ONCOGENOMICS

A pre-neoplastic epigenetic field defect in HCV-infected liver at transcription factor binding sites and polycomb targets

NA Wijetunga[1], M Pascual[1,2], J Tozour[1], F Delahaye[3], M Alani[4,5], M Adeyeye[1], AW Wolkoff[4,5], A Verma[6] and JM Greally[1]

The predisposition of patients with Hepatitis C virus (HCV) infection to hepatocellular carcinoma (HCC) involves components of viral infection, inflammation and time. The development of multifocal, genetically distinct tumours is suggestive of a field defect affecting the entire liver. The molecular susceptibility mediating such a field defect is not understood. One potential mediator of long-term cellular reprogramming is heritable (epigenetic) regulation of transcription, exemplified by DNA methylation. We studied epigenetic and transcriptional changes in HCV-infected livers in comparison with control, uninfected livers and HCC, allowing us to identify pre-neoplastic epigenetic and transcriptional events. We find the HCV-infected liver to have a pattern of acquisition of DNA methylation targeted to candidate enhancers active in liver cells, enriched for the binding sites of the FOXA1, FOXA2 and HNF4A transcription factors. These enhancers can be subdivided into those proximal to genes implicated in liver cancer or to genes involved in stem cell development, the latter distinguished by increased CG dinucleotide density and polycomb-mediated repression, manifested by the additional acquisition of histone H3 lysine 27 trimethylation (H3K27me3). Transcriptional studies on our samples showed that the increased DNA methylation at enhancers was associated with decreased local gene expression, results validated in independent samples from The Cancer Genome Atlas. Pharmacological depletion of H3K27me3 using the EZH2 inhibitor GSK343 in HepG2 cells suppressed cell growth and also revealed that local acquired DNA methylation was not dependent upon the presence of polycomb-mediated repression. The results support a model of HCV infection influencing the binding of transcription factors to cognate sites in the genome, with consequent local acquisition of DNA methylation, and the added repressive influence of polycomb at a subset of CG-dense *cis*-regulatory sequences. These epigenetic events occur before neoplastic transformation, resulting in what may be a pharmacologically reversible epigenetic field defect in HCV-infected liver.

INTRODUCTION

Hepatocellular carcinoma (HCC) is, worldwide, the sixth most common cancer and the third leading cause of cancer-related mortality, resulting in 600 000 deaths annually.[1] As a heterogeneous disease with a multistep pathogenesis, there are several risk factors that are mechanistically implicated in the development of HCC, including infection by Hepatitis B (HBV) and Hepatitis C (HCV) viruses. HCV-related HCC has emerged as a growing public health concern in both the developed and developing worlds, with cost being a major impediment to the widespread use of new, effective antiviral therapies.[2] HCC is commonly associated with decades of chronic inflammation and cirrhosis, but can occur in as many as 20% of individuals without cirrhosis.[3] Inflammation of the liver by itself does not appear to be sufficient to induce HCC, as the tumour rarely co-occurs with chronic autoimmune hepatitis.[4] In animal models, the expression of individual HCV proteins has led to inflammation-independent tumour induction, indicating that HCV infection alone may be enough to promote oncogenesis.[5,6] As ~ 75% of HCCs are multifocal,[7] the model that emerges is one in which hepatocytes throughout the liver have acquired a similar increased potential for tumorigenesis, a 'field defect' previously suggested by altered transcription patterns in the non-malignant areas of livers in individuals with HCC.[8]

The mechanism by which an oncogenic virus induces a field defect and cancer susceptibility is not known. While DNA sequence mutations characterize the malignant HCC,[9–12] there is sparse evidence for extensive mutational events in infected, non-neoplastic hepatocytes.[13] The widespread transcriptional changes of these infected hepatocytes[8,14] indicate disturbances with broad effects in the genomes of these cells. Epigenetic regulatory mechanisms represent an attractive candidate group of mediators of such a field defect, given their potential for widespread transcriptional effects (reviewed previously[15,16]), and their observed dysregulation in malignancies (reviewed previously[17]) and pre-malignant neoplasias.[18–20] Locus-specific epigenetic events are believed to occur early during hepatocarcinogenesis,[21] and several studies have found DNA methylation changes to occur at a limited number of loci tested in pre-neoplastic lesions associated with HCV infection.[22,23] More recent studies have investigated genome-wide epigenetic changes during the progression from HCV-associated cirrhosis to HCV-related HCC, finding potential prognostic value to identifying aberrant patterns of DNA methylation at specific gene promoters, CpG islands and CpG island shores.[24,25]

The most commonly used technique for studying the epigenome in cancer is the measurement of DNA methylation at as many

[1]Department of Genetics and Center for Epigenomics, Bronx, NY, USA; [2]Centro de Investigación Médica Aplicada (CIMA), IDISNA, Oncohematology Department, Pamplona, Spain; [3]Department of Obstetrics, Gynecology and Women's Health, Bronx, NY, USA; [4]Department of Medicine (Division of Gastroenterology and Liver Diseases), Bronx, NY, USA; [5]Marion Bessin Liver Research Center, Bronx, NY, USA and [6]Department of Medicine (Oncology), Albert Einstein College of Medicine, Bronx, NY, USA. Correspondence: Dr JM Greally, Albert Einstein College of Medicine, 1301 Morris Park Avenue, Bronx NY 10461, USA.
E-mail: john.greally@einstein.yu.edu
Received 21 October 2015; revised 26 July 2016; accepted 5 August 2016; published online 10 October 2016

sites as possible throughout the genome. Changes in DNA methylation can be global, with a general decrease in levels exemplified by colorectal carcinomas (reviewed previously[26]), but can also be locus-specific, as found with silencing of tumour-suppressor genes involving the abnormal acquisition of DNA methylation at their promoters.[27,28] Local changes in DNA methylation are associated with changes of nucleosome positioning at *cis*-regulatory elements in cancer.[29] A further observation in non-neoplastic cells is that the changes of DNA methylation occur associated with the presence or absence of transcription factor (TF) binding to that locus.[30,31] The causality of the relationship between DNA methylation and transcription factor binding is complex,[15,32,33] but the model linking these events is relatively straightforward – DNA methylation changes at *cis*-regulatory elements (which are defined by transcription factor binding) represent a footprint of changes of transcription factor binding. Distal (non-promoter) *cis*-regulatory elements are also being recognized to be the more informative loci for DNA methylation changes associated with transcriptional regulatory functions,[34–36] cancers[37,38] and other diseases.[39]

A problem in interpreting DNA methylation changes from genome-wide assays is that the locations of distal *cis*-regulatory elements are very variable between cell types[40,41] and not readily interpretable from fixed genomic annotations like the positions of genes, CpG islands or their shores.[42] CpG islands, regions with unusually high density of CG dinucleotides,[43,44] have been found to acquire increased DNA methylation in some tumours[45,46] (reviewed previously[47,48]). They also correspond to bivalent chromatin domains in pluripotent cells,[49] loci that are targets of the polycomb repressive complex in early development that are believed to be unusually prone to the acquisition of DNA methylation later in life as part of oncogenesis.[50–52] As genome-wide DNA methylation assays have begun to test more sites in the genome, it is now becoming apparent that DNA methylation changes appear to target *cis*-regulatory sites more frequently outside than within CpG islands.[42,53,54] We have previously shown that CpG island shores, where DNA methylation is unusually informative for changes associated with cancer or local transcription levels,[40,42,55,56] are enriched for chromatin features indicating enhancer functions.[57] However, these would represent only a proportion of the *cis*-regulatory elements within a cell, with many regulatory sequences located far from genes or CpG islands. To overcome this problem, publicly accessible data have been generated in which chromatin features have been mapped in primary cells, tissues or cell lines, allowing some insights into where *cis*-regulatory loci are likely to be located when studying a tissue like liver, and improving the ability to interpret changes of functional consequence.

The extent and significance of transcriptome dysregulation in HCV-associated HCC is relatively unknown, but gene expression analyses of HCCs from multiple etiologies show non-random effects, selecting certain genes with associated protein functions and pathway relationships. These include associations between poor survival and signatures related to aberrant cell signalling,[58–60] cell proliferation[61–64] and differentiation,[59,65–67] inflammation,[64] and angiogenesis.[58,63,68] Transcriptional studies of the pre-neoplastic liver tissue have found altered expression of genes associated with inflammation,[69,70] hepatic progenitor cells[71] and pathways involving Wnt/β-catenin[72] and hedgehog[14] signalling. These results suggest that a study of transcriptional regulatory processes in HCV-associated pre-neoplastic and malignant tissues should show functional coherence in terms of the loci targeted for dysregulation.

In this study, we use patient biopsies from uninfected individuals and from individuals with HCV infection and HCC, allowing us to compare HCV-infected non-neoplastic and HCV-infected malignant samples with control liver samples, creating a representation of the natural course of the disease. We integrated high-resolution DNA methylation studies with transcriptional profiling and cell-type specific reference epigenomic profiles, generating evidence for DNA methylation changes targeted to the binding sites of a group of specific transcription factors. These loci with altered DNA methylation are enriched near genes encoding proteins already associated with HCC but also include a group with stem cell functions. We find the DNA methylation changes to occur predominantly at the pre-neoplastic stage, with continued functional consequences in HCV-infected tumour tissue, and an association with independent polycomb-mediated repression at a subset of loci. Data from The Cancer Genome Atlas (TCGA)[73] are used as a replication cohort to validate our findings. The study reveals a relationship between the dysregulation of transcription factor activity, DNA methylation and transcription occurring before malignant transformation, potentially contributing to the field defect that characterizes the chronically HCV-infected liver.

RESULTS

DNA methylation changes in HCV-infected livers precede neoplastic transformation

We tested DNA methylation at ~2 million CpGs throughout the human genome in control (uninfected, non-neoplastic), infected (HCV+, non-neoplastic) and malignant (HCV+, HCC) samples (Supplementary Table S1). The infected and malignant samples were from different sites within the same individual's liver. To get an overview of the DNA methylation patterns characterizing control, infected and malignant samples, an unsupervised hierarchical clustering was performed of the 5000 loci with the most variable DNA methylation. The DNA methylation patterns revealed seven of the nine malignant samples to have a subset of loci with distinctively decreased DNA methylation (group A, Figure 1). The infected samples also tended to cluster together but were less distinctive from the control samples, with a different subset of loci showing a relatively moderate increase in DNA methylation compared with controls (group B, Figure 1). These results suggested distinct steps in a progression of DNA methylation changes in HCV-infected liver and associated HCC, prompting more detailed analyses.

A K-means clustering approach was used to identify the subsets of loci with distinctive patterns of DNA methylation change associated with a progression from control to the infected and malignant states. The optimum number of clusters was calculated to be approximately 8 (see Supplementary Methods, Supplementary Figure S1), allowing the grouping of loci into those with no changes of DNA methylation and those with early or later (infection or malignancy-associated) changes (increases and decreases). Genome-wide, the majority of loci tested (69.6%) show no changes in DNA methylation (groups I and V, Supplementary Figure S2), but there is a significant subset of loci with both early and late changes in DNA methylation. The transition from infected to malignant states is associated with subsets of loci making late changes in DNA methylation (7.8% of loci, groups II and VI, Supplementary Figure S2). The analysis also reveals a larger number of loci that make early and sustained changes in DNA methylation (17.1% of loci, groups III, IV and VII, Supplementary Figure S2), and a subset that appears to revert in the malignant samples to the control state (5.5% of loci, group VIII, Supplementary Figure S2).

The changes of DNA methylation were explored further using polytomous regression modelling at the individual CG dinucleotide level. This technique is effective for modelling relative DNA methylation proportions, is robust for detection of linear and non-linear DNA methylation events, and allows for adjustment for potential confounding covariates. A principal components analysis showed that subject age and race were associated with DNA methylation variability (Supplementary Figure S3), prompting the
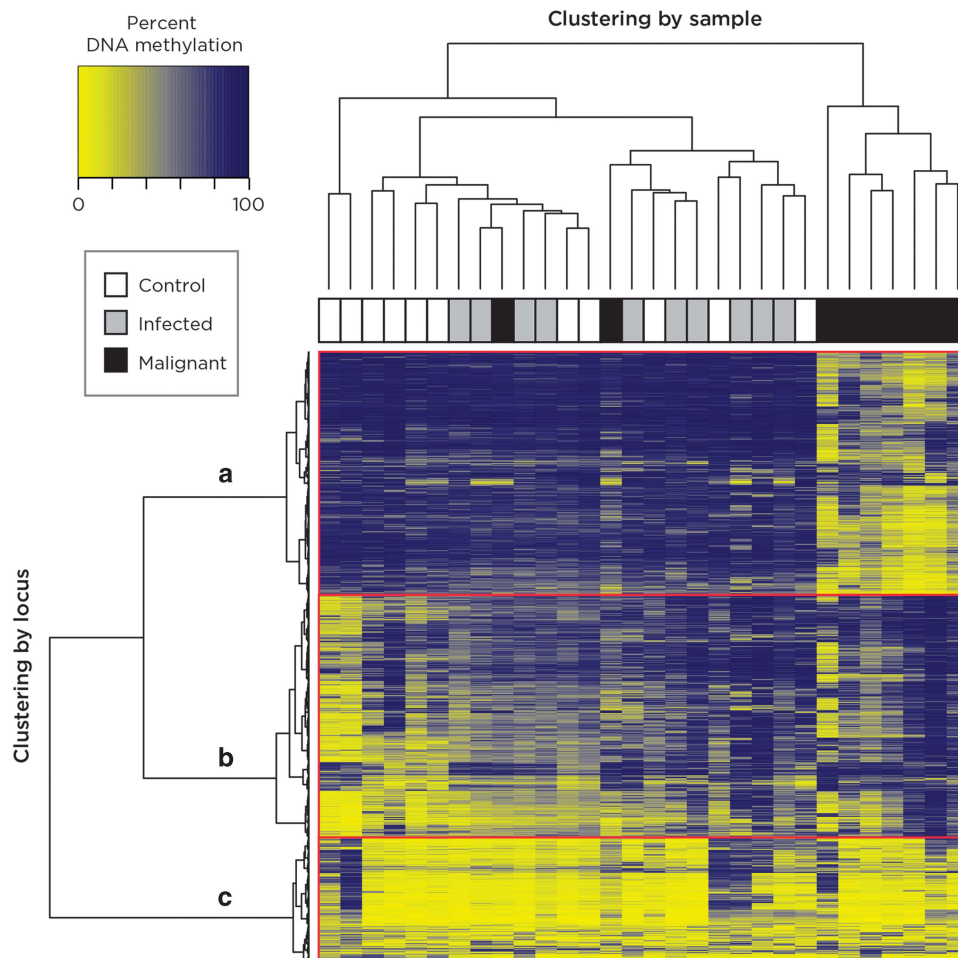
Figure 1. Unsupervised hierarchical clustering of DNA methylation. The 5000 loci with the most variable DNA methylation were used to explore DNA methylation changes in the 30 samples studied. The malignant (black) samples form a major cluster of the three clusters (**a**–**c**), the malignant (black) samples form a major cluster (**a**), largely due to the Group A loci that appear strikingly less methylated in the malignant samples, and a less striking distinction of infected from control samples.

inclusion of these variables in regression models to estimate the adjusted odds of DNA methylation for both infected and malignant samples relative to the control samples for each of the ~ 2 million assayed CG dinucleotides.

This analysis confirmed the impression of the heat map of Figure 1, that the malignant samples have a greater number of loci with significant loss (16,921) than gain (6894) of DNA methylation relative to control samples (Figure 2a). Linking the identified K-means clustering groups with this restricted set of loci shows the loss of DNA methylation to be a late event, mostly occurring in the transition from the infected to malignant states (groups V–VI, Figure 2b), whereas the smaller subset of loci gaining DNA methylation does so at the earlier, infected stage (groups III–IV, Figure 2b).

Targeting of DNA methylation changes to specific genomic contexts

Having identified loci with distinctive patterns of change of DNA methylation in the progression to HCV-associated HCC, it was possible to ask whether these changes were targeting certain types of loci in the genome. Genomic annotations were explored that are not cell type-specific (RefSeq gene promoters and bodies and intergenic regions, CpG islands and shores[42]) as well as liver-specific annotations based on publicly available chromatin immunoprecipitation sequencing (ChIP-seq) data from the NIH

Roadmap Epigenomic Mapping Consortium for three normal adult livers (histone H3 lysine 4 trimethylation (H3K4me3), H3K4me1, H3K27ac, H3K36me3 and H3K27me3). The significance of the enrichment or depletion of loci with altered DNA methylation for each annotation was determined using randomized permutation tests.

We studied the loci defined in Figure 2a as having significantly increased or decreased DNA methylation. The loci with increases in DNA methylation are significantly enriched at enhancers, characterized by local enrichment of H3K4me1 and H3K27ac (but not the promoter-associated mark H3K4me3), as well as loci likely to be undergoing active transcription (gene bodies and H3K36me3-enriched loci, Figure 3). In contrast, decreased DNA methylation is only significantly enriched at intergenic regions and not at any candidate regulatory elements (Figure 3). As increasing DNA methylation at enhancer sequences is potentially reflective of local transcription factor binding,[30,32] the permutation analysis was repeated using ChIP-seq data for transcription factor binding sites mapped in the HepG2 HCC cell line as part of the ENCODE project,[74] the best available hepatocyte surrogate with extensive transcription factor mapping. No significant associations were observed for loci with loss of DNA methylation, but permutation testing revealed significant enrichment for several transcription factors at loci that gain DNA methylation: FOXA1, FOXA2, HNF4A, MAFK, MAFF, CEBPB and RXRA (Figure 4). No significant changes in
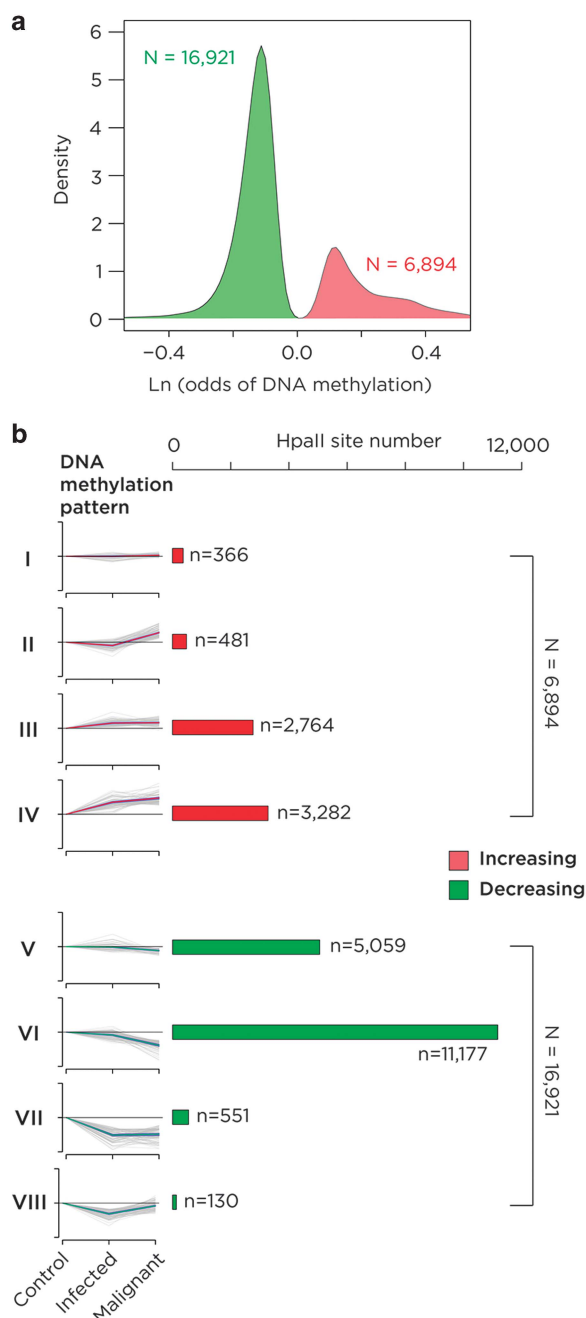
Figure 2. Identification of subsets if loci with distinctive progressive changes in DNA methylation. Panel (**a**) shows the comparison between malignant and control samples, with the majority of loci losing DNA methylation (green) with malignant transformation. In panel (**b**), K-means clustering shows the majority of loci gaining DNA methylation to do so in the infected stage (groups III and IV), but the loci losing DNA methylation only have this occur in the transition from infected to cancer (groups V and VI).

FOXA1/2 and HNF4A transcription factors have also been implicated in WNT signalling pathways that promote carcinogenesis and epithelial-to-mesenchymal transition,[88–91] which links the DNA methylation changes to the risk of neoplastic progression.

Peripheral blood B lymphocytes do not show DNA methylation changes in HCV infection

A question that arises is whether the changes of DNA methylation in the infected, non-malignant samples represent cellular responses to the chronic inflammatory state and are not due to the HCV infection or neoplastic process. To test this, peripheral blood CD19+ B lymphocytes were chosen as a reporter cell type that is not only readily accessible but also appears to be unusually influenced by chronic HCV infection, judging by the known associations of chronic HCV with mixed cryoglobulinemia, rheumatoid factor production and lymphoproliferative disorders including monoclonal gammopathies and lymphomas.[92] DNA was isolated from CD19+ B lymphocytes from an independent cohort of 10 chronically HCV-infected patients and 10 HCV-negative controls. The global DNA methylation profiles of the HCV-infected patients were not obviously different to those of uninfected patients (Supplementary Figure S5). Furthermore, after adjusting for multiple comparisons, there were no significant DNA methylation events associated with the inflammatory environment of chronic HCV infection at a false discovery rate (FDR) < 0.05. The results do not support the chronic inflammatory process itself inducing DNA methylation changes as a generalized effect in the body.

Enhancer-targeted DNA methylation is associated with decreased gene expression

As the acquisition of DNA methylation at enhancers is potentially associated with altered gene expression,[38,42] RNA-seq was therefore performed on the same samples on which DNA methylation studies had been performed. K-means clustering was used in the analysis of these data, with the optimal number of groups found to be approximately 6, allowing patterns of gene expression changes in infected and malignant samples to be compared with the controls (Supplementary Figure S1). In Figure 5a it is apparent that most genes (groups I and II) do not change levels of expression in infected or malignant cells ($n = 13\,507$), but that there are also groups of genes with early (group IV, $n = 1720$) and progressive (group III, $n = 4621$) patterns of increased gene expression, and further genes that progressively decrease expression levels by substantial (pattern VI, $n = 890$) or lesser extents (pattern V, $n = 4375$) (Figure 5a). Overall, more genes were represented by patterns with increased rather than decreased gene expression. A principal components analysis found potential covariates modifying gene expression, allowing subsequent polytomous regression to control for sequencing batch (Supplementary Figure S6). In all, 309 genes were identified to have significantly decreasing expression and only 193 genes with increased expression levels (Figure 5b, Supplementary Tables S2–S3). The 309 genes with significantly decreased expression were tested for proximity to the enhancers found to have increased DNA methylation. A conservative approach was used, linking a candidate enhancer (H3K4me1 positive, H3K4me3 negative) to a gene if the chromatin state was within 5 kb of the gene's transcription start site, finding a total of 644 genes fitting these criteria. A significant association was found linking enhancers with increased DNA methylation and genes with decreased expression levels (hypergeometric test $P = 0.002$), with no significant association found for the genes with increased expression levels (hypergeometric test $P = 0.763$) (Figure 5c). While influencing only a small proportion of genes, the increased DNA methylation at candidate enhancers is significantly associated with local transcriptional repression.
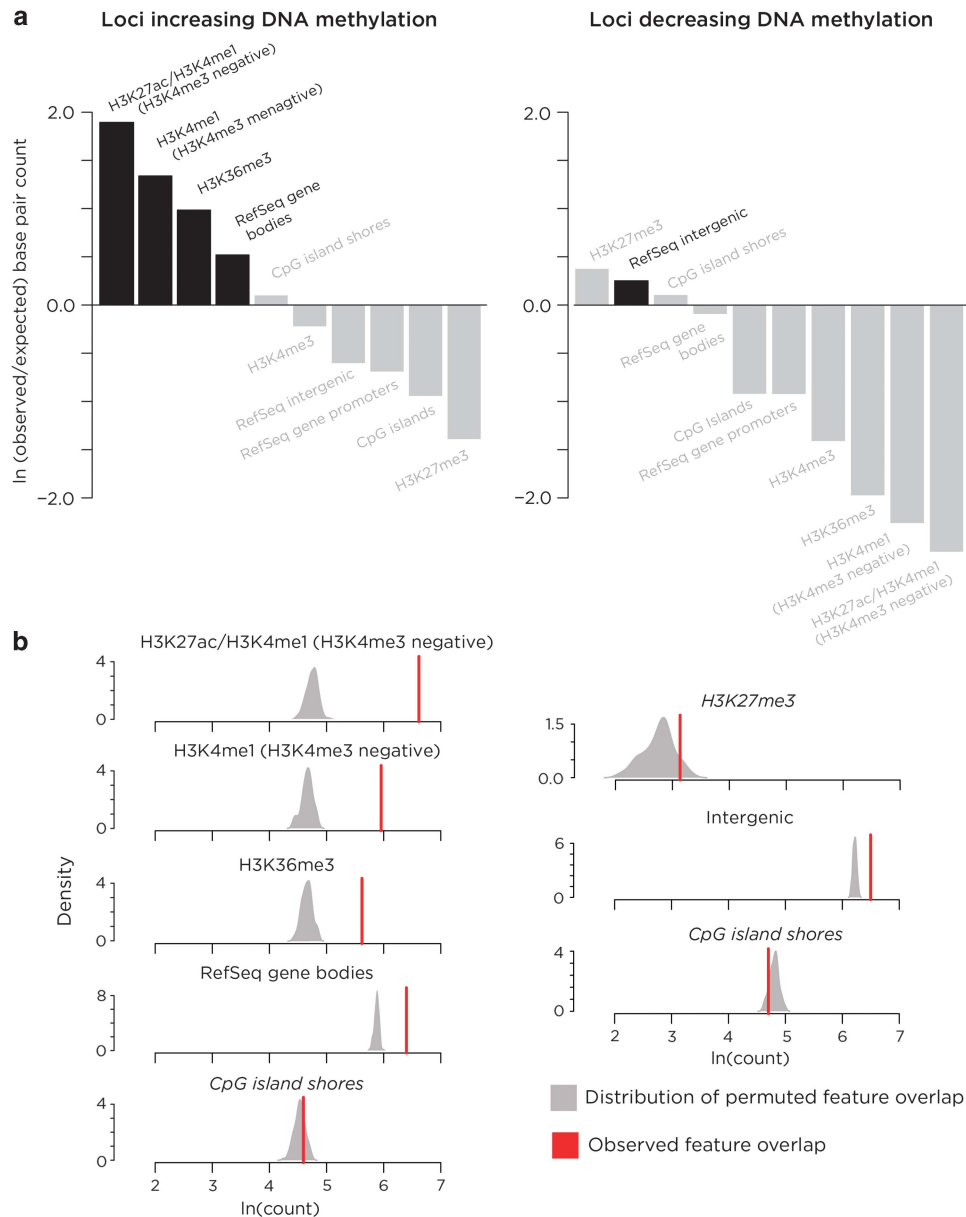
levels of expression of these genes were found at alpha = 0.05 (Supplementary Figure S4), so this acquisition of DNA methylation is not attributable to a simple model of down-regulation of production of these transcription factors. All seven transcription factors are involved in normal liver development and physiology,[75–80] with CEBPB, RXRA and MAF transcription factors also involved with inflammation and cellular stress.[81–87] The most strongly perturbed

Figure 3. The genomic contexts of loci gaining or losing DNA methylation. In panel (**a**) we show the observed/expected ratio for overlap of the loci with distinctive DNA methylation with different genomic features. The loci with increasing DNA methylation are enriched at candidate enhancers (with histone H3 lysine 4 monomethylation H3K4me1 and H3K27ac but not co-incident H3K4me3 that would indicate promoter function). These loci are also enriched at regions likely to be transcribed (with H3K36me3 and at RefSeq gene bodies). These enrichments were tested by permutation analyses (**b**), revealing them to be non-random (black). Loci where DNA methylation is lost during disease progression is enriched at several genomic features (**a**) but only that for intergenic sequences survives the permutation analysis of significance in panel (**b**).

Validation of findings using TCGA data

The Cancer Genome Atlas (TCGA) includes HCC as one of its studied tumour types, and includes both DNA methylation and expression studies,[73] testing the HCC tumour itself plus adjacent liver, thereby mimicking the malignant and infected samples of the current study in its HCV-related HCC tumour cohort ($n = 5$). We compared our results with those from these TCGA individuals, comparing their DNA methylation profiles between infected and malignant samples at the loci identified in the current study to have increased DNA methylation at enhancers. In Figure 5d we show that these loci are also significantly increased in DNA methylation in the TCGA subjects. When we focused on the genes identified to be significantly altered in expression in our study (Figure 5a), the 193 genes with increased expression do not show

significant differences in expression in TCGA data, but the 309 genes with decreased expression are also significantly decreased in expression in TCGA subjects (Figure 5d). We conclude that the changes in DNA methylation at enhancers and the downregulated genes that we find in the current study are robust and high confidence findings.

Epigenetic and transcriptional dysregulation target genes associated with liver cancer and stem cell functions

To understand the potential cellular consequences of altering transcription levels and dysregulating genes by DNA methylation changes, Gene Set Enrichment Analysis[93] was performed focusing on a comparison between control and malignant samples. Three
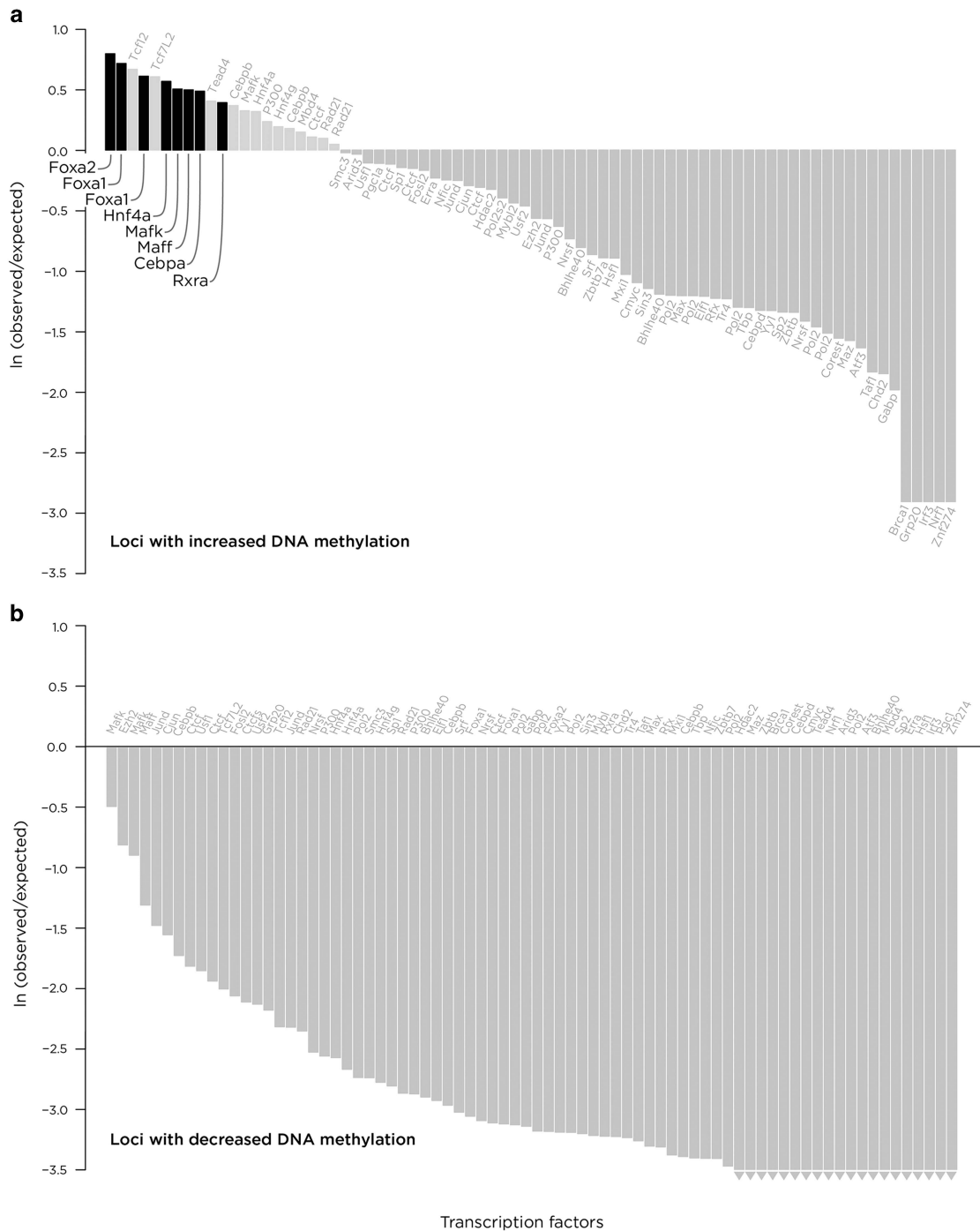
Figure 4. Overlap of loci with distinctive DNA methylation with sites of transcription factor binding in HepG2 cells. Whereas no loci with decreasing DNA methylation are enriched at transcription factor binding sites (**b**), those gaining DNA methylation (**a**) non-randomly (black) overlap the sites of binding of several transcription factors, FOXA2, FOXA1, HNF4A, MAFK, MAFF, CEBPA and RXRA.

lists of genes were generated – the 309 downregulated and the 193 upregulated genes, and the 644 genes where increased DNA methylation was found at H3K4me1-defined enhancers within 5 kb of the transcription start site (Supplementary Table S4). All three gene lists showed concordance for enrichment in gene sets related to altered expression in liver cancer, fetal liver expression and polycomb repressive complex 2 (PRC2) targeting in ES cells (Supplementary Figure S7). Interestingly, two of the gene sets were published as targets of the HNF4A and FOXA2 transcription factors in liver,[94,95] supporting the possibility illustrated in Figure 4

that the regulation exerted by these transcription factors is specifically targeted in HCV infection and HCC. Because we found pathway-enrichment for PRC2 gene targets associated H3K27me3 in embryonic stem cells, the pathway analysis for 644 genes with H3K4me1-defined enhancers in normal adult liver was stratified by the presence of H3K4me1 and H3K27me3 in HepG2, using publicly available ChIP-seq data. Of the 644 genes defined to have increasing DNA methylation at promoter-proximal enhancers, 326 retained those enhancers in HepG2 cells, but another 137 added the repressive H3K27me3 mark at those enhancers,
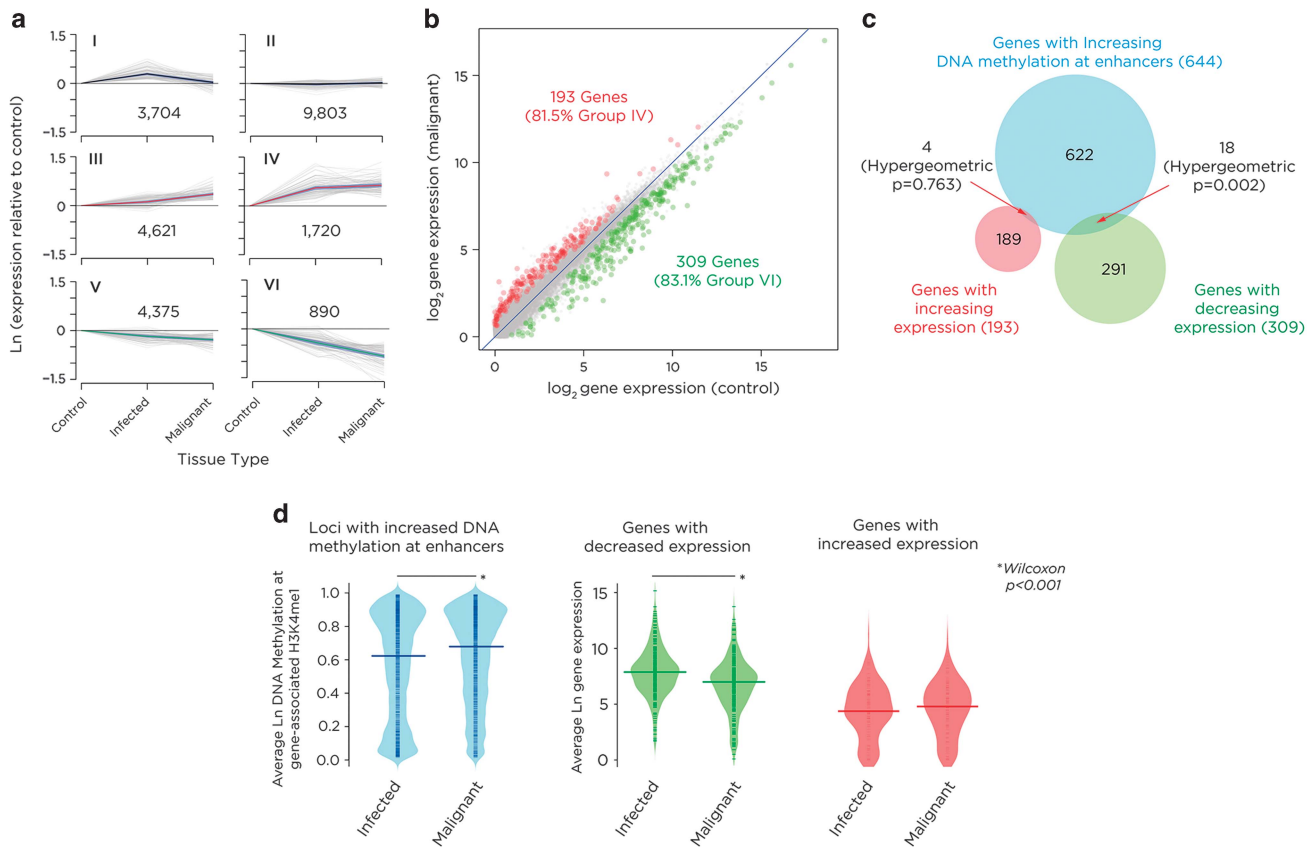
Figure 5. Analysis of gene expression data from the samples tested for DNA methylation changes. In (**a**) we again use K-means clustering to identify subsets of genes with distinctive progression of expression patterns during disease progression. Groups III and IV represent genes increasing their expression while groups V and VI show a decrease in expression levels. The significantly differentially expressed genes in a comparison of malignant and control samples (**b**) are mostly accounted for by genes in groups IV and VI. We associated a candidate enhancer (H3K4me1 positive, H3K4me3 negative) to a gene if the chromatin state was within 5 kb of the gene's transcription start site and showed increased DNA methylation in control versus malignant samples, finding a total of 622 genes fitting these criteria. When we tested to see whether these 622 include genes with significantly altered levels of expression, we showed that these genes with increased DNA methylation were significantly enriched for overlap with genes with decreased expression ($P = 0.002$) but not increased expression (**c**). In (**d**) we show the results of analysis of matched HCC and infected liver samples from HCV+ individuals, using data from The Cancer Genome Atlas (TCGA). We tested DNA methylation data from TCGA at candidate enhancers where we had identified increased DNA methylation and found the studies from TCGA to reveal a significant increase of DNA methylation in their samples also. We also compared the genes where we had found significant changes in levels of expression, showing the genes with decreased expression also to have significantly lower levels in TCGA samples, but no significant changes in the levels of expression of genes that we had found to be upregulated with disease progression.

while 63 genes gained H3K27me3 and lost the H3K4me1 mark altogether. This sub-categorization allowed further insights – the genes that do not gain the H3K27me3 mark are more likely to represent those previously associated with liver and other cancers, while those that do gain the H3K27me3 modification represent known targets of polycomb and other transcriptional regulators, especially in stem cells and early development (Figure 6a). When the CG dinucleotide content of the enhancers in each sub-category was tested, it became strikingly apparent that the acquisition of H3K27me3 and the loss of H3K4me1 were positively associated with increased CG content (Supplementary Figure S8). The cancer gene signature is therefore associated with acquisition of DNA methylation at CG-depleted loci, and the stem cell signature with acquisition of DNA methylation and H3K27me3 at CG-dense loci.

We then tested how integration of the transcriptional, DNA methylation and histone modification information could increase our power to detect pathway-level events in an unbiased manner. We applied our recently developed *SMITE* integrative network approach (manuscript in review) to derive gene scores using gene expression differences combined with DNA methylation measurements at gene promoters, gene bodies and gene-associated

H3K4me1 peaks. By testing for high-scoring sub-networks within a gene network, we identified 12 genomic modules, each with multiple interacting genes enriched for transcriptomic and epigenetic dysregulation. Using Kyoto Encyclopedia of Genes and Genomes pathway information, genes within each module were annotated to determine the likely functions of the modules. Out of the 12 genomic modules, 10 strongly associate with Kyoto Encyclopedia of Genes and Genomes annotations: seven related to liver metabolism, and one each to the cell cycle, immune system processes and to WNT signalling (Supplementary Table S5). Specifically, a module built around telomerase reverse transcriptase (TERT), transducin-Like enhancer of split (TLE2/4), pygopus family PHD fingers (PYGO1/2) included 10 WNT signalling-enriched genes, consistent with the acquisition of a stem cell-like phenotype in the HCV-related HCC genome (Figure 6b).

Polycomb inhibition suppresses cell growth but does not influence DNA methylation

We were interested in whether the increased DNA methylation at CG-dense enhancers that acquire H3K27me3 was associated with a decreased gene expression independently of the presence of
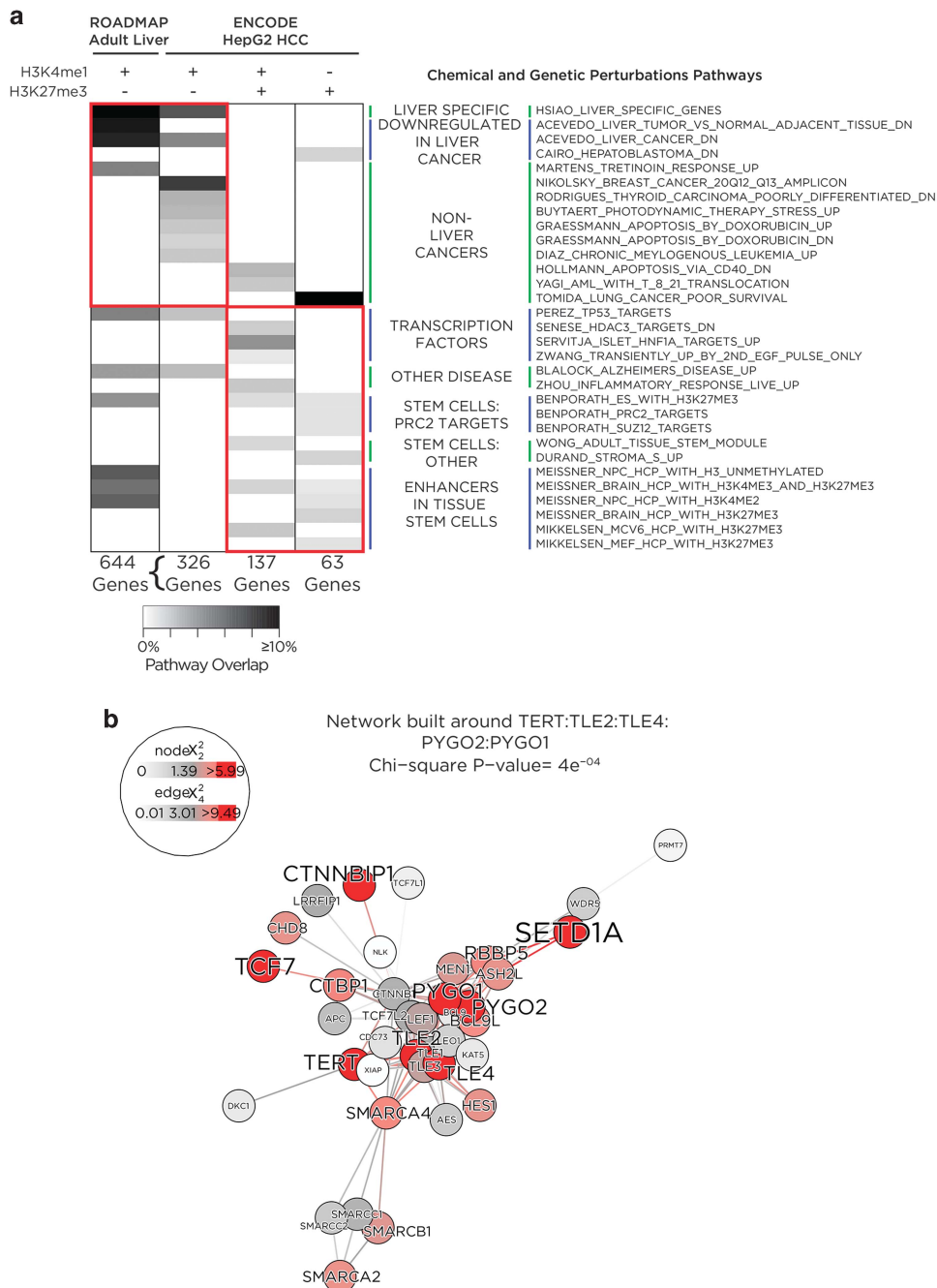
Figure 6. Functional properties of genes targeted for DNA methylation and transcriptional changes. We focus in (**a**) on the 644 genes with gain of DNA methylation during disease progression at promoter-proximal enhancers, but breaking down the associated pathways depending on whether the hypermethylated enhancer loci with H3K4me1 in normal liver continued to have this mark in the HepG2 HCC cells, and whether the repressive H3K27me3 polycomb mark is added at those sites in HepG2 cells. We show that the associated functional pathways are segregated into two groups (red boxes) when these patterns of change of histone modifications are studied. Genes acquiring DNA methylation at nearby H3K4me1-defined enhancers that do not lose this mark or gain polycomb repression are those associated with cancers in general and liver cancers in particular. The genes that acquire H3K27me3 are those encoding transcription factors and with stem cell properties. In (**b**) we show a result from our SMITE analysis based on integration of transcriptional information with genomic context-dependent DNA methylation changes. The module shown is highly enriched for proteins involved in WNT signalling, consistent with the targeting of genomic dysregulation to genes with stem cell functions.

H3K27me3. To test the dependence of gene expression on the presence of H3K27me3, we treated HepG2 cells with a PRC2-EZH2 inhibitor, GSK343,[96] and found that there was a decrease in cellular proliferation (Supplementary Figure S9), a finding consistent with a prior study involving a shRNA-mediated knockdown of *EZH2*.[97] We used ENCODE data to identify loci where

H3K27me3 was enriched in HepG2 cells and compared enrichment at these loci with negative control loci using quantitative PCR of chromatin immunoprecipitated DNA (qChIP). We showed that the GSK343 treatment of the HepG2 cells depleted H3K27me3, as expected. When we then tested DNA methylation at these loci using bisulphite MassArray[98] and local gene expression using qRT-

2038

PCR, we found both to be unaltered in the presence of local depletion of H3K27me3 (Supplementary Figure S9). The results indicate that, at least at the loci tested, the increased DNA methylation and transcriptional repression are not dependent upon the presence of polycomb-mediated H3K27me3 repressive marks, but that polycomb inhibition exerts repressive effects upon the proliferation of the HepG2 HCC cell line.

DISCUSSION

Our results show an intriguing pattern of acquisition of DNA methylation in HCV-infected livers prior to malignant transformation, at which time a second wave of global demethylation occurs. Whereas the global demethylation associated with malignancy is enriched at un-annotated intergenic sequences, the pre-neoplastic, infection-associated acquisition of DNA methylation is found at candidate *cis*-regulators of transcription. This pattern of global hypomethylation and local hypermethylation is considered typical of oncogenesis.[99] Prior studies have also found acquisition of DNA methylation at promoters and other *cis*-regulatory elements in liver prior to the development of HCC,[22,23,100–102] but were focused on individual loci and did not perform the genome-wide survey of our study. We defined the loci regulating transcription using fixed annotations (RefSeq gene promoters, CpG islands and shores) as well as tissue-specific annotations from ChIP-seq data of histone modifications and sites of transcription factor binding in primary liver samples or the HepG2 cell line. Our results support a model of interaction between transcription factor binding and local DNA methylation, with a possible model involving a primary effect on transcription factor binding resulting in DNA methylation changes as the downstream effect, similar to the processes observed during normal cell differentiation.[30,32] Transcriptional data from the same samples links the alterations in DNA methylation at *cis*-regulatory sites with local gene expression changes, and that the genes affected are those previously found to be associated with liver cancers as well as genes involved in hepatic and other types of tissue development. Furthermore, we show that the acquisition of the H3K27me3 repression-associated histone modification is enriched at loci where candidate enhancers (H3K4me1 positive) are lost in the HepG2 HCC cell line compared with normal liver.

The major goal of this study was to gain insights into events that create an organ-wide susceptibility to neoplastic transformation, reflected by tumour multifocality in the majority of individuals with HCC.[7] We have previously found that these kinds of tumour precursor states are associated with epigenetic dysregulation, in monoclonal gammopathy of undetermined significance as a precursor to multiple myeloma,[103] and in Barrett's esophagus as a field defect preceding esophageal adenocarcinoma.[104] A mouse model of colitis that predisposes to neoplasia also showed such early epigenetic events preceding neoplasia.[105] The current study used the HCV-infected state to represent the pre-neoplastic hepatocytes in which a field defect allows multifocal HCC to develop. Unlike previous studies in which DNA methylation was compared between the HCC and adjacent liver,[25,106–109] we added uninfected liver as a control, and restricted our study to HCV-infected individuals, with the assumption that tissue adjacent to the HCC does not, in fact, represent normal liver. By creating three stages representing disease progression, we could identify the early (infection associated) and late (malignancy associated) epigenetic and transcriptional events. The epigenetic changes included gains and losses of DNA methylation, with the losses associated with malignant transformation occurring at intergenic regions of uncertain functional significance, but the gains in DNA methylation taking place earlier, in the infected, pre-neoplastic state, and at *cis*-regulatory elements. The genes near the candidate enhancers identified to increase DNA methylation show a decrease in expression levels, with both epigenetic and transcriptional findings replicated in an independent TCGA cohort.

In pursuing the hypothesis that gains in DNA methylation are due to the lack of local binding of transcription factors,[110,111] we were able to implicate a group of transcription factors that have been implicated in WNT signalling and cancer. Why these transcription factors would fail to target their binding sites in infected cells is not known, but as the levels of expression of the genes encoding these specific transcription factors do not change, the amount of transcription factor is likely to be less influential than some qualitative property of these transcription factors or of the chromatin at these binding sites. As might be expected for genes regulated by specific transcription factors, the effect of increased DNA methylation is quite specific in terms of the genes selected for dysregulation, enriching for genes already implicated in liver cancer, but also a set of genes associated with polycomb targeting in stem cells or early development. The enhancers at these two groups of genes have different patterns of polycomb-mediated influences and DNA methylation in HepG2 cells, with the cancer genes acquiring just DNA methylation, but the developmentally important genes also acquiring the repressive H3K27me3 mark. Sequential ChIP-bisulphite studies have revealed a complex relationship between DNA methylation and H3K27me3, with mutual exclusivity of the two marks in CG dinucleotide-dense regions, but their co-localization in the remainder of the genome[112,113] and perturbations of this relationship in cancer.[112] Biochemically, the polycomb complex protein EZH2 has been found to associate with DNA methyltransferases,[114] but recruitment by EZH2 of DNA methyltransferase3a to a specific site has been found to be insufficient to cause local DNA methylation.[115] Our use of pharmacological inhibition of the polycomb protein EZH2 did not affect DNA methylation at sites of depletion of H3K27me3, indicating that the acquisition of DNA methylation is not dependent upon the presence of polycomb-mediated repressive mechanisms locally.

For an epigenome-wide approach to show changes in one group of samples and not another, the changes need to be recurring at the same loci in multiple individuals, and not randomly within the genome. This consistency of location of changes is of interest from a mechanistic standpoint – in cancer it could certainly reflect a primary event of changes at random locations followed by selection of cells with patterns favouring uncontrolled cell division. However, the alternative hypothesis is that the changes occur at specific loci because sequence-specific transcriptional regulators are altering their functions. In the human genome, sequence specificity of transcriptional regulation is best recognized to be mediated by transcription factors, which bind to definable DNA sequences. It has been proposed that many events currently referred to as 'epigenetic' are, in fact, due to primary effects mediated by transcription factors,[116,117] allowing the frequently studied epigenetic regulator DNA methylation to be viewed as a way of 'footprinting' where transcription factor binding is occurring in the genome.[15,30,32] It has been recognized for some time that HCV interacts with host transcription factors, the HCV NS5A protein binding to the host transcription factors SRCAP,[118] TBP and p53,[119] and the virus affecting the subcellular localization and function of forkhead transcription factors through alteration of post-translational modifications.[120] In the current study, the results are consistent with the HCV infection influencing the DNA binding of forkhead transcription factors (FOXA1, FOXA2) as well as HNF4A, MAFK, MAFF, CEBPB and RXRA, causing the loci normally bound by these transcription factors to acquire DNA methylation in infected cells. We find that the changes in DNA methylation are enriched at genes with pre-neoplastic transcriptional changes, and that independent samples from TCGA reveal similar patterns of DNA methylation and gene expression changes. Our model is therefore one of transcription factors having a primary role in mediating the transcriptional regulatory

reprogramming, including epigenetic regulation, that occurs in neoplastic transformation.

Our parallel study of B lymphocytes indicates that the transcriptional dysregulatory events are not constitutive throughout the body and reflective of a generally inflammatory state, but instead occur specifically in the cell type infected by HCV. B lymphocytes should have represented a cell type with a higher likelihood of being informative for HCV-associated cellular perturbations, given the unusually high frequency of a spectrum of B cell lymphoproliferative events associated with HCV infection.[92] The easy accessibility of these cells from peripheral blood also raised the possibility that they could serve a role as biomarkers of events occurring in the liver. Instead, we see no significant DNA methylation or transcriptional changes characterizing the B lymphocytes from HCV-infected individuals, indicating that the transcriptional regulatory effects observed in liver are more likely to be due to the HCV infection in their host cells and not to the chronic inflammatory process.

A limitation of our study is that the extensive transcription factor mapping by ChIP-seq is only available from HepG2 cells, which are derived from an HCC, and as such do not fully represent normal hepatocytes. However, the problem likely to arise when using these tumour-derived cells (as opposed to normal hepatocytes) is likely to be one of sensitivity – if tumour-associated changes have also taken place in HepG2, a comparison with uninfected, normal hepatocytes would probably have greater potential to find transcription factors enriched at sites found to be differentially methylated in our study. The ChIP-seq mapping data for H3K4me1 (candidate enhancers) and H3K27me3 (polycomb-mediated repressive chromatin) allowed us to exploit data from normal liver (Roadmap Epigenomics data), with HepG2 (ENCODE data) this time serving as an HCC model, allowing us to group the genes we found to be affected by DNA methylation and transcriptional changes. We find that the H3K4me1 enhancers in normal liver that gain DNA methylation in pre-neoplastic liver can continue to have this chromatin mark in HCC, or can acquire H3K27me3 at the same loci (to create 'poised' enhancers[121,122]) or can lose the H3K4me1 mark altogether, with only H3K27me3 remaining. The first group, in which the enhancers acquire DNA methylation but remain free of polycomb-mediated inactivation, is enriched for genes involved in liver and other cancers, while those that acquire the polycomb-mediated suppressive H3K27me3 mark are known targets of polycomb in stem cells and genes expressed in tissue progenitor cells. These links between polycomb and cell lineage commitment[123] and with targeting of DNA methylation in cancer[50–52] have been appreciated for some time. Our results suggest that the polycomb stem cell mark is targeted to a subset of genes undergoing DNA methylation changes in neoplastic transformation, and is not responsible for all targeting of DNA methylation observed.

The role of mutations is not addressed in the current study. However, there is little prior evidence for mutations occurring in infected, non-neoplastic hepatocytes,[13,124,125] suggesting that the driver mutations for HCC occur during and are responsible for the transition from the pre-neoplastic to the malignant stage. When HCC arises simultaneously in different parts of the infected liver, two potential mechanisms are possible, one in which a primary single tumour gives rise to intrahepatic metastases, the other in which there are multiple independent malignant transformation events. The latter model suggests a field defect throughout the liver, with current evidence suggesting that a substantial proportion of cases of multicentric HCC is due to such independent events (reviewed previously[126]). The intriguing model that arises is that some mutations may only be capable of promoting neoplastic transformation when the transcriptional regulatory and epigenetic properties of the cell have already been altered as we find in the infected hepatocytes. This finding opens up the possibility that the use of pharmacological agents targeting mediators of epigenetic repression may be worth exploring as chemoprevention in individuals during the long latency period of chronic HCV infection.

MATERIALS AND METHODS

Patient tissue samples for DNA methylation and gene expression studies

This study was approved by the institutional review board of the Montefiore Medical Center and the Committee on Clinical Investigation at the Albert Einstein College of Medicine and is in accordance with Health Insurance Portability and Accountability Act regulations. Written informed consent was obtained from all subjects prior to participation. We obtained liver biopsies from the Montefiore-Einstein Liver Center Hepatic Tissue and Serum Repository taken from control patients with non-virally (HCV, HBV and HIV) infected non-cancerous liver tissue and cases with HCV-infected HCC, from which we obtained paired tumour and non-tumour tissues. All cases were sex, race and approximately age (±15 years) matched to controls. After retrospective chart reviews, we reclassified two samples resulting in 11 HCV-uninfected non-neoplastic (controls), 10 HCV+ non-neoplastic (infected) and 9 HCV+ HCC (malignant) samples.

Patient blood samples for DNA methylation studies

We obtained 8–12 ml of whole blood from 10 HCV positive patients and 10 age, race and gender matched HCV negative control patients attending Gastroenterology and Liver Disease clinics at Montefiore Medical Center. We isolated lymphocytes through negative selection using immunomagnetic beads and magnetic columns (Miltenyi Biotec, Bergisch Gladbach, Germany).

Nucleic acid preparation

For each of the 30 biopsies, we divided the tissue into two portions for separate DNA and RNA extraction. For all peripheral lymphocytes only genomic DNA was extracted. Genomic DNA was extracted using proteinase K digestion, phenol–chloroform extraction, dialysis against 0.2X SSC and concentration by surrounding the dialysis bag with PEG 20 000 to reduce water content by osmosis. Total RNA was extracted using TRIzol. The quality and integrity of the RNA were measured using NanoDrop spectrophotometry and Bioanalyzer (Agilent, Santa Clara, CA, USA).

HELP-tagging assay and quantification of genome-wide DNA methylation

As previously described,[127] we used 1 μg of liver tissue DNA and 200 ng of lymphocyte DNA to generate HELP-tagging libraries. Genomic DNA was digested with either MspI or HpaII at 37 °C overnight, purified and ligated indexed Illumina adapters containing the T7 promoter sequence, as well as the EcoP15I recognition site (AE adapters). After ligation, the DNA samples were digested with EcoP15I at 37 °C overnight, end-filled, 3′ terminal A extended and ligated to the second Illumina adapter (AS adapter). Samples were then in vitro transcribed using the MEGAshort kit (Ambion, Thermo Fisher Scientific, Waltham, MA, USA), followed by retro-transcription (SuperScriptIII kit, Invitrogen, Carlsbad, CA, USA) before amplification. Libraries were multiplexed for 50 bp single-end sequencing on the Illumina HiSeq 2000 platform. For each sample, sequencing read counts were obtained at HpaII sites throughout the genome and compared to the previously generated MspI human reference. The read counts were log transformed and normalized by calculating Z scores before taking the ratio of the normalized HpaII count to normalized MspI count at each assayed HpaII site. The ratios are approximately Cauchy distributed, so the cumulative distribution function of the Cauchy distribution was used to obtain

2040

DNA methylation probabilities. Finally, mixture modelling of the distribution assuming three composite distributions was performed to scale the probabilities between 0 and 1. Low confidence CG sites were identified as sites with no reads in the HpaII channel and fewer than five reads in the MspI channel, and they were subsequently removed from the analysis.

Directional whole transcriptome RNA-sequencing assay and gene expression quantification

DNase-treated, rRNA-depleted (Ribozero, Epicentre, Madison, WI, USA) RNA was used as a template for SuperScript III first-strand cDNA synthesis (Invitrogen), using oligo-dT as well as random hexamers. Actinomycin D was added to the reaction to prevent any possible amplification from contaminating genomic DNA. During second-strand synthesis, a dU/VTP mix was used to create directional libraries. cDNA samples were fragmented with Covaris to 300-bp fragments. The samples were then end-filled, 3' terminal A extended and ligated to pre-annealed TruSeq-indexed Illumina adapters. Uracil-DNAglycosylase treatment preceded the PCR reaction to amplify exclusively the originally oriented transcripts. Libraries were amplified using P5 and P7 Illumina primers and gel-extracted for size selection and primer-dimer removal. Before sequencing, libraries were tested using the BioAnalyzer to assure library quality, in terms of size and primer-dimer depletion. Indexed libraries were multiplexed for 100 bp single-end sequencing on the Illumina HiSeq 2000 platform. For each sample, we obtained FASTQ files of past filter reads and aligned them to a composite reference exome of human (hg19), mirBase and HCV using GSNAP. Next, alignments were associated with genes using HTSeq. We scaled gene counts using DESeq implemented in R to estimate effective library size.

Unsupervised K-means clustering to determine summary patterning

For each HpaII locus and each gene, the average DNA methylation and gene expression, respectively, for each tissue type was calculated and used to compute the natural log-ratio of infected and malignant samples relative to the control samples. To identify patterns in DNA methylation and expression relative to tissue type, we used K-means clustering, an unsupervised clustering approach, available through the kmeans function in R. We determined an optimal number of clusters by plotting the total within cluster sums of squares against the number of clusters and selecting a cluster value that occurred at an inflection point.

Statistical method for DNA methylation and gene expression analysis

We used principal components analysis to identify biological and technical covariates that were associated with the variability in DNA methylation and gene expression, separately. Covariates found to be significantly associated with DNA methylation and gene expression were controlled for downstream analyses by inclusion in linear models to obtain covariate-adjusted estimates. For robust data analysis and visualization, we employed batch correction implemented through the ComBat package in R to obtain covariate adjusted DNA methylation values. Infected and malignant samples show more intrasample variability than intersample variability (Figure 1), so they were considered independent cohorts in all analyses. For liver tissue, we performed polytomous regression, a modelling technique that is suited for analysing DNA methylation proportions and log gene counts and allows simultaneous assessment of the odds of DNA methylation or expression for infected and malignant tissues relative to control, in R using the multinom function available through the nnet package. Odds of DNA methylation was adjusted for age and race, while odds of gene expression was adjusted for sequencing

batch. Overall model significance was assessed by adjustment for multiple comparisons using a false discovery rate controlled at alpha = 0.05. Because the overall model significance does not indicate which comparison between control, infected and malignant samples is driving the significance, we additionally use two criteria for further determination of models: (a) control compared with malignant effect significance < 0.05 and (b) an average effect difference between control and malignant samples of 10%. For lymphocyte DNA methylation, we used logistic regression predicting the odds of methylation during chronic HCV infection relative to uninfected samples.

Genomic context analysis

Using the UCSC Genome Browser, we obtained the hg19 RefSeq gene annotation track and the predicted CpG island track. Gene promoters were defined as the 2 kb region flanking the gene transcription start site, while gene bodies were defined as the end of the gene promoter to the transcription end site. All other genome sequence was defined as intergenic. The 2 kb regions flanking CpG Islands were defined as CpG island shores.[42] We obtained publically available ChIP-seq datasets from the ROADMAP Epigenomics project for normal adult liver corresponding to four histone marks: H3K4me3, H3K4me1, H3K36me3 and H3K27me3. For each histone mark, we obtained hg19 aligned reads from three individuals and called peaks using MACS2.[128] We only considered true peaks to be those occurring in at least two of the samples. For transcription factor binding sites, H3K4me3 and H3K27me3 mapping data from the HepG2 HCC cell line, we obtained publically available ChIP-seq datasets from the ENCODE project. The peaks for the HepG2 histone methylation datasets were already called relative to input in a pooled replicate approach, mirroring the approach we used for the liver studies. For relative enrichment studies, the observed overlap between loci of interest and a particular annotation was compared with the expected overlap given the total coverage of the annotation. For permutation tests, we compared the observed overlap of significant HpaII sites and a particular genomic interval (for example gene promoters, CpG islands) with distribution of the same overlap under the null hypothesis. To obtain the null distribution, we employed a random sampling approach where we randomly sampled HpaII sites 100 times from the population of all assayed HpaII sites, effectively permuting the HpaII sites. Thus, with loci of interest, n, and an observed overlap with a particular genomic interval, X, we sampled n loci from all HpaII sites and found the overlap of the random sample with the genomic interval, Y_k, for k in 1, 2, …, 100 (that is 100 random samples). Next, we compared X with the distribution of simulated annotation overlaps, $Y_{1, 2, …, 100}$. If the resulting null distribution, $Y_{1, 2, …, 100}$, contains the observed overlap, X, then we can conclude that there is no significant enrichment. Conversely, when the null distribution, $Y_{1, 2, …, 100}$, excludes the observed overlap, X, then we can conclude that there is significant enrichment beyond that of random chance.

Gene Set Enrichment Analysis

Using the Gene Set Enrichment Analysis tool from the Broad Institute, we were able to cross-reference our identified genes of interest to a database of common pathways, the Molecular Signatures Database v 4.0 (MSigDB). The MSigDB contains curated gene sets assembling online pathway databases, like Kyoto Encyclopedia of Genes and Genomes and REACTOME, publications from PubMed and domain experts. Additionally, the Gene Set Enrichment Analysis tool has catalogued all published genome-wide datasets pertaining to HCC. The overlap between our genes of interest and known gene pathways was analysed using a hypergeometric distribution with a false discovery rate used to correct for multiple hypothesis testing.

Verification

We verified methylation results using Sequenom's MassARRAY (Sequenom, San Diego, CA, USA) in six controls, five infected and four malignant samples at 24 different loci (Supplementary Figure S10). Using the UCSF *MethPrimer* tool, we designed primers to amplify bisulphite converted target DNA sequences. *BiSearch* was used to avoid primer sets that generate off-target amplicons. Following Sequenom's MALDI-TOF Mass Spectrometry, matched peak data was exported using EpiTYPER software and analysed for quality and DNA methylation. We verified gene expression using qRT-PCR of 10 genes in seven controls, five infected and five malignant samples (Supplementary Figure S10). We also show representative loci demonstrating the sites and the distributions of DNA methylation changes in Supplementary Figure S11.

Validation using TCGA data

Using TCGA datasets for HCC, we identified five samples with HCV infection listed as the primary reason for HCC. All five samples had Infinium HumanMethylation450 Methylation microarray and RNA-seq gene expression data available. For DNA methylation and RNA-seq we plotted the density of signal for all genes and our genes of interest.

Polycomb inhibition using the EZH2 inhibitor GSK343

We obtained GSK343, an EZH2 methyltransferase inhibitor, through Santa Cruz Biotechnology, Inc. (Santa Cruz, CA, USA; sc-397025). GSK343 was dissolved in DMSO at a stock concentration of 1.5 mM. HepG2 cells were maintained in Eagle's Minimum Essential Medium supplemented with 10% fetal bovine serum, streptomycin (100 µg/mL), and penicillin (100 units/mL). A WST-1 assay (Clontech, Takara Bio USA Inc., Mountain View, CA, USA) was performed following the manufacturer's recommendations. HepG2 cells for control and GSK343 treatment groups were seeded in triplicate in 96 well plates. After 24, 48, 72 and 96 h of treatment with either 5 µM GSK343 or a 0.1% DMSO vehicle control, absorbance was measured after 2 h of incubation with WST-1 to quantify the net metabolic activity of the cells in culture.

For control and treatment group HepG2 cells, chromatin and nucleic acids were isolated at 72 h. Using the primers in Supplementary Table S3, candidate genes were analysed for gene expression, and gene-associated candidate regions were analysed for the presence of DNA methylation and H3K27me3. Gene expression studies were performed using qRT-PCR with intron-spanning primers designed for *PCSK6, SLC27A2* and *CYP2A7*. We used SYBR Green (Roche 4707516001) with the recommended settings for the LightCycler 480. For DNA methylation studies, 1 µg of DNA was bisulphite treated (Zymo Research, Irvine, CA, USA; D5005) and for each gene, regions overlapping informative HpaII sites were chosen for amplification and analysis with MassArray (Sequenom). For histone studies, native ChIP was performed with immunoprecipitation by antibodies against H3K27me3 (Fisher, Waltham, MA, USA; 07449MI) and rabbit IgG. The relative amount of H3K27me3 was quantified using qRT-PCR for the same regions assayed for DNA methylation in *PCSK6, SLC27A2* and *CYP2A7*. All assays were performed in triplicate.

Data and code sharing

All custom code used is available at https://github.com/GreallyLab/HCV_Wijetunga_et_al_2016.

All data have been uploaded to the Gene Expression Omnibus under Accession Number GSE82178 at http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc = GSE82178.

Private reviewer link: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token = gbqnqyuahzghjaj&acc = GSE82178.

REFERENCES

1 Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015; **136**: E359–E386.

2 Colombo M. Interferon-free therapy for hepatitis C: The hurdles amid a golden era. *Dig Liver Dis* 2015; **47**: 727–733.

3 Schwartz ME, Shrager B. Surgical resection for hepatocellular carcinoma in the noncirrhotic: the Western experience. *Recent Results Cancer Res* 2013; **190**: 85–100.

4 Koike K. The oncogenic role of hepatitis C virus. *Recent Results Cancer Res* 2014; **193**: 97–111.

5 Moriya K, Nakagawa K, Santa T, Shintani Y, Fujie H, Miyoshi H et al. Oxidative stress in the absence of inflammation in a mouse model for hepatitis C virus-associated hepatocarcinogenesis. *Cancer Res* 2001; **61**: 4365–4370.

6 Lerat H, Honda M, Beard MR, Loesch K, Sun J, Yang Y et al. Steatosis and liver cancer in transgenic mice expressing the structural and nonstructural proteins of hepatitis C virus. *Gastroenterology* 2002; **122**: 352–365.

7 Jelic S, Sotiropoulos GCESMO Guidelines Working Group, Hepatocellular carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2010; **21**(Suppl 5): v59–v64.

8 Zogopoulos G, Chen L, Borozan I, Sun J, Heathcote J, Gallinger S et al. Gene expression profiling of hepatocellular cancer identifies a consistent pattern in cancerous and distant liver tissue: arguments for a field defect. *Cancer Res* 2005; **65**: 1434.

9 Li M, Zhao H, Zhang X, Wood LD, Anders RA, Choti MA et al. Inactivating mutations of the chromatin remodeling gene ARID2 in hepatocellular carcinoma. *Nat Genet* 2011; **43**: 828–829.

10 Tao Y, Ruan J, Yeh SH, Lu X, Wang Y, Zhai W et al. Rapid growth of a hepatocellular carcinoma and the driving mutations revealed by cell-population genetic analysis of whole-genome data. *Proc Natl Acad Sci USA* 2011; **108**: 12042–12047.

11 Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad IB et al. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat Genet* 2012; **44**: 694–698.

12 Huang J, Deng Q, Wang Q, Li KY, Dai JH, Li N et al. Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma. *Nat Genet* 2012; **44**: 1117–1121.

13 Machida K, Cheng KT, Sung VM, Lee KJ, Levine AM, Lai MM. Hepatitis C virus infection activates the immunologic (type II) isoform of nitric oxide synthase and thereby enhances DNA damage and mutations of cellular genes. *J Virol* 2004; **78**: 8835–8843.

14 Wurmbach E, Chen YB, Khitrov G, Zhang W, Roayaie S, Schwartz M et al. Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma. *Hepatology* 2007; **45**: 938–947.

15 Schübeler D. Function and information content of DNA methylation. *Nature* 2015; **517**: 321–326.

16 Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 2012; **13**: 484–492.

17 Baylin SB. DNA methylation and gene silencing in cancer. *Nat Clin Pract Oncol* 2005; **2**(Suppl 1): S4–S11.

18 Eads CA, Lord RV, Kurumboor SK, Wickramasinghe K, Skinner ML, Long TI et al. Fields of aberrant CpG island hypermethylation in Barrett's esophagus and associated adenocarcinoma. *Cancer Res* 2000; **60**: 5021–5026.

19 Wentzensen N, Sherman ME, Schiffman M, Wang SS. Utility of methylation markers in cervical cancer early detection: appraisal of the state-of-the-science. *Gynecol Oncol* 2009; **112**: 293–299.

20 Yagi K, Takahashi H, Akagi K, Matsusaka K, Seto Y, Aburatani H et al. Intermediate methylation epigenotype and its correlation to KRAS mutation in conventional colorectal adenoma. *Am J Pathol* 2012; **180**: 616–625.

21 Zhang YJ, Wu HC, Shen J, Ahsan H, Tsai WY, Yang HI et al. Predicting hepatocellular carcinoma by detection of aberrant promoter methylation in serum DNA. *Clin Cancer Res* 2007; **13**: 2378–2384.

22 Nishida N, Nagasaka T, Nishimura T, Ikai I, Boland CR, Goel A. Aberrant methylation of multiple tumor suppressor genes in aging liver, chronic hepatitis, and hepatocellular carcinoma. *Hepatology* 2008; **47**: 908–918.

23 Umer M, Qureshi SA, Hashmi ZY, Raza A, Ahmad J, Rahman M et al. Promoter hypermethylation of Wnt pathway inhibitors in hepatitis C virus-induced multistep hepatocarcinogenesis. *Virol J* 2014; **11**: 117.

24 Archer KJ, Mas VR, Maluf DG, Fisher RA. High-throughput assessment of CpG site methylation for distinguishing between HCV-cirrhosis and HCV-associated hepatocellular carcinoma. *Mol Genet Genomics* 2010; **283**: 341–349.

25 Shen J, Wang S, Zhang YJ, Kappil M, Wu HC, Kibriya MG et al. Genome-wide DNA methylation profiles in hepatocellular carcinoma. *Hepatology* 2012; **55**: 1799–1808.

26 Lao VV, Grady WM. Epigenetics and colorectal cancer. *Nat Rev Gastroenterol Hepatol* 2011; **8**: 686–700.

27 Baylin SB, Herman JG. DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet* 2000; **16**: 168–174.

28 Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 2002; **3**: 415–428.

29 Taberlay PC, Statham AL, Kelly TK, Clark SJ, Jones PA. Reconfiguration of nucleosome-depleted regions at distal regulatory elements accompanies DNA methylation of enhancers and insulators in cancer. *Genome Res* 2014; **24**: 1421–1432.

30 Feldmann A, Ivanek R, Murr R, Gaidatzis D, Burger L, Schübeler D. Transcription factor occupancy can mediate active turnover of DNA methylation at regulatory regions. *PLoS Genet* 2013; **9**: e1003994.

31 Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E et al. The accessible chromatin landscape of the human genome. *Nature* 2012; **489**: 75–82.

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**: 490–495.

33 Domcke S, Bardet AF, Adrian Ginno P, Hartl D, Burger L, Schübeler D. Competition between DNA methylation and transcription factors determines binding of NRF1. *Nature* 2015; **528**: 575–579.

34 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**: 1198–1206.

35 Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND et al. Global epigenomic reconfiguration during mammalian brain development. *Science* 2013; **341**: 1237905.

36 Hodges E, Molaro A, Dos Santos CO, Thekkat P, Song Q, Uren PJ et al. Directional DNA methylation changes and complex intermediate states accompany lineage specificity in the adult hematopoietic compartment. *Mol Cell* 2011; **44**: 17–28.

37 Gebhard C, Benner C, Ehrich M, Schwarzfischer L, Schilling E, Klug M et al. General transcription factor binding at CpG islands in normal cells correlates with resistance to de novo DNA methylation in cancer cells. *Cancer Res* 2010; **70**: 1398–1407.

38 Aran D, Hellman A. DNA methylation of transcriptional enhancers and cancer predisposition. *Cell* 2013; **154**: 11–13.

39 Delahaye F, Wijetunga NA, Heo HJ, Tozour JN, Zhao YM, Greally JM et al. Sexual dimorphism in epigenomic responses of stem cells to extreme fetal growth. *Nat Commun* 2014; **5**: 5187.

40 Ziller MJ, Gu H, Müller F, Donaghey J, Tsai LT, Kohlbacher O et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature* 2013; **500**: 477–481.

41 Won KJ, Zhang X, Wang T, Ding B, Raha D, Snyder M et al. Comparative annotation of functional regions in the human genome using epigenomic data. *Nucleic Acids Res* 2013; **41**: 4423–4432.

42 Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 2009; **41**: 178–186.

43 Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. *J Mol Biol* 1987; **196**: 261–282.

44 Bird AP. CpG-rich islands and the function of DNA methylation. *Nature* 1986; **321**: 209–213.

45 Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP. CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci USA* 1999; **96**: 8681–8686.

46 Costello JF, Frühwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X et al. Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat Genet* 2000; **24**: 132–138.

47 Jones PA, Laird PW. Cancer epigenetics comes of age. *Nat Genet* 1999; **21**: 163–167.

48 Esteller M. Epigenetics in cancer. *N Engl J Med* 2008; **358**: 1148–1159.

49 Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* 2008; **4**: e1000242.

50 Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C et al. Epigenetic stem cell signature in cancer. *Nat Genet* 2007; **39**: 157–158.

51 Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L et al. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nat Genet* 2007; **39**: 237–242.

52 Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J et al. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. *Nat Genet* 2007; **39**: 232–236.

53 Hu CY, Mohtat D, Yu Y, Ko YA, Shenoy N, Bhattacharya S et al. Kidney cancer is characterized by aberrant methylation of tissue-specific enhancers that are prognostic for overall survival. *Clin Cancer Res* 2014; **20**: 4349–4360.

54 Aran D, Sabato S, Hellman A. DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. *Genome Biol* 2013; **14**: R21.

55 Rao X, Evans J, Chae H, Pilrose J, Kim S, Yan P et al. CpG island shore methylation regulates caveolin-1 expression in breast cancer. *Oncogene* 2013; **32**: 4519–4528.

56 Bockmühl Y, Patchev AV, Madejska A, Hoffmann A, Sousa JC, Sousa N et al. Methylation at the CpG island shore region upregulates Nr3c1 promoter activity after early-life stress. *Epigenetics* 2015; **10**: 247–257.

57 Wijetunga NA, Delahaye F, Zhao YM, Golden A, Mar JC, Einstein FH et al. The meta-epigenomic structure of purified human stem cell populations is defined at cis-regulatory sequences. *Nat Commun* 2014; **5**: 5195.

58 Kaposi-Novak P, Lee JS, Gòmez-Quiroz L, Coulouarn C, Factor VM, Thorgeirsson SS. Met-regulated expression signature defines a subset of human hepatocellular carcinomas with poor prognosis and aggressive phenotype. *J Clin Invest* 2006; **116**: 1582–1595.

59 Hoshida Y, Nijman SM, Kobayashi M, Chan JA, Brunet JP, Chiang DY et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res* 2009; **69**: 7385–7392.

60 Villanueva A, Chiang DY, Newell P, Peix J, Thung S, Alsinet C et al. Pivotal role of mTOR signaling in hepatocellular carcinoma. *Gastroenterology* 2008; **135**: 1972–1983; 1983.e1.

61 Lee JS, Chu IS, Heo J, Calvisi DF, Sun Z, Roskams T et al. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. *Hepatology* 2004; **40**: 667–676.

62 Chen X, Cheung ST, So S, Fan ST, Barry C, Higgins J et al. Gene expression patterns in human liver cancers. *Mol Biol Cell* 2002; **13**: 1929–1939.

63 Boyault S, Rickman DS, de Reyniès A, Balabaud C, Rebouissou S, Jeannot E et al. Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. *Hepatology* 2007; **45**: 42–52.

64 Chiang DY, Villanueva A, Hoshida Y, Peix J, Newell P, Minguez B et al. Focal gains of VEGFA and molecular classification of hepatocellular carcinoma. *Cancer Res* 2008; **68**: 6779–6788.

65 Yamashita T, Forgues M, Wang W, Kim JW, Ye Q, Jia H et al. EpCAM and alphafetoprotein expression defines novel prognostic subtypes of hepatocellular carcinoma. *Cancer Res* 2008; **68**: 1451–1461.

66 Cairo S, Armengol C, De Reyniès A, Wei Y, Thomas E, Renard CA et al. Hepatic stem-like phenotype and interplay of Wnt/beta-catenin and Myc signaling in aggressive childhood liver cancer. *Cancer Cell* 2008; **14**: 471–484.

67 Lee JS, Heo J, Libbrecht L, Chu IS, Kaposi-Novak P, Calvisi DF et al. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. *Nat Med* 2006; **12**: 410–416.

68 Mínguez B, Hoshida Y, Villanueva A, Toffanin S, Cabellos L, Thung S et al. Gene-expression signature of vascular invasion in hepatocellular carcinoma. *J Hepatol* 2011; **55**: 1325–1331.

69 Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med* 2008; **359**: 1995–2004.

70 Budhu A, Forgues M, Ye QH, Jia HL, He P, Zanetti KA et al. Prediction of venous metastases, recurrence, and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment. *Cancer Cell* 2006; **10**: 99–111.

71 Kim JW, Ye Q, Forgues M, Chen Y, Budhu A, Sime J et al. Cancer-associated molecular signature in the tissue samples of patients with cirrhosis. *Hepatology* 2004; **39**: 518–527.

72 Llovet JM, Chen Y, Wurmbach E, Roayaie S, Fiel MI, Schwartz M et al. A molecular signature to discriminate dysplastic nodules from early hepatocellular carcinoma in HCV cirrhosis. *Gastroenterology* 2006; **131**: 1758–1767.

73 Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013; 45: 1113–1120.

74 ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. Nature 2012; 489: 57–74.

75 Alder O, Cullum R, Lee S, Kan AC, Wei W, Yi Y et al. Hippo signaling influences HNF4A and FOXA2 enhancer switching during hepatocyte differentiation. Cell Rep 2014; 9: 261–271.

76 DeLaForest A, Nagaoka M, Si-Tayeb K, Noto FK, Konopka G, Battle MA et al. HNF4A is essential for specification of hepatic progenitors from human pluripotent stem cells. Development 2011; 138: 4143–4153.

77 Lee CS, Friedman JR, Fulmer JT, Kaestner KH. The initiation of liver development is dependent on Foxa transcription factors. Nature 2005; 435: 944–947.

78 Nagy P, Bisgaard HC, Thorgeirsson SS. Expression of hepatic transcription factors during liver development and oval cell differentiation. J Cell Biol 1994; 126: 223–233.

79 Yamazaki H, Katsuoka F, Motohashi H, Engel JD, Yamamoto M. Embryonic lethality and fetal liver apoptosis in mice lacking all three small Maf proteins. Mol Cell Biol 2012; 32: 808–816.

80 Sucov HM, Dyson E, Gumeringer CL, Price J, Chien KR, Evans RM. RXR alpha mutant mice establish a genetic basis for vitamin A signaling in heart morphogenesis. Genes Dev 1994; 8: 1007–1018.

81 Katsuoka F, Motohashi H, Ishii T, Aburatani H, Engel JD, Yamamoto M. Genetic evidence that small maf proteins are essential for the activation of antioxidant response element-dependent genes. Mol Cell Biol 2005; 25: 8044–8051.

82 Meir O, Dvash E, Werman A, Rubinstein M. C/EBP-beta regulates endoplasmic reticulum stress-triggered cell death in mouse and human models. PLoS ONE 2010; 5: e9516.

83 Alam T, An MR, Papaconstantinou J. Differential expression of three C/EBP isoforms in multiple tissues during the acute phase response. J Biol Chem 1992; 267: 5021–5024.

84 Zimmerman TL, Thevananther S, Ghose R, Burns AR, Karpen SJ. Nuclear export of retinoid X receptor alpha in response to interleukin-1beta-mediated cell signaling: roles for JNK and SER260. J Biol Chem 2006; 281: 15434–15440.

85 Arensdorf AM, Rutkowski DT. Endoplasmic reticulum stress impairs IL-4/IL-13 signaling through C/EBPβ-mediated transcriptional suppression. J Cell Sci 2013; 126: 4026–4036.

86 Tanigawa S, Lee CH, Lin CS, Ku CC, Hasegawa H, Qin S et al. Jun dimerization protein 2 is a critical component of the Nrf2/MafK complex regulating the response to ROS homeostasis. Cell Death Dis 2013; 4: e921.

87 Palomer X, Salvadó L, Barroso E, Vázquez-Carrera M. An overview of the crosstalk between inflammatory processes and metabolic dysregulation during diabetic cardiomyopathy. Int J Cardiol 2013; 168: 3160–3172.

88 Colletti M, Cicchini C, Conigliaro A, Santangelo L, Alonzi T, Pasquini E et al. Convergence of Wnt signaling on the HNF4alpha-driven transcription in controlling liver zonation. Gastroenterology 2009; 137: 660–672.

89 Villacorte M, Suzuki K, Hirasawa A, Ohkawa Y, Suyama M, Maruyama T et al. β-Catenin signaling regulates Foxa2 expression during endometrial hyperplasia formation. Oncogene 2013; 32: 3477–3482.

90 Song Y, Washington MK, Crawford HC. Loss of FOXA1/2 is essential for the epithelial-to-mesenchymal transition in pancreatic cancer. Cancer Res 2010; 70: 2115–2125.

91 Yang M, Li SN, Anjum KM, Gui LX, Zhu SS, Liu J et al. A double-negative feedback loop between Wnt-β-catenin signaling and HNF4α regulates epithelial-mesenchymal transition in hepatocellular carcinoma. J Cell Sci 2013; 126: 5692–5703.

92 Dammacco F, Sansonno D, Piccoli C, Racanelli V, D'Amore FP, Lauletta G. The lymphoid system in hepatitis C virus infection: autoimmunity, mixed cryoglobulinemia, and overt B-cell malignancy. Semin Liver Dis 2000; 20: 143–157.

93 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005; 102: 15545–15550.

94 Bochkis IM, Rubins NE, White P, Furth EE, Friedman JR, Kaestner KH. Hepatocyte-specific ablation of Foxa2 alters bile acid homeostasis and results in endoplasmic reticulum stress. Nat Med 2008; 14: 828–836.

95 Ohguchi H, Tanaka T, Uchida A, Magoori K, Kudo H, Kim I et al. Hepatocyte nuclear factor 4alpha contributes to thyroid hormone homeostasis by cooperatively regulating the type 1 iodothyronine deiodinase gene with GATA4 and Kruppel-like transcription factor 9. Mol Cell Biol 2008; 28: 3917–3931.

96 Verma SK, Tian X, LaFrance LV, Duquenne C, Suarez DP, Newlander KA et al. Identification of potent, selective, cell-active inhibitors of the histone lysine methyltransferase EZH2. ACS Med Chem Lett 2012; 3: 1091–1096.

97 Zhao H, Xu Y, Mao Y, Zhang Y. Effects of EZH2 gene on the growth and migration of hepatocellular carcinoma HepG2 cells. Hepatobiliary Surg Nutr 2013; 2: 78–83.

98 Thompson RF, Suzuki M, Lau KW, Greally JM. A pipeline for the quantitative analysis of CG dinucleotide methylation using mass spectrometry. Bioinformatics 2009; 25: 2164–2170.

99 Jones PA, Baylin SB. The epigenomics of cancer. Cell 2007; 128: 683–692.

100 Lu CY, Hsieh SY, Lu YJ, Wu CS, Chen LC, Lo SJ et al. Aberrant DNA methylation profile and frequent methylation of KLK10 and OXGR1 genes in hepatocellular carcinoma. Genes Chromosomes Cancer 2009; 48: 1057–1068.

101 Couvert P, Carrié A, Pariès J, Vaysse J, Miroglio A, Kerjean A et al. Liver insulin-like growth factor 2 methylation in hepatitis C virus cirrhosis and further occurrence of hepatocellular carcinoma. World J Gastroenterol 2008; 14: 5419–5427.

102 Zhao W, Liu H, Liu W, Wu Y, Chen W, Jiang B et al. Abnormal activation of the synuclein-gamma gene in hepatocellular carcinomas by epigenetic alteration. Int J Oncol 2006; 28: 1081–1088.

103 Heuck CJ, Mehta J, Bhagat T, Gundabolu K, Yu Y, Khan S et al. Myeloma is characterized by stage-specific alterations in DNA methylation that occur early during myelomagenesis. J Immunol 2013; 190: 2966–2975.

104 Alvarez H, Opalinska J, Zhou L, Sohal D, Fazzari MJ, Yu Y et al. Widespread hypomethylation occurs early and synergizes with gene amplification during esophageal carcinogenesis. PLoS Genet 2011; 7: e1001356.

105 Katsurano M, Niwa T, Yasui Y, Shigematsu Y, Yamashita S, Takeshima H et al. Early-stage formation of an epigenetic field defect in a mouse colitis model, and non-essential roles of T- and B-cells in DNA methylation induction. Oncogene 2012; 31: 342–351.

106 Hernandez-Vargas H, Lambert MP, Le Calvez-Kelm F, Gouysse G, McKay-Chopin S, Tavtigian SV et al. Hepatocellular carcinoma displays distinct DNA methylation signatures with potential as clinical predictors. PLoS ONE 2010; 5: e9749.

107 Stefanska B, Huang J, Bhattacharyya B, Suderman M, Hallett M, Han ZG et al. Definition of the landscape of promoter DNA hypomethylation in liver cancer. Cancer Res 2011; 71: 5891–5903.

108 Shen J, Wang S, Zhang YJ, Wu HC, Kibriya MG, Jasmine F et al. Exploring genome-wide DNA methylation profiles altered in hepatocellular carcinoma using Infinium HumanMethylation 450 BeadChips. Epigenetics 2013; 8: 34–43.

109 Song MA, Tiirikainen M, Kwee S, Okimoto G, Yu H, Wong LL. Elucidating the landscape of aberrant DNA methylation in hepatocellular carcinoma. PLoS ONE 2013; 8: e55761.

110 Hogart A, Lichtenberg J, Ajay SS, Anderson SNIH Intramural Sequencing Center, Margulies EH et al. Genome-wide DNA methylation profiles in hematopoietic stem and progenitor cells reveal overrepresentation of ETS transcription factor binding sites. Genome Res 2012; 22: 1407–1418.

111 Choy MK, Movassagh M, Goh HG, Bennett MR, Down TA, Foo RS. Genome-wide conserved consensus transcription factor binding motifs are hyper-methylated. BMC Genomics 2010; 11: 519.

112 Brinkman AB, Gu H, Bartels SJ, Zhang Y, Matarese F, Simmer F et al. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome Res 2012; 22: 1128–1138.

113 Statham AL, Robinson MD, Song JZ, Coolen MW, Stirzaker C, Clark SJ. Bisulfite sequencing of chromatin immunoprecipitated DNA (BisChIP-seq) directly informs methylation status of histone-modified DNA. Genome Res 2012; 22: 1120–1127.

114 Viré E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C et al. The Polycomb group protein EZH2 directly controls DNA methylation. Nature 2006; 439: 871–874.

115 Rush M, Appanah R, Lee S, Lam LL, Goyal P, Lorincz MC. Targeting of EZH2 to a defined genomic site is sufficient for recruitment of Dnmt3a but not de novo DNA methylation. Epigenetics 2009; 4: 404–414.

116 Ptashne M. On the usse of the word "epigenetic". Curr Biol 2007; 17: R233–R236.

117 Ptashne M. Faddish stuff: epigenetics and the inheritance of acquired characteristics. FASEB J 2013; 27: 1–2.

118 Ghosh AK, Majumder M, Steele R, Yaciuk P, Chrivia J, Ray R et al. Hepatitis C virus NS5A protein modulates transcription through a novel cellular transcription factor SRCAP. J Biol Chem 2000; 275: 7184–7188.

119 Qadri I, Iwahashi M, Simon F. Hepatitis C virus NS5A protein binds TBP and p53, inhibiting their DNA binding and p53 interactions with TBP and ERCC3. Biochim Biophys Acta 2002; 1592: 193–204.

120 Banerjee A, Meyer K, Mazumdar B, Ray RB, Ray R. Hepatitis C virus differentially modulates activation of forkhead transcription factors and insulin-induced metabolic gene expression. J Virol 2010; 84: 5936–5946.

121 Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. Nature 2011; 470: 279–283.

122 Zentner GE, Tesar PJ, Scacheri PC. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res* 2011; **21**: 1273–1283.

123 Signolet J, Hendrich B. The function of chromatin modifiers in lineage commitment and cell fate specification. *FEBS J* 2015; **282**: 1692–1702.

124 Machida K, Cheng KT, Sung VM, Shimodaira S, Lindsay KL, Levine AM *et al*. Hepatitis C virus induces a mutator phenotype: enhanced mutations of immunoglobulin and protooncogenes. *Proc Natl Acad Sci USA* 2004; **101**: 4262–4267.

125 Machida K, McNamara G, Cheng KT, Huang J, Wang CH, Comai L *et al*. Hepatitis C virus inhibits DNA damage repair through reactive oxygen and nitrogen species and by interfering with the ATM-NBS1/Mre11/Rad50 DNA repair pathway in monocytes and hepatocytes. *J Immunol* 2010; **185**: 6985–6998.

126 Feo F, Pascale RM. Multifocal hepatocellular carcinoma: intrahepatic metastasis or multicentric carcinogenesis? *Ann Transl Med* 2015; **3**: 4.

127 Suzuki M, Jing Q, Lia D, Pascual M, McLellan A, Greally JM. Optimized design and data analysis of tag-based cytosine methylation assays. *Genome Biol* 2010; **11**: R36.

128 Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE *et al*. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008; **9**: R137.

Supplementary Information accompanies this paper on the Oncogene website (http://www.nature.com/onc)