JGH Open Open Access An open access journal of gastroenterology and hepatology

JGHF Journal of Gastroenterology and Hepatology Foundation

doi:10.1002/jgh3.12833

ORIGINAL ARTICLE

Persistent epigenetic alterations in transcription factors after a sustained virological response in hepatocellular carcinoma

Keishi Sugimachi,* Hiromitsu Araki,†,‡ Hideyuki Saito,§ Takaaki Masuda,§ Fumihito Miura,† Kentaro Inoue,†,¶ Tomonari Shimagaki,* Yohei Mano,* Tomohiro Iguchi,* Masaru Morita,‖ Yasushi Toh,‖ Tomoharu Yoshizumi,¶ Takashi Ito† and Koshi Mimori§

Departments of *Hepatobiliary-Pancreatic Surgery, ‖Gastroenterological Surgery, National Hospital Organization Kyushu Cancer Center, Departments of †Biochemistry, Graduate School of Medical Sciences, ‡Business and Technology Management, Faculty of Economics, ¶Surgery and Science, Graduate School of Medical Sciences, Kyushu University, Fukuoka and §Department of Surgery, Kyushu University Beppu Hospital, Beppu, Japan

Key words
hepatocellular carcinoma, methylation, sustained virologic response, transcription factor.

Accepted for publication 8 October 2022.

Correspondence
Keishi Sugimachi, Department of Hepatobiliary-Pancreatic Surgery, National Hospital Organization Kyushu Cancer Center, 3-1-1 Notame, Minami-ku, Fukuoka 811-1395, Japan.
Email: ksugimachi@gmail.com

Declaration of conflict of interest: None.

Funding support: Japan Society for the Promotion of Science 18K08722; 21K08745

Funding support: Platform Project for Supporting Drug Discovery and Life Science Research JP21am0101103

Funding support: Taiju Life Social Welfare Foundation

Abstract

Background and Aim: The risk of hepatocellular carcinoma (HCC) persists in a condition of sustained virologic response (SVR) after hepatitis C virus (HCV) eradication. Comprehensive molecular analyses were performed to test the hypothesis that epigenetic abnormalities present after an SVR play a role in hepatocarcinogenesis.
Methods: Whole-genome methylome and RNA sequencing were performed on HCV, SVR, and healthy liver tissue. Integrated analysis of the sequencing data focused on expression changes in transcription factors and their target genes, commonly found in HCV and SVR. Identified expression changes were validated in demethylated cultured HCC cell lines and an independent validation cohort.
Results: The coincidence rates of the differentially methylated regions between the HCV and SVR groups were 91% in the hypomethylated and 71% in the hypermethylated regions in tumorous tissues, and 37% in the hypomethylated and 36% in the hypermethylated regions in non-tumorous tissues. These results indicate that many epigenomic abnormalities persist even after an SVR was achieved. Integrated analysis identified 61 transcription factors and 379 other genes that had methylation abnormalities and gene expression changes in both groups. Validation cohort specified gene expression changes for 14 genes, and gene ontology pathway analysis revealed apoptotic signaling and inflammatory response were associated with these genes.
Conclusion: This study demonstrates that DNA methylation abnormalities, retained after HCV eradication, affect the expression of transcription factors and their target genes. These findings suggest that DNA methylation in SVR patients may be functionally important in carcinogenesis, and could serve as biomarkers to predict HCC occurrence.

Introduction

Hepatocellular carcinoma (HCC) is the fifth most common cancer worldwide, and results in death of more than 30 000 people annually in Japan. The 5-year survival rate of HCC is 30%, despite multiple treatment strategies, including hepatectomy.[1] Thus, cancer prevention and early detection are of paramount importance in controlling HCC. Chronic hepatitis and liver cirrhosis, associated with infection from the hepatitis C virus (HCV), are major causes of hepatocarcinogenesis worldwide and in Japan; however, several breakthrough treatments for HCV infection have emerged over the last few decades.[2] In 2009, an interferon (IFN)-based treatment regimen produced a sustained viral response (SVR) and eliminated the HCV from the body in nearly 70% of patients. In addition, a direct-acting antiviral agent (DAA) recently enabled an SVR in almost 100% of patients. Multiple studies have proven that achieving an SVR significantly reduced (more than 70%) the risk of HCC in patients who had a history of HCV infection.[3]

Despite the progress of anti-HCV therapies, a number of primary or recurrent HCCs occur after an HCV infection is eliminated.[4,5] HCCs that occur in patients who have achieved an SVR (SVR-HCC) present new issues and clinical problems to be addressed. Since there are currently no biomarkers that define patients with a high risk for developing SVR-HCC, all patients must undergo surveillance for life after achieving an SVR, despite the relatively low carcinogenic rate. Identification of patients at a high risk for developing HCC would improve prognosis and reduce unnecessary medical costs. It remains unclear whether there is a difference in risk for hepatocarcinogenesis

based on different treatment regimens (IFN or DAA), considering the potential effect of DAAs on angiogenesis and cellular proliferation.[6,7]

Challenging clinical questions associated with HCC patients who have achieved an SVR include, why HCC develops and what differences exist between patients who develop HCC after an SVR and those who do not. HCV-induced epigenetic changes associated with hepatocarcinogenesis were found to persist after an SVR in clinical samples[8] and after virus eradication in cell lines.[9] These observations have generated a focus on epigenetic changes, which may elucidate a mechanism for carcinogenesis in HCC after SVR.

In this study, we sought to identify the molecular mechanisms behind hepatocarcinogenesis after SVR, utilizing integrated epigenomic and transcriptomic analyses. The methylation abnormalities identified herein may help to elucidate the molecular mechanisms underlying HCC occurrence, after an SVR has been achieved, and significantly contribute to the diagnosis and treatment of HCC.

Methods

Clinical samples. This study was approved by the institutional review board (the Ethics and Indications Committee of the Kyushu Cancer Center, 24-160). Written informed consent to participate in the study was obtained from each patient. For the methylation analysis, tumors and the corresponding adjacent non-tumoral liver samples were collected from 14 patients who underwent hepatectomy for primary HCC at the Kyushu Cancer Center and Kyushu University. The clinical samples collected included eight patients with an HCV infection and six with an SVR. Additionally, normal (non-tumorous) liver tissue samples were obtained from 10 patients who underwent hepatectomy for metastatic liver cancer or intrahepatic cholangiocarcinoma. These samples were selected from patients who did not have hepatitis, fatty liver disease, or any other chronic liver disease. For the validation analysis, samples were obtained from 13 patients with HCC who underwent hepatectomy at the Kyushu Cancer Center. Disease etiology for the validation cohort included five patients with HCV and eight patients with an SVR. In this study, only fresh frozen clinical samples were used in the assays described below. As a quality control, the resected specimens were confirmed by histology and the part with less necrosis and fibrosis and more viable tumor cells was selected for nucleic acid extraction refereeing previous study.[9]

Target methylome sequencing. Library preparation for targeted methylome sequencing was performed based on the post-bisulfite adaptor tagging (PBAT)[10] adapted for Sure Select-based target enrichment as described previously.[11] All libraries for the shotgun bisulfite sequencing were prepared from 330 ng of purified genomic DNA. After amplification with minimum cycles of polymerase chain reaction (PCR) (4 cycles), these libraries were served for sequencing on HiSeq 2500 with 101 cycles of single-end sequencing in rapid run mode. Two libraries were mixed and assigned for a lane.

RNA sequencing. The library preparation for RNA sequencing (RNA-seq) analysis of HCC patients' tissues was performed using NEBNext Ultra Directional RNA Library Prep Kit for Illumina combined with NEBNext rRNA Depletion Kit (human/mouse/rat). One hundred nanograms of total RNA was served for the library preparation, and 12 cycles of PCR amplification were performed. Sequencing was performed with Illumina HiSeq 2500 with 101 cycles of single-end sequencing in high output mode. We obtained RNA-seq data on 14 normal liver tissue samples from GSE126848.[12]

Analysis of differentially methylated regions. Targeted bisulfite sequencing (TGBS) reads were mapped to the Human genome build 19 (hg19) reference genome using Bmap (http://itolab.med.kyushu-u.ac.jp/BMap/index.html). Methylation levels were calculated for individual CpG sites. The methylation levels of probe regions were determined by averaging the methylation levels of CG sites in individual features. Hierarchical clustering of the mean methylation levels of probe regions was performed by calculating Euclidian distance matrices and using Ward's linkage method. Differentially methylated regions (DMRs) were identified using metilene, with 0.2–6 for the default parameters.[13] DMRs were filtered to retain those containing a minimum of 10 CpG sites with a *P*-value <0.01 and a methylation difference larger than 20%. Genes located within 5 kbs of DMRs were defined as being associated with the corresponding DMR.

RNA-seq in validation cohort. Total RNA was isolated using the TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) and purified with the SV Total RNA Isolation System (Promega, Madison, WI, USA) according to the manufacturer's instructions. RNA samples were quantified using an ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and the quality was confirmed on a TapeStation (Agilent, Santa Clara, CA, USA). The sequencing libraries were prepared from 200 ng of total RNA using the MGIEasy rRNA Depletion Kit and the MGIEasy RNA Directional Library Prep Set (MGI, Shenzhen, China) following the manufacturer's instructions. The libraries were sequenced on the DNBSEQ-G400 FAST Sequencer (MGI) using a paired-end 150nt strategy. Quality trimming and adapter clipping of the read data were performed using Trimmomatic version 0.38.[14] Trimmed reads were mapped to the transcript using the hg38 and hg19 reference genomes with the Bowtie2 aligner software within RSEM.[15] Abundance estimation of the genes and isoforms from RSEM produced basic (expected) data counts.

RNA-seq analysis of gene expression. Raw RNA-seq fastq files were preprocessed using fastp 0.20.0 to filter out low-quality reads and cut adapter sequences.[16] Preprocessed reads were mapped to the hg19 reference genome using STAR 2.6.0c,[17] then the mapped reads were assigned to exons using featureCounts v1.4.6.[18] Genes with <10 reads were discarded. Differential analysis was conducted using the Bioconductor package edgeR 3.26.8.[19] A gene was defined as differentially expressed if the false discovery rate (FDR), corrected with the Benjamini–Hochberg method, was less than 0.01 and if the log2 fold change (log2FC) was greater than 1 (up-regulated) or less than −1 (down-regulated).

Cell culture and demethylation treatment. The human HCC cell lines Huh7 and PLC/PRF/5 were obtained from the American Type Culture Collection and the Japanese Collection of Research Bioresources. Cells were grown in DMEM media supplemented with 100 g/L of fetal bovine serum and incubated at 37 °C in 50 mL/L CO_2. Prior to demethylation,

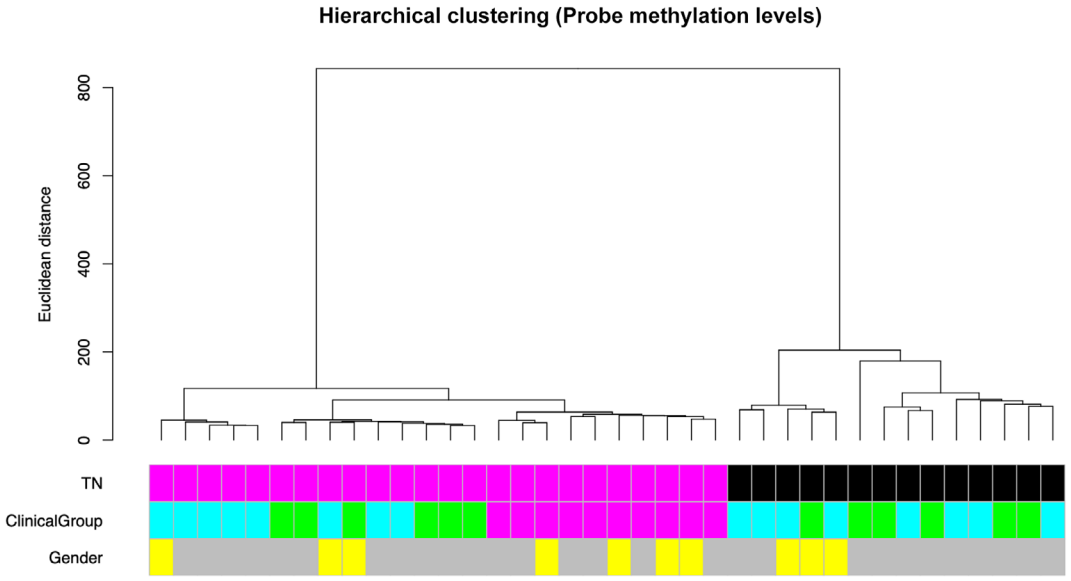

Figure 1 Hierarchical clustering of whole-genome methylation status in hepatocellular carcinoma (HCC). All samples were clustered based on their whole-genome methylation status, as defined by a comprehensive genome-wide analysis using highly sensitive methylome sequencing through post-bisulfite adaptor tagging. Upper row (TN): (■), T—tumorous tissue; (■), N—non-tumorous tissue. Middle row (clinical group): (■), HCV—samples from patients with chronic hepatitis C; (■), SVR—samples from patients with a sustained viral response; (■), normal—normal liver tissue. Lower row (gender): (■), F—female; (■), M, male.

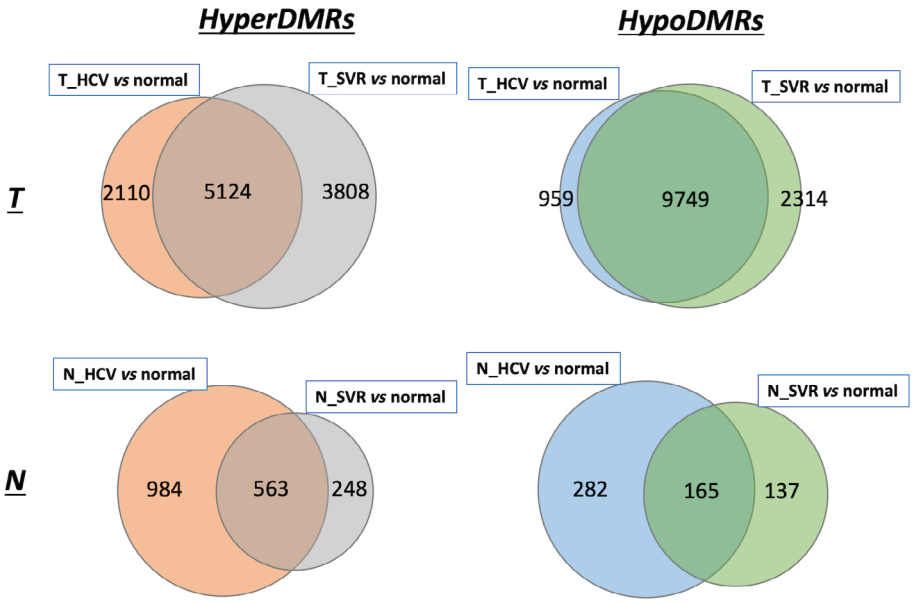

Figure 2 Venn diagram comparing the methylation status of hepatitis C virus (HCV) and sustained virologic response (SVR) cases. The number of differentially methylated regions (DMRs) identified in tissue from HCV and SVR cases are shown in the Venn diagram. The upper panels represent the comparison for tumorous tissues (T), and the lower panels show the comparison for non-tumor tissue (N). The panels on the left depict the number of hypermethylated DMRs for each comparison, and the panels on the right show the number of hypomethylated DMRs.

the cells were split to a concentration of 5×10^5 cells per 75-cm^2 culture bottle and incubated overnight in the growth media. To treat the cells the media was replaced with the growth media supplemented with the demethylating agent, 5-aza-2′-deoxycytidine (5-Aza) (Sigma, St Louis, MO, USA), at a final concentration of 10 μM for 6 days. The media was changed on day four. Cells cultured with vehicle alone served as a negative control for the 5-Aza treatment. After the culturing period, the cells were harvested for total RNA extraction. Three independent replicate assays were performed.

Gene ontology pathway analysis. Pathway analysis was performed using the statistical software R version 4.1.2 and the R package clusterProfiler version 4.2.1.[20] Gene ontology (GO) analysis was performed using genes with variable expression in HCV and SVR groups to analyze pathways of significant enrichment. Of the GO terms, only those related to cancer-related pathways were selected and shown in the figure.

Results

Whole-genome methylation status of HCV-associated HCC. A flowchart of the methods and the analyses performed for this study can be found in Figure S1, Supporting information. To investigate whether chronic HCV infection induces persistent epigenetic alterations even after viral eradication, a comprehensive genome-wide analysis using highly sensitive methylome sequencing through PBAT was performed. Significant clustering of whole-genome methylation status, among tumorous and non-tumorous tissue obtained from HCC

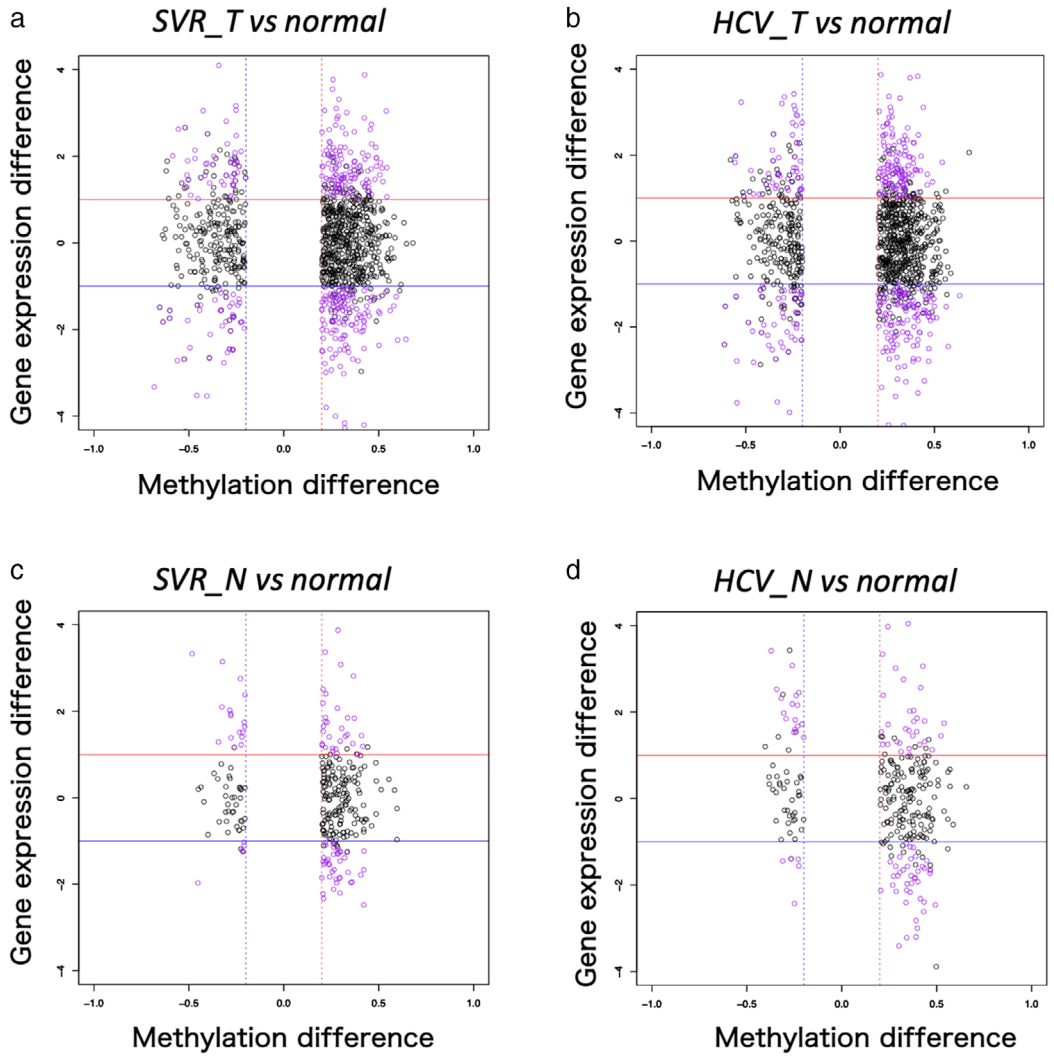

Figure 3 Plot depicting the differentially methylated genes and differentially expressed genes identified in hepatocellular carcinoma. Gene expression differences (vertical axis) and methylation level differences (horizontal axis) were plotted for each gene. Only genes with a difference in methylation level of ±20% were plotted. (a) Comparison of tumorous tissue from sustained virologic response (SVR) cases to normal liver tissue. (b) Comparison of tumorous tissue from hepatitis C virus (HCV) cases to normal liver tissue. (c) Comparison of non-tumorous tissue from SVR cases to normal liver tissue. (d) Comparison of non-tumorous tissue from HCV cases to normal liver tissue.

Table 1 Transcriptional factors and target genes identified in the comprehensive analysis of clinical samples and cell line demethylation experiments

TF	TF_DMR_T	TF_DMR_N	TF_DEG_T	TF_DEG_N	Target genes	Regulation	Target genes_DEG_T	Target genes_DEG_N	PLF-PRF5	Huh7	Consistent with TF function
RXRA	Hyper	Hyper	Down	Down	HSD17B2	Activation, unknown	Down	n.s.	n.s.	Up	Yes
RELA	Hyper		Down		CXCL12	Activation	Down	Down	Down	n.s.	Yes
RELA	Hyper		Down		TF	Activation, unknown	Down	Down	Up	n.s.	Yes
RELA	Hyper		Down		TNFRSF10B	Activation, unknown	Up	Up	Down	n.s.	Yes
RELA	Hyper		Down		TRIB3	Repression	Up	Up	Down	Down	Yes
KLF4	Hyper		Up		CDH5	Activation	Down	n.s.	n.s.	Up	No
KLF4	Hyper		Up		IFITM3	Repression	Down	Down	n.s.	Up	Yes
KLF4	Hyper		Up		SOD1	Repression	Down	Down	n.s.	Up	Yes
KLF4	Hyper		Up		ATF3	Activation	Up	Up	Down	Down	Yes
RUNX1	Hyper		Up		KLF4	Unknown	Up	Up	Down	Down	Yes
RORA		Hyper		Up	TUBA1B	Unknown	n.s.	Down	Up	Up	Yes

DEG, differentially expressed gene; DMR, differentially methylated region; N, non-tumorous tissue; n.s., not significant; T, tumorous tissue; TF, transcription factor.

patients, as well as a clustering of normal liver tissues, was observed (Fig. 1). However, methylation status did not cluster based on clinical groups (HCV or SVR) in either the tumorous or non-tumorous liver tissues (Fig. 1). Next, DMRs were identified in HCV, SVR, and normal liver tissues. When comparing the DMRs in tumorous tissues compared with non-tumorous tissues, 6572 hypermethylated and 15 148 hypomethylated tumor/non-tumor DMRs were identified. Examination of the DMRs in tumorous HCV and SVR tissues compared with normal liver tissues found that a total of 5124 out of 7234 hypermethylated tumor/normal DMRs (70.8%), and 9749 of 10 708 hypomethylated tumor/normal DMRs (91%), were shared between the HCV and SVR cases, indicating that the methylation profiles of the tumorous tissue are similar for both HCV and SVR (Fig. 2).

In order to analyze the epigenetic changes in HCC carcinogenesis, the methylome of the non-tumorous tissues was examined. When comparing the DMRs identified in the HCV and SVR tissues to normal liver tissue without hepatitis, a total of 1547 hypermethylated non-tumor/normal DMRs were found in HCV non-cancerous tissue compared with normal liver tissues, while only 811 hypermethylated non-tumor/normal DMRs were found in SVR non-cancerous tissue compared with normal liver tissues. Hypomethylated regions followed a similar trend, with 447 non-tumor/normal DMRs identified in the HCV non-cancerous tissue and 302 non-tumor/normal DMRs found in SVR non-cancerous tissue compared with normal liver tissues. Additionally, a total of 563 (69.4%) of the 811 hypermethylated non-tumor/normal DMRs, and 165 (54.6%) of the 302 hypomethylated non-tumor/normal DMRs, in the non-tumorous tissue isolated from the SVR patients were also found in non-tumorous tissue from HCV patients (Fig. 2). These findings suggest that a significant number of the epigenomic changes caused by an HCV infection are retained after achieving an SVR, and that these changes do not revert back to a normal state. Thus, the retained epigenomic changes may be associated with hepatocarcinogenesis in SVR patients.

Methylation-dependent differentially expressed genes in noncancerous liver tissues. Genes located within 5 kb of a DMR were interpreted as related genes (Figure S2A), and an integrated epigenetic and transcriptomic analysis was performed to identify possible affected genes whose regulation was directly or indirectly impacted by the epigenetic alterations (Figure S2B). It is well established that DNA methylation in promoter regions involves in gene regulation, but DNA methylation in intra- and inter-genic regions including downstream of gene transcriptional termination sites also play important roles in its gene functions.[21] In this study, we defined DMR-associated genes refereeing to the study by Bhasin *et al.*[22] They related DMRs to genes within 2 kb of the DMRs, but we expanded the range to 5 kb to capture DMR-associated genes more widely. The DMRs and differentially expressed genes (DEGs) identified through methylome and RNA sequencing are depicted in graphs comparing the SVR and HCV samples to the results for normal liver tissue (Fig. 3). Genes that showed alterations in both methylation level (DMRs) and gene expression level (DEGs) were identified, and the genes present in both the HCV and SVR cases are reported in Table S1. A total of 212 hypermethylated and 88 hypomethylated DEGs were identified in the tumorous tissues, and 61 hypermethylated and 18 hypomethylated DEGs were identified in non-tumorous tissue (Table S1).

Identification of potential oncogenic target genes regulated by methylation status and transcription factors. Genes with alterations in methylation status and gene expression levels (DMR-DEGs) were then classified into groups, transcription factors (TFs), and not TFs, utilizing public databases,[23] since methylation is an important mechanism regulating TF expression and is known to be involved in both liver cancer and stem cell development.[24] Although DNA methylation is generally associated with transcriptional silencing, the effects of methylation on binding affinity for most TFs are still unknown.

Previous studies reported that the TF binding affinity of individual TFs can either be increased or decreased on methylation.[25] As this study explores global methylation status, it contains both activation and repression of TF functions on target genes.We identified 40 hypermethylated and 6 hypomethylated TFs in tumorous tissues and identified 11 hypermethylated and 4 hypomethylated TFs in non-tumorous tissues (Table S1). The representative methylation status of 2 TFs (*KLF4* and *RXRA*) is shown in Figure S3. Methylation level was significantly higher in tumorous SVR and HCV tissues compared with normal liver tissue. To address the potential functional significance of alterations to these specific TFs, 661 downstream target genes were identified utilizing the TRRUST database.[26] Through a comprehensive analysis of their methylation and gene expression status, these target genes were narrowed down to 111 genes, whose gene expression was either significantly increased or decreased, corresponding to their specific TF's expression. Thus, these genes were identified as differentially expressed target genes (DETGs) (Table S2). There are several target genes of each TF, and the epigenetic transcriptional regulation includes both activation and repression defined by the TRRUST database. These functions can dynamically change in tissues and cells, and the expression of each target gene can be affected directly and indirectly by various signaling pathways.To confirm that the DETGs were functionally affected by methylation, cultured HCC cell lines were used. A global demethylation assay was performed in two HCC cell lines, Huh7 and PLF-PRF5. Complete DNA demethylation was achieved by treating the HCC cells with 5-aza-2′-deoxycytidine, as described above, and gene expression changes were determined by RNA-seq analysis. We selected genes whose expression changes in cell lines were consistent with those identified in clinical samples with hypomethylated DMRs since the cell lines were globally demethylated. Conversely, the DETGs that showed inverse expression were identified through hypermethylated DMRs in clinical samples. Significant changes were

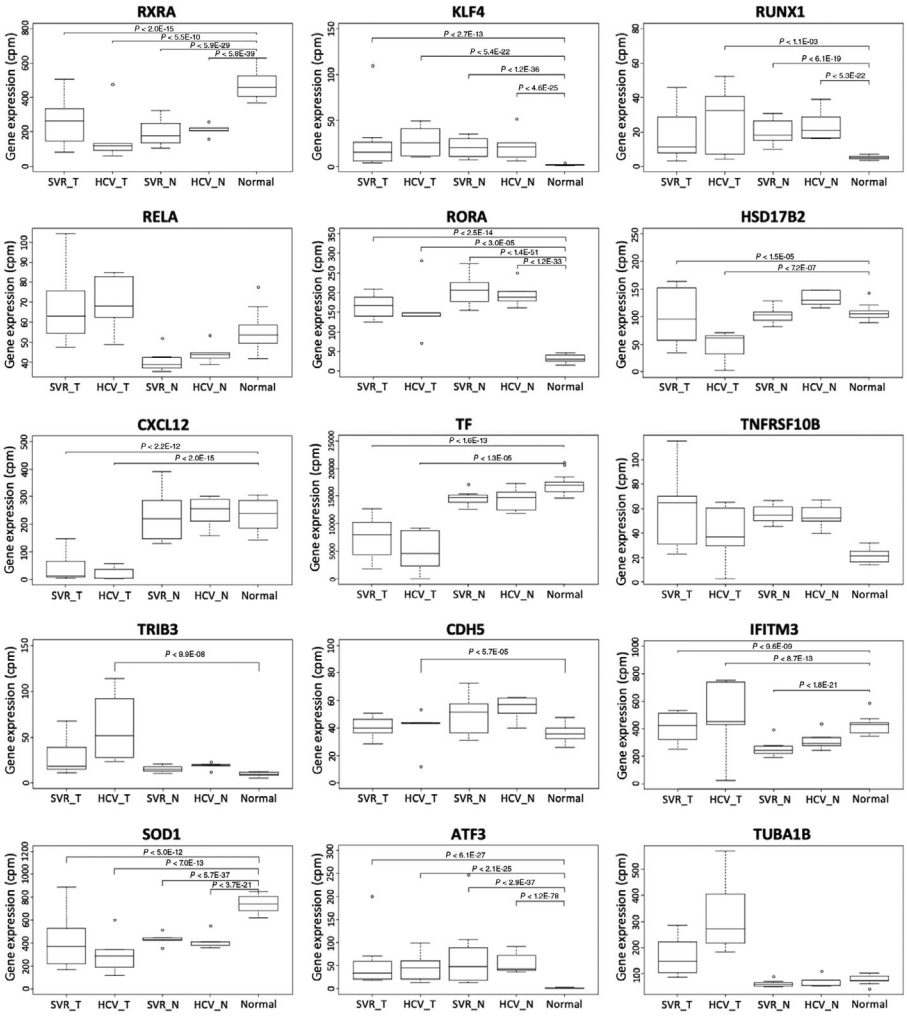

Figure 4 Gene expression levels in the validation cohort. Gene expression levels for the 5 transcription factors (TFs) and 11 target genes (*KLF4* is included as both a TF and target gene) were identified in the study. Whisker plots depict the expression level (cpm, counts per million) for each gene on the vertical axis, and the horizontal axis shows the clinical group and tissue type: sustained virologic response (SVR), hepatitis C viral infection (HCV), tumorous (T), and non-tumorous (N). Differences in the adjusted *P*-value (adjp) are shown as calculated using the Bioconductor edgeR 3.26.8 software package.

Table 2 Comparison of gene expression levels between the first systematic analysis and the second validation cohort

Gene name	Function	Derivation_T	Derivation_N	Validation_T	Validation_N
RXRA	Transcription factor	Down	Down	Down	Down
KLF4	Transcription factor	Up		Up	Up
RUNX1	Transcription factor	Up		Up	Up
RORA	Transcription factor		Up	Up	Up
RELA	Transcription factor	Down		Up	
HSD17B2	Target gene	Down		Down	
CXCL12	Target gene	Down	Down	Down	
TF	Target gene	Down	Down	Down	
TNFRSF10B	Target gene	Up	Up	Up	Up
TRIB3	Target gene	Up	Up	Up	
CDH5	Target gene	Down		Down	
IFITM3	Target gene	Down	Down	Down	
SOD1	Target gene	Down	Down	Down	Down
ATF3	Target gene	Up	Up	Up	Up
TUBA1B	Target gene		Down	Up	

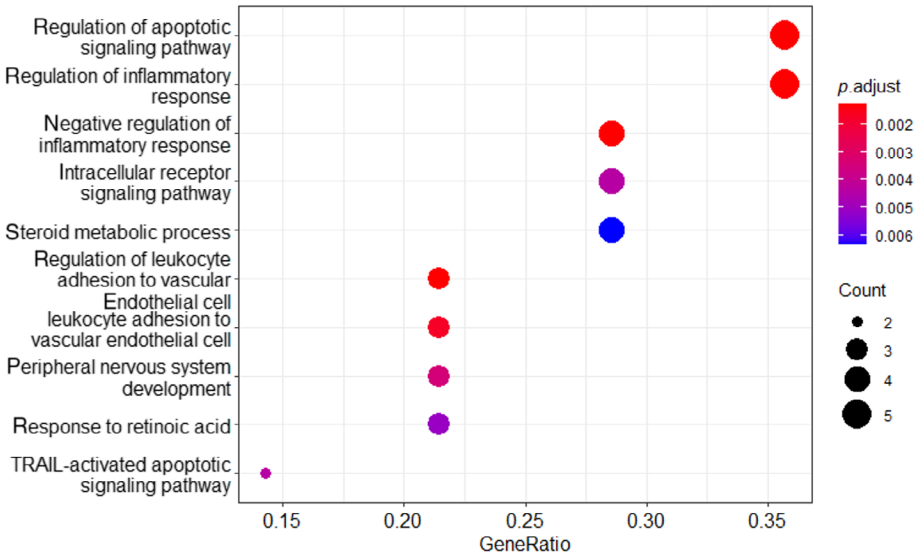

Figure 5 Gene ontology pathway analysis of 14 significant transcription factors and target genes. The gene ontology analysis shows a higher concentration of genes related to apoptotic signaling pathway and sis and inflammatory response. The vertical axis is the higher rank of the 2241 pathways, and the horizontal axis is the ratio of how many of the 14 genes belong to that pathway. The color of the circle shows the *P* values of the association of gene sets and pathways, and the size of the circle shows the number of genes that belong to the pathway.

identified in the expression of 11 genes in both the clinical samples and the cell lines (Table 1). Of these 11 genes, 10 showed alterations in target gene expression and along with their corresponding TFs, as identified with the TRRUST database[26] (Table 1). These data suggest a mechanism whereby gene function, as regulated by TFs, is altered through methylation changes after an HCV infection. Further supporting the theory that epigenetic modifications, which persist after HCV cure, could play a role in hepatocarcinogenesis.

Gene expression analysis of HCV and SVR-HCC in a validation cohort. Through a systematic analysis, several genes with significant gene expression changes have been identified (Table 1). The validity of these changes was subsequently confirmed by performing an RNA-seq analysis on an independent cohort containing five clinical samples from patients with HCV-HCC and eight patients with SVR-HCC. Gene expression levels for 5 TFs and 11 DETGs (including 1 TF; *KLF4*) identified in the first analysis were examined in the second validation cohort (Fig. 4 and Table 2). With the exception of the *RELA* transcription factor, gene expression changes of 4 TFs (*RXRA*, *KLF4*, *RUNX1*, and *RORA*) and 10 target genes (*HSD17B2*, *CXCL12*, *TF*, *TNFRSF10B*, *TRIB3*, *CDH5*, *IFITM3*, *SOD1*, *ATF3*, and *TUBA1B*) in both tumorous and non-tumorous tissue were confirmed in the validation cohort. Aberrant DNA methylation is caused by dysregulation of DNA methyltransferase (DNMT) and/or ten-eleven translocation (TET) pathways. We investigated the expression of DNMTs and TETs in HCV-HCC, SVR-HCC, and

normal liver tissue in the validation cohort. The expression levels of *DNMT1*, *DNMT3A*, *TET2*, and *TET3* in tumorous SVR and HCV tissues were significantly higher compared with normal liver tissues (Figure S4). Interestingly, the expression levels of non-tumorous SVR and HCV tissues were also higher compared with normal liver tissue (Figure S4). These results indicated that methylation/demethylation activities increased in HCCs and sustain after SVR.

GO analysis of SVR-related transcription factors and target genes. We examined how these 14 extracted genes increased the risk of HCC using the GO pathway sets. First, we listed 111 pathways containing two or more of the 14 genes of the 2241 molecular biology pathways using the cluster Profiler (Table S3). We found 10 pathways as statistically significantly enriched (Fig. 5). Notably, 2 GOs of 111 gene sets, regulation of apoptotic signaling pathway and regulation of inflammatory response, were found to include 5 of those 14 genes and showed highly significant associations ($P = 0.002$, Fig. 5). The present analysis indicated that apoptotic and inflammatory pathways were possibly associated with carcinogenesis of hepatocytes in SVR.

Discussion

For the first time, genome-wide methylation profiles were characterized and an integrated analysis of methylation and gene expression data was performed on HCC tissue samples from patients cured of HCV infection. In this study, methylation changes that persisted even after the HCV was eliminated were detected, and TFs and their corresponding downstream targets affected by aberrant methylation were identified. Thus, these genes may be associated with HCC carcinogenesis.

Previous studies have reported that DNA methylation occurs during both early and advanced HCC development.[27] Recently, epigenomic alterations associated with HCCs after eradication of HCV (SVR) were observed in clinical samples and cell lines.[8,28] The reason HCC occurs after resolution of an HCV infection is still unknown. The novelties of this study are that we investigated the comprehensive genome-wide methylation profile in clinical SVR-HCC samples and that we did in silico analysis focusing on TF and target genes. One reason this study focused on examining the epigenome was the similarity between the occurrence of HCC as a result of the HCV, and that the development of gastric cancer due to *Helicobacter pylori*. *H. pylori* is widely known to cause gastric cancer; however, after eradication of bacteria, the incidence of gastric cancer decreases, yet a certain number of gastric cancers still occur.[29]

Abnormal methylation plays an important role in carcinogenesis.[29] As was hypothesized, this study revealed that epigenetic changes, characterized by DNA methylation changes, persisted even after an SVR was achieved. These changes may play a role in carcinogenesis. Whether these epigenetic alterations are a direct effect of the HCV, or a secondary effect of fibrosis or chronic inflammation, must be determined in future studies.

In this study, the association between DMRs and DEGs was analyzed with a focus on TFs and their downstream target genes. It is known that global methylation is maintained throughout HCC progression; thus, methylated genes could be used for the early detection of, or as predictive biomarkers of, HCC.[30,31] This study focused on TFs for three reasons. First, methylation is associated with TF function. Second, since TFs are related to the maintenance of stemness, we hypothesized that stemness might also play a role in SVR carcinogenesis. Third, to date, no research has focused on the role of TFs in methylation changes in SVR-associated HCC.

This study identified four TFs (*RXRA*, *KLF4*, *RUNX1*, and *RORA*) and 10 target genes (*HSD17B2*, *CXCL12*, *TF*, *TNFRSF10B*, *TRIB3*, *CDH5*, *IFITM3*, *SOD1*, *ATF3*, and *TUBA1B*) that exhibited significantly different expression associated with methylation changes found in both HCV- and SVR-associated HCCs. Thus, these genes have the potential to play a role in tumorigenesis, specifically after eradication of the HCV. The genes, *RORA* and *TUBA1B*, were discovered in non-tumorous liver tissue, while all of the other genes were identified in cancerous tissues. The aberrant methylation of 3 TFs (*KLF4*, *RUNX1*, and *RORA*) have been well described in cancers that suggested critical functions in carcinogenesis.[32–34] Both DNA hypomethylation and mRNA overexpression of *TUBA1B* (*tubulin alpha 1b*) have been previously observed in non-alcoholic steatohepatitis-related HCCs[35] and are associated with poor prognosis in HCCs.[36] *ATF3* is a stress-induced transcription factor that plays vital roles in modulating metabolism, immunity, and oncogenesis; it was reported to have dual functions in tumor suppression and oncogenesis of liver cancer.[37] *TF* and *CXCL12* are also known to play critical roles in tumor progression.[38,39] Additionally, *KLF4* and *TRIB3* are genes that have been associated with stemness in cancer[40,41]; *IFITM3* and *CDH5* are genes associated with TGF-beta and MAPK[42,43]; and *TNFRSF10B* is associated with apoptosis.[44] While this study did not define individual functional roles for these genes, our GO analysis suggests that apoptosis and inflammation pathways maintain epigenetic abnormalities and are associated with hepatocarcinogenesis. There have been many evidences suggesting that liver cell damage in chronic HCV infection is mediated by apoptosis. HCV core protein promoted activated anti-apoptotic signaling pathway and it might play a critical role in hepatocyte growth regulation and development of HCC mediated by chronic HCV infection.[45] There also have been many evidences that HCC is associated with chronic inflammation and fibrosis arising from various etiologies including HCV.[46] It is convincing that our combined analysis finally led us to the conclusion that abnormal pathways of apoptosis and inflammation are related to hepatocarcinogenesis.

Study limitations include the relatively small sample size; the cohorts examined contained 14 and 13 HCC cases, respectively. Although this is an exploratory study with a relatively small number of cases, we attempted to refine our analysis using derivation and validation cohorts: narrowing down significant methylation in the derivation cohort and confirming the expression level in the validation cohort. In order to actually practice our results as a clinical diagnostic biomarker, we believe that a larger-scale validation study with an increased number of cases is necessary. However, it is difficult to prospectively collect a large number of samples, since incidences of HCV-positive HCCs are decreasing, and liver biopsies are not required for an SVR. Therefore, the use of public databases would be considered for future validation studies although whole-genome methylation

sequence database of large-scale SVR-HCC has not been available yet. We did not analyze noncoding RNAs (ncRNAs) although dysregulation of ncRNA has been implicated in the pathogenesis and progression of HCC.[47] Further analysis to elucidate the function of ncRNA on SVR-HCC is warranted. Finally, this study did not show the direct functional analysis of each methylated TF on their target genes, and did not complete functional analysis of the genes identified. The main purpose of this study was not to reveal the direct transcriptional regulation of each gene, but to show the methylation landscape by comprehensive analysis in order to find methylation biomarker of SVR-HCC. It is a limitation of this paper that there is a trade-off between comprehensive analysis and detailed functional analysis of each gene to some extent. Future studies should focus on how alterations to methylation status play a role in gene expression, and how gene function relates to carcinogenesis.

Importantly, this study identified a select group of previously undiscovered genes, including TFs and their target genes, that exhibit methylation abnormalities associated with carcinogenesis in HCCs after an SVR has been achieved. We consider the genes identified herein could be employed as early serum biomarkers for the prediction and early detection of SVR-HCC, through the detection of genes with altered methylation profiles in the blood. As we revealed apoptosis and inflammation pathway were associated with SVR-HCC, we believe that it is necessary to clarify the effects of this epigenetic alteration on the immune environment in HCC in the future.

In conclusion, this study is the first to apply a genome-wide methylome profiling approach to examine an SVR in HCC tissue samples. We demonstrated that aberrant DNA methylation was retained after eradication of the HCV, particularly in functionally relevant TFs and their target genes, which may result in HCC carcinogenesis after an SVR. Our findings indicate that DNA methylation in patients who have achieved an SVR could be functionally important for carcinogenesis and may also serve as biomarkers to predict the occurrence of HCCs.

Acknowledgments

This project was financially supported by JSPS KAKENHI (18K08722 and 21K08745), the Taiju Life Social Welfare Foundation (Tokyo, Japan), and Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research [BINDS]) from AMED under Grant Number JP21am0101103 (support number 0690). The authors would like to thank M. Kasagi, K. Toyokawa, and E. Miyamoto for their technical assistance.

References

1 Kudo M, Izumi N, Kokudo N *et al*. Report of the 22nd nationwide follow-up survey of primary liver cancer in Japan (2012-2013). *Hepatol. Res.* 2021; **51**: 355–405.
2 McGlynn KA, Petrick JL, El-Serag HB. Epidemiology of hepatocellular carcinoma. *Hepatology*. 2021; **73**: 4–13.
3 D'Ambrosio R, Della Corte C, Colombo M. Hepatocellular carcinoma in patients with a sustained response to anti-hepatitis C therapy. *Int. J. Mol. Sci.* 2015; **16**: 19698–712.
4 Sugimachi K, Kinjo N, Ikebe M *et al*. Significance of hepatic resection for hepatocellular carcinoma with sustained virological response to interferon therapy for chronic hepatitis C. *Hepatol. Res.* 2013; **43**: 605–9.
5 Shirabe K, Sugimachi K, Harada N *et al*. Favorable prognosis in patients with sustained virological response to antiviral therapy, including interferon, for chronic hepatitis C before hepatic resection for hepatocellular carcinoma. *Anticancer Res*. 2015; **35**: 6963–9.
6 Lee HW, Han DH, Shin HJ *et al*. Hepatocellular carcinoma risk according to regimens for eradication of hepatitis C virus; interferon or direct acting antivirals. *Cancers*. 2020; **12**: 3414.
7 Debes JD, van Tilborg M, Groothuismink ZMA *et al*. Levels of cytokines in serum associate with development of hepatocellular carcinoma in patients with HCV infection treated with direct-acting antivirals. *Gastroenterology*. 2018; **154**: 515–7 e3.
8 Hamdane N, Juhling F, Crouchet E *et al*. HCV-induced epigenetic changes associated with liver cancer risk persist after sustained virologic response. *Gastroenterology*. 2019; **156**: 2313–29 e7.
9 Cancer Genome Atlas Research Network. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. *Cell*. 2017; **169**: 1327–41 e23.
10 Miura F, Enomoto Y, Dairiki R, Ito T. Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res.* 2012; **40**: e136.
11 Miura F, Ito T. Highly sensitive targeted methylome sequencing by post-bisulfite adaptor tagging. *DNA Res.* 2015; **22**: 13–8.
12 Suppli MP, Rigbolt KTG, Veidal SS *et al*. Hepatic transcriptome signatures in patients with varying degrees of nonalcoholic fatty liver disease compared with healthy normal-weight individuals. *Am. J. Physiol. Gastrointest. Liver Physiol.* 2019; **316**: G462–G72.
13 Juhling F, Kretzmer H, Bernhart SH, Otto C, Stadler PF, Hoffmann S. metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Res.* 2016; **26**: 256–62.
14 Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; **30**: 2114–20.
15 Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* 2011; **12**: 323.
16 Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018; **34**: i884–i90.
17 Dobin A, Davis CA, Schlesinger F *et al*. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013; **29**: 15–21.
18 Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014; **30**: 923–30.
19 Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010; **26**: 139–40.
20 Wu T, Hu E, Xu S *et al*. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation*. 2021; **2**: 100141.
21 Nanavaty V, Abrash EW, Hong C *et al*. DNA methylation regulates alternative polyadenylation via CTCF and the cohesin complex. *Mol. Cell*. 2020; **78**: 752–64 e6.
22 Bhasin JM, Lee BH, Matkin L *et al*. Methylome-wide sequencing detects DNA hypermethylation distinguishing indolent from aggressive prostate cancer. *Cell Rep.* 2015; **13**: 2135–46.
23 Lambert SA, Jolma A, Campitelli LF *et al*. The human transcription factors. *Cell*. 2018; **175**: 598–9.
24 Domovitz T, Gal-Tanamy M. Tracking down the epigenetic footprint of HCV-induced hepatocarcinogenesis. *J. Clin. Med.* 2021; **10**: 551.
25 Yin Y, Morgunova E, Jolma A *et al*. Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science*. 2017; **356**: eaaj2239.

26 Han H, Cho JW, Lee S *et al*. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* 2018; **46**: D380–D6.
27 Jee BA, Choi JH, Rhee H *et al*. Dynamics of genomic, epigenomic, and transcriptomic aberrations during stepwise hepatocarcinogenesis. *Cancer Res.* 2019; **79**: 5500–12.
28 Perez S, Kaspi A, Domovitz T *et al*. Hepatitis C virus leaves an epigenetic signature post cure of infection by direct-acting antivirals. *PLoS Genet.* 2019; **15**: e1008181.
29 Padmanabhan N, Ushijima T, Tan P. How to stomach an epigenetic insult: the gastric cancer epigenome. *Nat. Rev. Gastroenterol. Hepatol.* 2017; **14**: 467–78.
30 Villanueva A, Portela A, Sayols S *et al*. DNA methylation-based prognosis and epidrivers in hepatocellular carcinoma. *Hepatology*. 2015; **61**: 1945–56.
31 Midorikawa Y, Yamamoto S, Tatsuno K *et al*. Accumulation of molecular aberrations distinctive to hepatocellular carcinoma progression. *Cancer Res.* 2020; **80**: 3810–9.
32 Shen Y, Chen TJ, Lacorazza HD. Novel tumor-suppressor function of KLF4 in pediatric T-cell acute lymphoblastic leukemia. *Exp. Hematol.* 2017; **53**: 16–25.
33 Goyama S, Huang G, Kurokawa M, Mulloy JC. Posttranslational modifications of RUNX1 as potential anticancer targets. *Oncogene*. 2015; **34**: 3483–92.
34 Watanabe Y, Kim HS, Castoro RJ *et al*. Sensitive and specific detection of early gastric cancer with DNA methylation analysis of gastric washes. *Gastroenterology*. 2009; **136**: 2149–58.
35 Tian Y, Arai E, Makiuchi S *et al*. Aberrant DNA methylation results in altered gene expression in non-alcoholic steatohepatitis-related hepatocellular carcinomas. *J. Cancer Res. Clin. Oncol.* 2020; **146**: 2461–77.
36 Lu C, Zhang J, He S *et al*. Increased alpha-tubulin1b expression indicates poor prognosis and resistance to chemotherapy in hepatocellular carcinoma. *Dig. Dis. Sci.* 2013; **58**: 2713–20.
37 Ku HC, Cheng CF. Master Regulator Activating Transcription Factor 3 (ATF3) in Metabolic Homeostasis and Cancer. *Front. Endocrinol.* 2020; **11**: 556.
38 Hisada Y, Mackman N. Tissue factor and cancer: regulation, tumor growth, and metastasis. *Semin. Thromb. Hemost.* 2019; **45**: 385–95.
39 Teicher BA, Fricker SP. CXCL12 (SDF-1)/CXCR4 pathway in cancer. *Clin. Cancer Res.* 2010; **16**: 2927–31.
40 Cui J, Shi M, Quan M, Xie K. Regulation of EMT by KLF4 in gastrointestinal cancer. *Curr. Cancer Drug Targets*. 2013; **13**: 986–95.
41 Yu JM, Sun W, Wang ZH *et al*. TRIB3 supports breast cancer stemness by suppressing FOXO1 degradation and enhancing SOX2 transcription. *Nat. Commun.* 2019; **10**: 5720.
42 Min J, Hu J, Luo C *et al*. IFITM3 upregulates c-myc expression to promote hepatocellular carcinoma proliferation via the ERK1/2 signalling pathway. *Biosci. Trends*. 2020; **13**: 523–9.
43 Higuchi K, Inokuchi M, Takagi Y *et al*. Cadherin 5 expression correlates with poor survival in human gastric cancer. *J. Clin. Pathol.* 2017; **70**: 217–21.
44 Lee YS, Lee DH, Jeong SY *et al*. Ferroptosis-inducing agents enhance TRAIL-induced apoptosis through upregulation of death receptor 5. *J. Cell. Biochem.* 2019; **120**: 928–39.
45 Jahan S, Ashfaq UA, Khaliq S, Samreen B, Afzal N. Dual behavior of HCV Core gene in regulation of apoptosis is important in progression of HCC. *Infect. Genet. Evol.* 2012; **12**: 236–9.
46 Yang YM, Kim SY, Seki E. Inflammation and liver cancer: molecular mechanisms and therapeutic targets. *Semin. Liver Dis.* 2019; **39**: 26–42.
47 Manna D, Sarkar D. Non-coding RNAs: regulating disease progression and therapy resistance in hepatocellular carcinoma. *Cancers*. 2020; **12**: 1243.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's website:

Figure S1. Flowchart of the analysis.

Figure S2. Schematic diagram depicting the integrated analysis of the epigenome and transcriptome. (A) Schematic demonstrating that genes, located within 5 kb of an identified differentially methylated region (DMR), were interpreted as related genes. (B) Illustration summarizing the comprehensive analysis of the DMR-related genes and differentially expressed genes (DEGs) identified in the study.

Figure S3. The representative methylation status of 2 transcription factors (*KLF4* and *RXRA*). SVR, sustained virologic response; HCV, hepatitis C viral infection; T, tumorous; and N, non-tumorous.

Figure S4. Gene expression levels of DNA methyltransferases and ten-eleven translocations in the validation cohort. SVR, sustained virologic response; HCV, hepatitis C viral infection; T, tumorous; and N, non-tumorous.

Table S1. List of genes with differences in both methylation status and gene expression levels for both the HCV and SVR cases.

Table S2. List of transcription factor target genes with differences in both methylation status and gene expression levels.

Table S3. Gene Ontology pathways enriched with genes associated with HCC after sustained virologic response.