.

As a tumor suppressor gene, BAP1 exerts tumor-suppressive functions on cell-cycle control, DNA damage repair, and differentiation (59). A previous study on uveal melanoma proved that BAP1 loss is related to DNA methylation repatterning (60). Because BAP1 alterations were found to be S4-specific (**Fig. 4A**) and negatively correlated with gene expression at both RNA and protein levels (**Fig. 4D**), we sought to explore the effects of BAP1 on global and CGI methylations, which showed that BAP1 expression at the RNA or protein level negatively correlated global methylation of iCCAs, while exhibited negative association with CGI methylation only at the RNA level (**Fig. 4D** and **E**; Supplementary Fig. S15A). Nevertheless, no significant DMR was found between BAP1-mutation and wild-type samples (same methodology as used for iCCA subtype-normal control comparison), which was further supported by the high similarity of global methylation pattern between these two groups (Supplementary Fig. S15B). Unsupervised hierarchical clustering based on common methylation blocks of FU-iCCAs also showed that BAP1-mutation samples merely clustered with S4 samples sporadically, rather than being remotely separated from wild-type iCCAs (Supplementary Fig. S15C), suggesting the BAP1-mediated DNA hypermethylation as the symbolic features of S4. Furthermore, we also analyzed the gene-expression pattern between BAP1-mutation and wild-type iCCAs, which confirmed the correlation between BAP1 alteration and downregulated activity of carbohydrate metabolism (Supplementary Fig. S15D), indicating the involvement of BAP1 loss in one-carbon metabolism and/or related nutrient pathways inducing DNA hypermethylation (59).

Epigenetically silenced genes with prognostic and therapeutic implication

To understand which genes underwent transcriptional repression and consequential downregulation of protein abundance in association with methylation changes, we performed an integrative analysis involving WGBS, RNA-seq, and proteomics data to evaluate the inner interactions in DNA methylation/mRNA/protein (see Supplementary Methods). Comparison of genes/proteins with both methylation–mRNA and methylation–protein correlation data showed that the performances of methylation changes on mRNA-abundance and protein-abundance predictions were deemed equivalent (median Spearman rho $= -0.045$ and -0.051, respectively). To assess the disagreement in the effect of methylation changes between transcriptomics and proteomics data, we calculated a reversion potential, corresponding to the difference between Spearman coefficients (see Supplementary Methods). A high reversion potential suggests genes being transcriptionally silenced but buffering at the protein level. Generally, 2,989 genes were reversed at the protein level (2,885 lowly reversed and 104 highly reversed), corresponding to 38.0% of the 7,857 genes analyzed (Supplementary Fig. S16A). These reversed proteins were characterized by the highest enrichment of metabolism pathways (Supplementary Fig. S16B), suggesting the high prevalence


A
log10 (Log rank P value)
HR
GBP4
SH3BP1
BCAM
IGF2BP3
HCK
AKR1B1
HMGB2
ECHDC1
MLF2
AMPD3
AKR7A2
PROX1
JAG1
TOP2B
GDAP1
LCP1
B
S1
S2
S3
S4
Promoter methylated
Yes
No
P = 4.3E-03
P = 4.3E-05
P = 1.2E-07
P = 1.3E-09
P = 1.5E-01
P = 4.1E-02
P = 2.9E-03
P = 4.2E-02
P = 9.3E-02
P = 1.0E-03
P = 1.9E-10
P = 2.7E-01
P = 5.3E-04
P = 3.1E-02
P = 4.5E-03
P = 4.2E-01
JAG1
HCK
MLF2
SH3BP1
TOP2B
AKR7A2
LCP1
AKR1B1
PROX1
HMGB2
GBP4
ECHDC1
IGF2BP3
BCAM
AMPD3
GDAP1
C
Overall survival rate
GBP4
Methylated
Unmethylated
P = 3.7E-04
Time (months)
D
Overall survival rate
GBP4
Methylated
Unmethylated
P = 2.9E-02
Time (months)
E
GBP4
promoter unmethylated
GBP4
promoter methylated
F
rho = −0.36
P = 4.7E-03
Methylation level
IHC score
P = 4.3E-04
IHC score
Methylated
Unmethylated
S1
S2
S3
S4
G
HuCC-T1
si-NC
si-GBP4#1
si-GBP4#2
GBP4
β-Actin
70 kDa
40 kDa
HCCC-9810
si-NC
si-GBP4#1
si-GBP4#2
GBP4
β-Actin
70 kDa
40 kDa
H
HuCC-T1
si-GBP4#1
si-GBP4#2
si-NC
Absorbance (450 mm)
Time (days)
HCCC-9810
si-GBP4#1
si-GBP4#2
si-NC
Absorbance (450 mm)
Time (days)
I
si-NC
si-GBP4#1
si-GBP4#2
Invasion
Migration
HuCC-T1
Number of cells/field
Invasion
Migration
HuCC-T1
si-NC
si-GBP4#1
si-GBP4#2
si-NC
si-GBP4#1
si-GBP4#2
Invasion
Migration
HCCC-9810
Number of cells/field
Invasion
Migration
HCCC-9810
si-NC
si-GBP4#1
si-GBP4#2

of other mechanisms rescuing metabolism protein abundance from methylation changes in iCCAs.

Given that cancer proteome more closely links genotype to phenotype than mRNA, we set a stringent criterion to identify significantly epigenetically silenced genes in iCCAs using both RNA-seq and proteomics data (see Supplementary Methods), resulting in 16 genes with significantly downregulated protein abundance in relation to increased methylation (Supplementary Table S11). Among these genes, the epigenetic silencing of GBP4, which belongs to a family of GTPases induced by interferon-gamma (61), demonstrated the most significant prognostic implication (**Fig. 5A**) and the most variable prevalence among methylation subtypes (**Fig. 5B**). Thus, we focused on GBP4 methylation and its biological relevance. Previous studies revealed significant but heterogeneous correlations between GBP4 expression and survival outcomes in various cancer types (61). To recognize the potential role of GBP4 methylation in iCCA, we first analyzed the association of methylation states in GBP4 ("silenced" or "not silenced") with the survival outcomes of patients in this study using KM curve, proving that cases with silenced GBP4 had better survival outcomes (log-rank $P = 0.00037$), which was also validated by the array-based cohort of Jusakul and colleagues (log-rank $P = 0.029$; ref. 12), suggesting GBP4 methylation as an epigenetic prognostic indicator for favorable survival in iCCAs (**Fig. 5C** and **D**). Moreover, IHC staining also showed that the GBP4-methylated samples had a significantly lower level of GBP4 expression than those unmethylated ones (**Fig. 5E** and **F**). Tumors with demethylated GBP4 were significantly associated with higher expression of genes involved in carbohydrate metabolism, drug metabolism, ECM interaction, bile acid/fatty acid synthesis, and various innate/adaptive immune pathways (Supplementary Fig. S17A and S17B). Further function analysis demonstrated that silencing the expression of GBP4 significantly inhibited iCCA cell proliferation, migration, and invasion (**Fig. 5G**-**I**). These results implied that the demethylation of GBP4, as a robust prognostic biomarker, functions as a strong oncogenic factor for iCCA.

Discussion

By integrating WGBS with multiomics approaches, including genetic, transcriptomic, and proteomic analyses, our study has demonstrated that iCCAs can be classified into four DNA methylation subtypes, each of which exhibits distinct biological features and clinical outcomes (**Fig. 6**). iCCA S1, phenocopying the IDH-mutation group (11), is almost completely composed of patients with IDH mutation, which accounts for ∼15% of the iCCA population worldwide (48), suggesting patients of this subtype as ideal candidates for IDH inhibitors (e.g., ivosidenib; ref. 62). iCCA S2 is characterized by the lowest methylation level and the highest mutation burden among the four subtypes, and a gene-expression pattern associated with cell cycle/DNA replication, whereas iCCA S3 is characterized by considerably high interpatient heterogeneity of tumor immunity, a gene-expression pattern associated with carbohydrate metabolism and a preference for KRAS alterations, which closely associates with the response of anti–PD-1 therapy synergistically enhanced by interleukin-1 receptor antagonist (IL1 Ra; ref. 63). Based on the results from *in vitro* function analyses, the high frequency of GBP4 demethylation in S2 and S3 also highlights the urgent need for future anti-GBP4 treatment, and the poor survival of S2 and S3 also raises the necessity of combined surgical-adjuvant therapies for these two subtypes. Although characterized by the highest SCNV burdens, the fourth subtype, iCCA S4, has the best survival benefits from surgical resection and shows a preference for FGFR2 fusions/BAP1 mutations, suggesting them as ideal candidates for FGFR inhibitors (e.g., pemigatinib; ref. 64). Moreover, the statistically significant overlap of S3/S4 with proteomics subtypes (Dong_2022; ref. 15) also indicates the sensitivity of these two subtypes to EGFR inhibitors (S3) and gemcitabine/paclitaxel (S4). Despite interpatient heterogeneity in tumor immunity, the major proportion of immunogenic/immune-activated patients in the S3 and S4 subtypes suggests that it is worth trying immunotherapies on these two subtypes. Meanwhile, the controversy between the superior survival and high frequency of genome alterations (including somatic mutations and SCNV in oncogenes and tumor suppressors) in S4 also suggests the mightiness of epigenetic mechanisms rescuing patients from the oncogenic forces by genome alterations.

The global CGI methylation gain across most tumors in this study suggests a CGI methylator phenotype (CIMP) present in iCCAs. Indeed, hierarchical clustering on common CGIs across iCCAs (**Fig. 3F**) shows a small proportion of S4 iCCAs demonstrates hypermethylation in all of these CGIs, denoting a CIMP of these patients. Traditional methods for characterizing CIMP rely on DNA methylation levels of a panel of marker genes or variable CpG probes on the Illumina Infinium Array (usually around 1,000 CpGs), which do not consider the methylation levels of genome-wide CGIs and therefore have limited interpretability (65, 66). Presently, only a few studies have reported the genome-wide identification of CIMP using WGBS data in cancer populations (7, 67). However, the definitions of CIMP-positive or -negative tumors in these studies remain elusive, as it is based on a small fraction of CpGs in a limited set of CGIs defined explicitly for a specific population. Thus, further pan-cancer analyses on the whole-genome level are still needed to identify consensus CIMP CGIs and explore their biological relevance.

Although gene-expression regulation by DNA methylation has been thoroughly investigated by combined transcriptomics–methylome analysis in various cancer types (4, 7, 44), limited studies have reported the effect of DNA methylation alteration on consequent global protein abundance and posttranslational modifications, especially for patients with iCCA. To our knowledge, this study represents the first attempt to examine the correlation between proteomics and DNA methylation in iCCAs. By integrated methylome–transcriptomics–proteomics analysis, we identified the significantly positive correlation between GBP4 promoter demethylation and protein abundance, which has both prognostic and therapeutic implications.

Figure 5.

Identifying epigenetically silenced genes with prognostic and therapeutic implications. **A,** Dot plot shows the prognostic analysis results of the epigenetically silenced genes identified. **B,** List of epigenetically silenced genes across iCCAs. *P* values (listed on the left side) were calculated using the Chi-square or Fisher exact test comparing the difference among iCCA subtypes. Each square represents the methylated states (methylated or unmethylated, cutoff = 0.3) of the gene promoter in this iCCA sample. **C,** KM curve for the OS of GBP4-methylated and -unmethylated samples in this study. **D,** KM curve for the OS of GBP4-methylated and -unmethylated samples in an array-based external validation cohort (12). **E,** Representative IHC staining of GBP4 in GBP4-methylated and -unmethylated samples ($n = 59$). **F,** Correlation between GBP4 promoter methylation and IHC score ($n = 59$). **G,** Western blot of GBP4 in iCCA cells transfected with GBP4 siRNA. **H,** CCK-8 assays show that silencing of GBP4 inhibited the proliferation of iCCA cells. **I,** Transwell Matrigel invasion and migration assay performed in GBP4-silenced iCCA cells. Representative images (left) and statistical analyses (right) of the migrated/invaded cells are shown. Scale bars, 200 μm. Data represent the mean ± SD and are representative of 5 independent experiments (one-way ANOVA). ***, $P < 0.001$; NC, normal control.

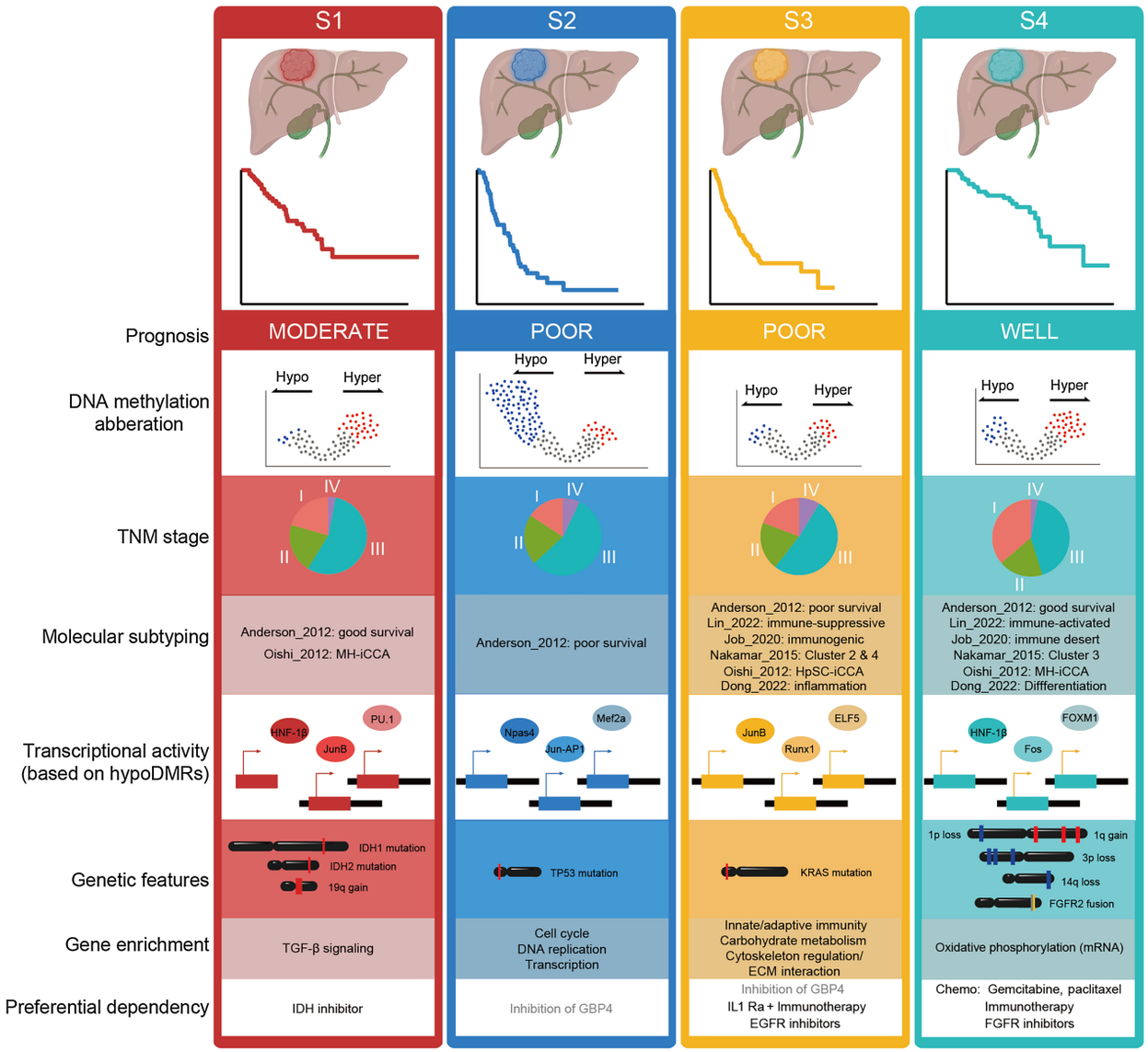


Figure 6.
Overview of biological features of each iCCA methylation subtype.

Additionally, the prognostic value of other GBP family members, such as GBP1, GBP2, and GBP5, has also been investigated, in which there is also a disparity in the correlations between protein abundance and survival outcomes (61). However, this disparity is largely unknown, the explanation for which might be the distinct cellular mechanisms used under different conditions (61). For example, GBP1 can prevent cytoskeletal polymerization by binding to F-actin and may initiate a resistance pathway to paclitaxel in ovarian cancer (68, 69), but suppresses matrix metalloproteinase 1 expression and thus inhibit glioblastoma metastasis (70). In this study, the correlation between GBP4 demethylation and upregulated genes involved in various metabolism pathways, including carbohydrate metabolism, lipid metabolism, amino acid metabolism, and glycan biosynthesis, suggests the oncogenic role of GBP4-dependent metabolism reprogramming in iCCAs, which allows cancer cells to adapt to drastic changes in the tumor environment, thus initiating the process of transformation and promoting the growth of malignant cells (71). On the other hand, GBP4 also serves as an immune-related biomarker predicting prognoses of cancer. For example, GBP4 is implicated in the orchestration of tumor immunity of colorectal cancer, wherein it symbolizes a proinflammatory microenvironment (72, 73). More importantly, GBP4 plays a pivotal role in mediating type-I interferon responses and demonstrates a notable positive association with macrophage infiltration (74, 75). These facts suggest an immune-activated microenvironment in iCCAs with demethylated GBP4, as evidenced by the enrichment of various innate/adaptive immune pathways (Supplementary Fig. S17), which inspires the adoption of immunotherapies for patients with demethylated GBP4. However, the "immune desert" states of subtype S2 compared with S3 (as shown in **Fig. 4B**; Supplementary Fig. S13A) highly recommend

consideration given to these patients' DNA methylation subtyping when applying immunotherapies.

In summary, this study constitutes the inaugural effort to integrate WGBS with an expansive range of multiomics data in the context of iCCA. We identified major molecular mechanisms that are affected by alterations in genome-wide DNA methylation. These insights hold the potential for guiding the development of innovative, patient-tailored therapeutic strategies for iCCA, ultimately benefiting clinical practice. Moreover, the oncogenic GBP4 demethylation may inspire further studies on the metabolism pattern of iCCA from the epigenetics perspective. In addition, we identified subtype-specific blocks with cross-validated high accuracy as symbolic and handful biomarkers to classify iCCAs into one of the four methylation subtypes (Supplementary Fig. S18A–S18D; see Supplementary Methods), laying foundations for developing targeted methylation assays for future noninvasive iCCA subtyping using small amount of tumor DNA, such as cell-free DNA (76) and circulating tumor DNA (77).

Authors' Disclosures

H. Jiang reports grants from the National Natural Science Foundation of China and China Postdoctoral Science Foundation outside the submitted work. No disclosures were reported by the other authors.

Authors' Contributions

H. Liao: Resources, data curation, supervision, funding acquisition, investigation, visualization, methodology, project administration, writing–review and editing. **X. Chen:** conceptualization, resources, supervision, methodology, writing–original draft, writing–review and editing. **H. Wang:** Data curation, supervision, funding acquisition, methodology, writing–original draft, writing–review and editing. **Y. Lin:** Resources, data curation, formal analysis, investigation, Methodology. **L. Chen:** Conceptualization, resources. **K. Yuan:** Conceptualization, resources, supervision. **M. Liao:** Resources, investigation. **H. Jiang:** Supervision, writing–review and editing. **J. Peng:** Methodology. **Z. Wu:** Resources. **J. Huang:** Conceptualization, resources, supervision, funding acquisition. **J. Li:** Conceptualization, resources, supervision, methodology. **Y. Zeng:** Conceptualization, resources, supervision, funding acquisition.

Acknowledgments

This work was supported by grants from the National Multidisciplinary Collaborative Diagnosis and Treatment Capacity Building Project for Major Diseases (TJZ202104), the Natural Science Foundation of China (82173248, 82373400, 82372660, 82272685, 82002967, and 82202260), the Postdoctoral Science fellowship of China National Postdoctoral Program for Innative Talents (BX20200225), the Project funded by China Postdoctoral Science Foundation (2022TQ0221), the Science and Technology Major Program of Sichuan Province (2022ZDZX0019), the Sichuan Science and Technology Program (2023YFS0128, 2023NSFSC1874, and 2021YJ0420), 1.3.5 project for disciplines of excellence, West China Hospital, Sichuan University (ZYJC18008 and ZYGD22006), the Sichuan University postdoctoral interdisciplinary Innovation Fund (10822041A2103). We are most grateful to West China Biobanks, Department of Clinical Research Management, West China Hospital, Sichuan University for their support of human samples. We also thank Prof. Qiang Gao [Department of Liver Surgery and Transplantation, Liver Cancer Institute, Zhongshan Hospital, and Key Laboratory of Carcinogenesis and Cancer Invasion (Ministry of Education), Fudan University] and Dr. Dongming Liu (Department of Hepatobiliary Cancer, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer) for providing human tissue/blood samples and expert guidance to our study, and Dr. KailingTu (National Frontier Center of Disease Molecular Network, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University) for valuable discussions.

Note

Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Received October 23, 2023; revised January 17, 2024; accepted March 8, 2024; published first March 12, 2024.

References

1. Robertson KD. DNA methylation and human disease. Nat Rev Genet 2005;6:597–610.
2. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 1999;99:247–57.
3. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nat Rev Genet 2002;3:415–28.
4. Jammula S, Katz-Summercorn AC, Li X, Linossi C, Smyth E, Killcoyne S, et al. Identification of subtypes of Barrett's esophagus and esophageal adenocarcinoma based on DNA methylation profiles and integration of transcriptome and genome data. Gastroenterology 2020;158:1682–97.
5. Hetzel S, Mattei AL, Kretzmer H, Qu C, Chen X, Fan Y, et al. Acute lymphoblastic leukemia displays a distinct highly methylated genome. Nat Cancer 2022;3:768–82.
6. Klughammer J, Kiesel B, Roetzer T, Fortelny N, Nemc A, Nenning KH, et al. The DNA methylation landscape of glioblastoma disease progression shows extensive heterogeneity in time and space. Nat Med 2018;24:1611–24.
7. Zhao SG, Chen WS, Li H, Foye A, Zhang M, Sjöström M, et al. The DNA methylation landscape of advanced prostate cancer. Nat Genet 2020;52:778–89.
8. Choudhury A, Magill ST, Eaton CD, Prager BC, Chen WC, Cady MA, et al. Meningioma DNA methylation groups identify biological drivers and therapeutic vulnerabilities. Nat Genet 2022;54:649–59.
9. Sirica AE, Gores GJ, Groopman JD, Selaru FM, Strazzabosco M, Wei Wang X, et al. Intrahepatic cholangiocarcinoma: continuing challenges and translational advances. Hepatology 2019;69:1803–15.
10. Lamarca A, Barriuso J, McNamara MG, Valle JW. Molecular targeted therapies: ready for "prime time" in biliary tract cancer. J Hepatol 2020;73:170–85.
11. Goeppert B, Toth R, Singer S, Albrecht T, Lipka DB, Lutsik P, et al. Integrative analysis defines distinct prognostic subgroups of intrahepatic cholangiocarcinoma. Hepatology 2019;69:2091–106.
12. Jusakul A, Cutcutache I, Yong CH, Lim JQ, Huang MN, Padmanabhan N, et al. Whole-genome and epigenomic landscapes of etiologically distinct subtypes of cholangiocarcinoma. Cancer Discov 2017;7:1116–35.
13. Cleary JM, Raghavan S, Wu Q, Li YY, Spurr LF, Gupta HV, et al. FGFR2 extracellular domain in-frame deletions are therapeutically targetable genomic alterations that function as oncogenic drivers in cholangiocarcinoma. Cancer Discov 2021;11:2488–505.
14. Farshidfar F, Zheng S, Gingras MC, Newton Y, Shih J, Robertson AG, et al. Integrative genomic analysis of cholangiocarcinoma identifies distinct IDH-mutant molecular profiles. Cell Rep 2017;19:2878–80.
15. Dong L, Lu D, Chen R, Lin Y, Zhu H, Zhang Z, et al. Proteogenomic characterization identifies clinically relevant subgroups of intrahepatic cholangiocarcinoma. Cancer Cell 2022;40:70–87.
16. Galli GG, Carrara M, Yuan WC, Valdes-Quezada C, Gurung B, Pepe-Mooney B, et al. YAP drives growth by controlling transcriptional pause release from dynamic enhancers. Mol Cell 2015;60:328–37.
17. Kitchen P, Lee KY, Clark D, Lau N, Lertsuwan J, Sawasdichai A, et al. A runaway PRH/HHEX-Notch3-positive feedback loop drives cholangiocarcinoma and determines response to CDK4/6 inhibition. Cancer Res 2020;80:757–70.
18. Loyfer N, Magenheim J, Peretz A, Cann G, Bredno J, Klochendler A, et al. A DNA methylation atlas of normal human cell types. Nature 2023;613:355–64.
19. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26:841–2.
20. Ishibashi H, Suzuki T, Suzuki S, Moriya T, Kaneko C, Takizawa T, et al. Sex steroid hormone receptors in human thymoma. J Clin Endocrinol Metab 2003;88:2309–17.

21. Zhou W, Dinh HQ, Ramjan Z, Weisenberger DJ, Nicolet CM, Shen H, et al. DNA methylation loss in late-replicating domains is linked to mitotic cell division. Nat Genet 2018;50:591–602.
22. Feinberg AP, Koldobskiy MA, Göndör A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. Nat Rev Genet 2016;17:284–99.
23. Flavahan WA, Gaskell E, Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. Science 2017;357:eaal2380.
24. Baylin S, Bestor TH. Altered methylation patterns in cancer cell genomes: cause or consequence? Cancer Cell 2002;1:299–305.
25. Brennan K, Flanagan JM. Is there a link between genome-wide hypomethylation in blood and cancer risk? Cancer Prev Res (Phila) 2012;5:1345–57.
26. Nishiyama A, Nakanishi M. Navigating the DNA methylation landscape of cancer. Trends Genet 2021;37:1012–27.
27. Sproul D, Kitchen RR, Nestor CE, Dixon JM, Sims AH, Harrison DJ, et al. Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns. Genome Biol 2012;13:R84.
28. Brinkman AB, Nik-Zainal S, Simmer F, Rodríguez-González FG, Smid M, Alexandrov LB, et al. Partially methylated domains are hypervariable in breast cancer and fuel widespread CpG island hypermethylation. Nat Commun 2019; 10:1749.
29. Burger L, Gaidatzis D, Schübeler D, Stadler MB. Identification of active regulatory regions from DNA methylation data. Nucleic Acids Res 2013;41: e155.
30. Bock C, Beerman I, Lien WH, Smith ZD, Gu H, Boyle P, et al. DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. Mol Cell 2012;47:633–47.
31. Hon GC, Rajagopal N, Shen Y, McCleary DF, Yue F, Dang MD, et al. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. Nat Genet 2013;45:1198–206.
32. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature 2011; 480:490–5.
33. Sheffield NC, Pierron G, Klughammer J, Datlinger P, Schönegger A, Schuster M, et al. DNA methylation heterogeneity defines a disease spectrum in Ewing sarcoma. Nat Med 2017;23:386–95.
34. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 2010;38:576–89.
35. Bhutia YD, Babu E, Ramachandran S, Yang S, Thangaraju M, Ganapathy V. SLC transporters as a novel class of tumour suppressors: identity, function and molecular mechanisms. Biochem J 2016;473:1113–24.
36. Zhang X, Ma W, Cui J, Yao H, Zhou H, Ge Y, et al. Regulation of p21 by TWIST2 contributes to its tumor-suppressor function in human acute myeloid leukemia. Oncogene 2015;34:3000–10.
37. Ayrault O, Zhao H, Zindy F, Qu C, Sherr CJ, Roussel MF. Atoh1 inhibits neuronal differentiation and collaborates with Gli1 to generate medulloblastoma-initiating cells. Cancer Res 2010;70:5618–27.
38. Segueni J, Noordermeer D. CTCF: a misguided jack-of-all-trades in cancer cells. Comput Struct Biotechnol J 2022;20:2685–98.
39. Gurzov EN, Bakiri L, Alfaro JM, Wagner EF, Izquierdo M. Targeting c-Jun and JunB proteins as potential anticancer cell therapy. Oncogene 2008;27:641–52.
40. Nakayama T, Hieshima K, Arao T, Jin Z, Nagakubo D, Shirakawa AK, et al. Aberrant expression of Fra-2 promotes CCR4 expression and cell proliferation in adult T-cell leukemia. Oncogene 2008;27:3221–32.
41. Wan X, Guan S, Hou Y, Qin Y, Zeng H, Yang L, et al. FOSL2 promotes VEGF-independent angiogenesis by transcriptionnally activating Wnt5a in breast cancer-associated fibroblasts. Theranostics 2021;11:4975–91.
42. Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol 2009;27:361–8.
43. Aran D, Toperoff G, Rosenberg M, Hellman A. Replication timing-related and gene body-specific methylation of active human genes. Hum Mol Genet 2011;20: 670–80.
44. Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G. Gene body methylation can alter gene expression and is a therapeutic target in cancer. Cancer Cell 2014;26:577–90.
45. Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 2010;466:253–7.
46. Nakamura H, Arai Y, Totoki Y, Shirota T, Elzawahry A, Kato M, et al. Genomic spectra of biliary tract cancer. Nat Genet 2015;47:1003–10.
47. Lowery MA, Ptashkin R, Jordan E, Berger MF, Zehir A, Capanu M, et al. Comprehensive molecular profiling of intrahepatic and extrahepatic cholangiocarcinomas: potential targets for intervention. Clin Cancer Res 2018; 24:4154–61.
48. Brindley PJ, Bachini M, Ilyas SI, Khan SA, Loukas A, Sirica AE, et al. Cholangiocarcinoma. Nat Rev Dis Primers 2021;7:65.
49. Gurjar A, Colman H, Holmen S. EXTH-69. astrocytoma cells harboring mutant IDH1 are uniquely sensitive to inhibition of oxidative phosphorylation. Neuro-oncol 2022;24:vii225–vii.
50. Ni Y, Shen P, Wang X, Liu H, Luo H, Han X. The roles of IDH1 in tumor metabolism and immunity. Future Oncol 2022;18:3941–53.
51. Izquierdo-Garcia JL, Viswanath P, Eriksson P, Cai L, Radoul M, Chaumeil MM, et al. IDH1 mutation induces reprogramming of pyruvate metabolism. Cancer Res 2015;75:2999–3009.
52. Affar EB, Carbone M. BAP1 regulates different mechanisms of cell death. Cell Death Dis 2018;9:1151.
53. Han A, Chua V, Baqai U, Purwin TJ, Bechtel N, Hunter E, et al. Pyruvate dehydrogenase inactivation causes glycolytic phenotype in BAP1 mutant uveal melanoma. Oncogene 2022;41:1129–39.
54. Lin J, Dai Y, Sang C, Song G, Xiang B, Zhang M, et al. Multimodule characterization of immune subgroups in intrahepatic cholangiocarcinoma reveals distinct therapeutic vulnerabilities. J Immunother Cancer 2022;10: e004892.
55. Andersen JB, Spee B, Blechacz BR, Avital I, Komuta M, Barbour A, et al. Genomic and genetic characterization of cholangiocarcinoma identifies therapeutic targets for tyrosine kinase inhibitors. Gastroenterology 2012; 142:1021–31.
56. Job S, Rapoud D, Dos Santos A, Gonzalez P, Desterke C, Pascal G, et al. Identification of four immune subtypes characterized by distinct composition and functions of tumor microenvironment in intrahepatic cholangiocarcinoma. Hepatology 2020;72:965–81.
57. Oishi N, Kumar MR, Roessler S, Ji J, Forgues M, Budhu A, et al. Transcriptomic profiling reveals hepatic stem-like gene signatures and interplay of miR-200c and epithelial-mesenchymal transition in intrahepatic cholangiocarcinoma. Hepatology 2012;56:1792–803.
58. Guardia T, Zhang Y, Thompson KN, Lee SJ, Martin SS, Konstantopoulos K, et al. OBSCN restoration via OBSCN-AS1 long-noncoding RNA CRISPR-targeting suppresses metastasis in triple-negative breast cancer. Proc Natl Acad Sci USA 2023;120:e2215553120.
59. Han A, Purwin TJ, Aplin AE. Roles of the BAP1 tumor suppressor in cell metabolism. Cancer Res 2021;81:2807–14.
60. Field MG, Kuznetsov JN, Bussies PL, Cai LZ, Alawa KA, Decatur CL, et al. BAP1 loss is associated with DNA methylomic repatterning in highly aggressive class 2 uveal melanomas. Clin Cancer Res 2019;25:5663–73.
61. Tretina K, Park ES, Maminska A, MacMicking JD. Interferon-induced guanylate-binding proteins: Guardians of host defense in health and disease. J Exp Med 2019;216:482–500.
62. Zhu AX, Macarulla T, Javle MM, Kelley RK, Lubner SJ, Adeva J, et al. Final overall survival efficacy results of ivosidenib for patients with advanced cholangiocarcinoma with IDH1 mutation: the phase 3 randomized clinical ClarIDHy trial. JAMA Oncol 2021;7:1669–77.
63. Zhang M, Huang Y, Pan J, Sang C, Lin Y, Dong L, et al. An inflammatory checkpoint generated by IL1RN splicing offers therapeutic opportunity for KRAS mutant intrahepatic cholangiocarcinoma. Cancer Discov 2023;13: 2248–69.
64. Abou-Alfa GK, Sahai V, Hollebecque A, Vaccaro G, Melisi D, Al-Rajabi R, et al. Pemigatinib for previously treated, locally advanced or metastatic cholangiocarcinoma: a multicentre, open-label, phase 2 study. Lancet Oncol 2020;21:671–84.
65. Ogino S, Cantor M, Kawasaki T, Brahmandam M, Kirkner GJ, Weisenberger DJ, et al. CpG island methylator phenotype (CIMP) of colorectal cancer is best characterised by quantitative DNA methylation analysis and prospective cohort studies. Gut 2006;55:1000–6.
66. Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, Berman BP, et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell 2010;17:510–22.
67. Li J, Xu C, Lee HJ, Ren S, Zi X, Zhang Z, et al. A genomic and epigenomic atlas of prostate cancer in Asian populations. Nature 2020;580:93–9.

68. Wadi S, Tipton AR, Trendel JA, Khuder SA, Vestal DJ. hGBP-1 expression predicts shorter progression-free survival in ovarian cancers, while contributing to paclitaxel resistance. J Cancer Ther 2016;7:994–1007.
69. Andreoli M, Persico M, Kumar A, Orteca N, Kumar V, Pepe A, et al. Identification of the first inhibitor of the GBP1:PIM1 interaction. Implications for the development of a new class of anticancer agents against paclitaxel resistant cancer cells. J Med Chem 2014;57:7916–32.
70. Li M, Mukasa A, Inda MM, Zhang J, Chin L, Cavenee W, et al. Guanylate binding protein 1 is a novel effector of EGFR-driven invasion in glioblastoma. J Exp Med 2011;208:2657–73.
71. Faubert B, Solmonson A, DeBerardinis RJ. Metabolic reprogramming and cancer progression. Science 2020;368:eaaw5473.
72. Friedman K, Brodsky AS, Lu S, Wood S, Gill AJ, Lombardo K, et al. Medullary carcinoma of the colon: a distinct morphology reveals a distinctive immunoregulatory microenvironment. Mod Pathol 2016;29:528–41.
73. Elsayed I, Elsayed N, Feng Q, Sheahan K, Moran B, Wang X. Multi-OMICs data analysis identifies molecular features correlating with tumor immunity in colon cancer. Cancer Biomark 2022;33:261–71.
74. Ottenhoff TH, Dass RH, Yang N, Zhang MM, Wong HE, Sahiratmadja E, et al. Genome-wide expression profiling identifies type 1 interferon response pathways in active tuberculosis. PLoS One 2012;7:e45839.
75. Liu Z, Sun J, Gong T, Tang H, Shen Y, Liu C. The prognostic and immunological value of guanylate-binding proteins in lower-grade glioma: potential markers or not? Front Genet 2021;12:651348.
76. Gou Q, Zhang CZ, Sun ZH, Wu LG, Chen Y, Mo ZQ, et al. Cell-free DNA from bile outperformed plasma as a potential alternative to tissue biopsy in biliary tract cancer. ESMO open 2021;6:100275.
77. Lee MS, Kaseb AO, Pant S. The emerging role of circulating tumor DNA in non-colorectal gastrointestinal cancers. Clin Cancer Res 2023;29:3267–74.