Published online 17 April 2013 *Nucleic Acids Research, 2013, Vol. 41, No. 11* ***5639–5654***
doi:10.1093/nar/gkt232

Dynamic changes in 5-hydroxymethylation signatures underpin early and late events in drug exposed liver

John P. Thomson[1,2], Jennifer M. Hunter[1], Harri Lempiäinen[2,3], Arne Müller[2,3], Rémi Terranova[2,3], Jonathan G. Moggs[2,3,*] and Richard R. Meehan[1,2,*]

[1]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Crewe Road, Edinburgh EH4 2XU, UK, [2]Member of MARCAR Consortium and [3]Discovery and Investigative Safety, Preclinical Safety, Novartis Institutes for Biomedical Research, Klybeckstrasse, Basel, Switzerland, CH-4002

Received January 28, 2013; Revised March 12, 2013; Accepted March 13, 2013

ABSTRACT

Aberrant DNA methylation is a common feature of neoplastic lesions, and early detection of such changes may provide powerful mechanistic insights and biomarkers for carcinogenesis. Here, we investigate dynamic changes in the mouse liver DNA methylome associated with short (1 day) and prolonged (7, 28 and 91 days) exposure to the rodent liver non-genotoxic carcinogen, phenobarbital (PB). We find that the distribution of 5mC/5hmC is highly consistent between untreated individuals of a similar age; yet, changes during liver maturation in a transcriptionally dependent manner. Following drug treatment, we identify and validate a series of differentially methylated or hydroxymethylated regions: exposure results in staged transcriptional responses with distinct kinetic profiles that strongly correlate with promoter proximal region 5hmC levels. Furthermore, reciprocal changes for both 5mC and 5hmC in response to PB suggest that active demethylation may be taking place at each set of these loci via a 5hmC intermediate. Finally, we identify potential early biomarkers for non-genotoxic carcinogenesis, including several genes aberrantly expressed in liver cancer. Our work suggests that 5hmC profiling can be used as an indicator of cell states during organ maturation and drug-induced responses and provides novel epigenetic signatures for non-genotoxic carcinogen exposure.

INTRODUCTION

In the mammalian genome, the dinucleotide sequence CpG is frequently modified by the addition of a methyl group to the fifth position of cytosine base to form 5-methylcytosine (5mC). Sequences modified in this way are often associated with the silencing of transposable elements and gene promoters as well as playing a role in X-inactivation, tissue specific gene regulation and the regulation of imprinted alleles (1–3). In recent years, there has been intense interest in a second modified form of cytosine, that of 5-hydroxymethylcytosine (5hmC) found in both cultured cells, tissues samples and cancer (4–17). A consensus view is that 5hmC modified CpGs are typically found enriched in the bodies of actively transcribing genes, are present at enhancer elements and at a small cohort of regions spanning an annotated Transcription Start Site (TSS) (8,17,18). The Ten-eleven translocation 1-3 (Tet1-3) proteins (TET) family of Fe(II) and α-ketoglutarate-dependent dioxygenases use molecular oxygen to transfer a hydroxyl group to methylated cytosine bases to form 5hmC (15,19–22). Furthermore, it was shown that Tet proteins can oxidize 5mC or 5hmC further, converting them to 5-formylcytosine (5 fC) and/or 5-carboxylcytosine (5 caC), which are proposed to be intermediates in a demethylation pathway, which may be removed through base excision repair mechanisms by enzymes such as thymine-DNA glycosylase (23–26). The gradual disappearance of 5hmC, 5 fC and 5 caC in pre-implantation embryos indicates that DNA methylation may be removed in part by a replication-dependent passive loss mechanism (26,27). Erasure of CpG methylation in PGCs might also occur as a consequence of passive demethylation during migration and through conversion to 5hmC as a consequence of TET1 and TET2 activity (28,29). Disruption of the TET proteins has been reported to result in much reduced 5hmC levels, a phenomenon also seen during carcinogenesis (30). For example, knockdown of TET1 in ES cells leads to an increase in 5mC at TSS regions of its target genes and a partial decrease in 5hmC at specific promoters and within

*To whom correspondence should be addressed. Tel: +44 1313 322471; Fax: +44 1314 678456; Email: Richard.Meehan@igmm.ed.ac.uk
Correspondence may also be addressed to Jonathan G. Moggs. Tel: +41 6132 47494; Email: jonathan.moggs@novartis.com

gene bodies of TET1 target genes (5,7,9), whereas knockdown of Tet2 in hematopoietic progenitor cells results in a block of myeloid differentiation possibly through a failure to activate critical genes in the differentiation pathway (31,32). Conversely, activation of Tet2 target genes in pre-B cells was also seen to accompany increased promoter hydroxymethylation (33).

The widely used anticonvulsant phenobarbital (PB), is a well characterized non-genotoxic carcinogen (NGC) used to investigate the initiation and progression of non-genotoxic carcinogenesis in the rodent liver, with prolonged exposure (28 days) resulting in the mis-regulation of gene expression of a cohort of genes as well as the perturbation of both DNA methylation and histone modification patterns that ultimately result in tumour formation (34–38). Typically, PB is believed to exhibit its effects through the regulation of nuclear receptors, including the constitutive androstane receptor (*CAR*), which mediates the transcriptional regulation of enzymes involved in response to drug exposure (39). Both DNA and histone modification profiles are perturbed over a handful of strongly PB induced genes in mice receiving PB in their drinking water every day for 28 days (17). Strikingly, there appeared to be a reciprocal gain in 5hmC modified DNA over the promoter regions, which had lost 5mC, resembling an active demethylation event via a 5hmC intermediate.

Here, we expand on this early work by further investigating the hydroxymethylomes of both control and PB-treated mice throughout a time course of PB exposure. In addition, we characterize the individual variations of 5hmC and 5mC patterns between several individual mouse livers of a similar age as well as throughout normal liver development from young mice (aged 30–33 days) through to mature adults 13 weeks later. Our study highlights the potential use of 5hmC profiles as an identifier of liver developmental stage. Analysis of the PB-induced epigenetic perturbations present over the promoter regions of strongly induced genes following both short (1 day) and long-term (7, 28 and 91 days PB) exposure hints at active and ongoing demethylation via a 5hmC intermediate. Combined epigenomic and transcriptomic approaches identify a set of potential biomarkers for the progression of non-genotoxic carcinogenesis, including several genes involved in cell cycle regulation, DNA damage response as well as a few with previously reported perturbation in many cancers including hepatocellular carcinomas (HCC) (40–42). Overall, this work provides novel insights into the 5-hydroxymethylcytosine patterns of normal mouse livers and the dynamic changes resulting from temporal (1–91 days) exposure to a tumor-promotor such as PB.

MATERIALS AND METHODS

Ethics statement

This study was performed in conformity with the Swiss Animal Welfare Law and specifically under the Animal Licenses No. 2345 by 'Kantonales Veterinäramt Basel-Stadt' (Cantonal Veterinary Office, Basel) and No. 5041 by 'Kantonales Veterinäramt Baselland' (Cantonal Veterinary Office, Basel Land).

Animal treatment and sample preparation

In all, 29–32 days old male B6C3F1/Crl (C57BL/6 ♂ × C3H/He ♀) mice were obtained from Charles River Laboratories (Germany). Animals were allowed to acclimatise for 5 days before being randomly divided into two treatment groups of five animals each. PB [Sigma 04710, 0.05% (w/v) in drinking water] was administered to one group through *ad libitum* access to drinking water for either 1, 7, 28 or 91 days. Mice were checked daily for activity and behaviour and sacrificed on the last day of dosing for each required time point of interest. Livers were removed before freezing in liquid nitrogen and −80°C storage.

Purification of 5hmC and 5mC enriched DNA fragments

Genomic DNA was extracted from frozen (−80°C) ground-up livers and fragmented to range between 300 and 1000 bp in size (Bioruptor, Diagenode) before immunoprecipitation. For full HmeDIP and MeDIP protocols, see Thomson *et al*. (17) and Lempiainen *et al*. (34). For Chemical capture 5hmC enrichment (Active Motif Hydroxymethyl collector Kit), please refer to manufacturers' protocol. Following purification, samples were prepared for microarray analysis by whole genome amplification using WGA2:GenomePlex Complete Whole Genome Kit (Sigma). For hmeDIP and meDIP arrays, amplified material was sent to Roche Nimblegen (Iceland) for Cy3 and Cy5 labelling and hybridization on 2.1 M Deluxe mouse promoter tiling arrays. Chemical capture enriched DNA was hybridized in house on 2.1 M whole genome mouse tiling array 2 of 4 (Nimblegen), which contains a proportion of chromosomes 4 and 9 and the entirety of chromosomes 5, 6, 7 and 8.

RNA extraction for expression array analysis

RNA was extracted from liver samples and Affymetrix expression arrays performed following the methods outlined in earlier work (34).

Bioinformatic techniques

Processing of Nimblegen promoter microarrays

Nimblegen 2.1 M deluxe mouse promoter arrays (mm9 build) contain 2 056 330 unique probes of 50–70 bp in length with ~50 bp spacing distributed over 21 562 tiled regions spanning 52 016 annotated TSS regions over 20 718 unique genes. Signals for each probe of the 5hmC enriched samples (Cy5 labelled) were compared with input samples (Cy3 labelled) to generate log2 (IP/Input) scores (fold changes). These log2 scores were then normalized first by Loess normalization and then by scale normalization using the Limma package in R/Bioconductor (43). For all samples, each probe was then mapped to one of six regions of the genome based on Refseq gene annotations (see Supplementary Figure S7a). Owing to the

highly reproducible nature of the patterns, average probe values were calculated for both the control and PB exposed groups. Subsequent analysis was carried out by comparing the differences between the mean control and mean PB data sets at each time point (1, 7, 28 or 91 days) and were plotted as changes in the log2 probe values in PB mice relative to control mice.

Defining regions of differential hydroxymethylation and methylation

A Student's *t*-test was carried out on each probe in the first group with respect to the second and the *t*-statistic calculated (i.e. *t*-statistic calculated between five control mice and five PB-exposed mice). Using the *t*-statistic as a value of significance, regions where at least three probes in a four probe window contained *t*-statistic values in the top 5% of all values generated were deemed significantly differentially modified [termed differentially methylated or hydroxymethylated regions (dHMRs and dMRs)]. In addition, all hyper peak probes are found to contain *t*-statistic scores >3, whereas hypo probe scores are all below −3. By contrast, the mean value of the *t*-statistics in the entire data sets is found to be heavily centred on zero. These peak probes were then mapped to genomic coordinates based on the criteria set in Supplementary Figure S7a).

Affymetrix gene expression analysis

Affymetrix gene expression data sets were normalized with RMA (44). The *P*-value for the log2 fold change was calculated with the R/Bioconductor LIMMA package using a moderated *t*-statistic. The expression changes following PB exposure were represented as log2 fold change expression relative to the control mice.

Sliding window analysis of DNA modifications

Sliding windows of hmC and 5mC profiles were characterized over strongly induced genes (>log2 1.5-fold) at each time point of PB dosing. Sliding window analysis was carried out using tools on the University of Edinburgh Wellcome Trust Centre for Cell biology Galaxy server. Sliding window analysis plots the average signal taken from a data files of interest (e.g. mean 5hmC normal probe values) and slides across regions of interest (chromosome, start, stop) in user-defined steps (in bp). Sliding window analysis was carried out using a window size of 200 bp and with a step size of 50 bp and average signals plotted.

Heat map analysis of PB driven 5hmC perturbation

Average changes to promoter proximal region (PPR) 5hmC or 5mC levels following PB exposure were calculated for each of five drug-treated mice versus the average data set of control mice. Differences in each mark were calculated by subtracting the log2 score for each probe in the PB mouse versus the average of the control mice ($n = 5$). The top 100 PPRs were selected based on the total delta from control across all five mice. Heat maps were drawn using R with colours taken using the Colour Brewer package ('RdYlBu') ranging from values of −1 to +1 with a 0.4 interval. Hundred random PPRs were generated and plotted alongside the PPRs gaining and losing each mark.

Dendrogram clustering of 5hmC and 5mC patterns

Clustering of samples based off of their DNA modification samples was calculated from 500 000 random probes (24.3% total unique probes on array). Dendrogram plots were carried out using R, and distances were calculated through both Euclidian and Ward methods.

Uniprot tissue annotation database ('UP-Tissue') expression analysis

Fisher Exact *P*-values were calculated for the sets of genes containing either a PPR dHMR or dMR using the Uniprot tissue annotation database ('UP-Tissue') analysis tool on the DAVID bioinformatic database website (http://david.abcc.ncifcrf.gov), which plots the tissues in which sets of genes are typically found to be expressed (low *P*-value = expression is highly correlated to the tissue).

Analysis of enhancer elements

Probes on our array, which correspond to regions defined as liver enhancers by Shen *et al.* (45), were taken and crossed to our 5hmC and 5mC data sets. Nearest genes to an enhancer element were calculated by using the 'operate on genomic intervals' function on the University of Edinburgh Wellcome Trust Centre for Cell biology Galaxy server relative to mouse genome build mm9 gene coordinates.

q-PCR validation of array results

Standard quantitative or real time PCR (qPCR) analysis was carried out on a Roche Lightcycler480 qPCR machine (Roche). A list of primers is available in Supplementary Table S6.

Access to data sets

All 5hmC, 5mC and expression data sets are available on GEO under the super-series accession number Series GSE40540.

RESULTS

Global liver 5-hydroxymethylome patterns are highly conserved between individuals

We first set out to fully characterize the 5mC and 5hmC patterns in normal control liver over the same time course of PB dosing (13 weeks total). During this period, the livers of young mice (30–33 days old) continue to mature allowing us to also study any associated epigenetic changes. To study the 5hmC and 5mC patterns in the liver, we enriched for DNA fragments containing either mark using antibodies specific for either 5hmC (Hydroxymethyl-DNA immunoprecipitation; HmeDIP) or 5mC (Methyl-DNA immunoprecipitation; MeDIP) from a group of mice of aged 30–33 days old ($n = 5$, see 'Materials and Methods' section). The enriched material was then mapped back to the genome through hybridization to high density promoter microarrays (Nimblegen 2.1 Million probe Deluxe promoter array, Figure 1a), which allow us to investigate the

promoter regions of a large number of mice. The advantages of the promoter array strategy that we used is that the high density (2.1 million probes) enabled us to visualize changes to both 5mC and 5hmC marks following PB exposure at promoter-associated regions in addition to regions 8 kb upstream and 3 kb downstream of annotated transcriptional start sites. An important point is that we were in a position to compare multiple samples (five mice per time point; both control and drug exposed) to investigate individual variation between liver 5-hydroxymethylomes. The availability of multiple sample analysis with this array method also provides for a significant statistical power to this study; in total, there are 120 data sets (80 DNA modifications and 40 expression arrays) in this study.

We initially set out to investigate if the liver 5mC and 5hmC patterns differ between individual mice of a similar age. Overall, Pearson correlations of the individual $\log_2$ probe intensity values on the array between each liver sample (ages 30–33 days) were high for the 5hmC patterns between all individual livers (average correlation value of 0.88 ± 0.01; (Figure 1b and Supplementary Table S1). In all, 5mC levels also exhibit significant Pearson correlation values for all probes between individuals (average correlation value 0.69 ± 0.064), albeit the patterns are less well conserved than 5hmC. This is likely to be in part due to cited issues with CpG density dependant binding of the 5mC antibody (46–49).

Visualization of these data sets further highlights the highly conserved nature of the hydroxymethylome and methylome between individual livers (Figure 1c and Supplementary Figure S1); however, to test this further, we set about defining regions of significant epigenetic change, which occur between the five mice aged 30–33 days old (dHMRs and dMRs). In brief, this was based on scoring each probe based on its departure from its notional value as well as its standard error (the *t*-statistic), followed by grouping significantly changing adjacent probes into dHMRs and dMRs (see 'Materials and Methods' section). Globally, there were very few regions, which exhibited a strong change in either hydroxymethylation (117 hyper-dHMRs, 51 hypo-dHMRs) or methylation (92 hyper-dMRs, 84 hypo-dMRs) across the control group of animals aged 30–33 days (Supplementary Table S2). Additionally, only a handful of the differentially methylated and/or hydroxymethylated regions were found to occur within 1 kb of an annotated gene transcriptional start site (defined as PPRs) in at least one of the mice. Indeed, only seven such PPRs were found to contain a dMR in at least one of the mice (four hyper, three hypo), and just one was found to contain an associated dHMR (Supplementary Table S2). Together, these results highlight global and promoter specific conservation of hydroxymethylation patterns, and to a lesser extent, methylation patterns in the livers of mice at a similar juvenile stage.

5hmC patterns during late stage liver maturation

We then turned our attention to investigating the changes to the hydroxymethylome and methylome during a 13-week period of late stage liver maturation. Liver mass is known to increase 3-fold between the day of birth (19 days postcoitus) to postnatal day 9, rapid cellular proliferation can be observed up to day 20, but by postnatal day 30, mouse livers are histologically identical to livers of mice at 3 months of age (50,51). At this point, the developmental patterns of gene expression of many liver specific genes are stably established (52); however, it is believed that liver maturation still occurs over the following weeks. To test this further and to study the epigenetic events associated with late stage liver maturation, we studied the changes to liver 5hmC and 5mC patterns in young mice aged 30–33 days through to those aged 13 weeks (91 days) later. Interestingly, global 5hmC patterns are sufficient to stratify the young livers from those 91 days later in their maturation (Figure 1d), which is in agreement with previous studies highlighting 5hmC as an identifier of tissue and cell state (16,17). In contrast to 5hmC, global 5mC patterns alone were unable to fully separate the juvenile and mature livers from each other (Supplementary Figure S2).

Using the same approach described earlier in the text, we find a large number of differential hydroxymethylation (dHMRs) and methylation (dMRs) events occur during 91 days of liver maturation. Globally, we find many more changes occur in the hydroxymethylome (15052; 6775 hypo- and 8277 hyper-) compared with the methylome (2484; 716 hypo- and 1768 hyper-), indicating that 5hmC is more dynamic through maturation than its more studied 5mC counterpart (Figure 1e). This is true both for the number of differential hydroxymethylated and methylated regions genome wide, over annotated enhancer elements, as well as a number of genes, which gain or lose either mark over their PPRs (Figure 1e and Supplementary Table S3). It has been reported that 5hmC can mark enhancer elements in both human and mouse cell lines (8,12) as well as in the mouse liver (17). Although 5hmC enrichment at enhancers has been proposed to act as an early event of enhancer activation based on the analysis of select histone modifications (8,12), the relationship between gene expression changes at nearby genes has not been tested. Here, we investigated the levels of both 5hmC and 5mC over liver enhancer elements [reported by (45)] and relate these to the expression levels of nearby genes during 13 week liver maturation. From this analysis, we find that enhancers remain enriched for 5hmC and depleted for 5mC in a manner, which is independent of the transcriptional activity of the nearest gene (Supplementary Figure S3). Further, to this, we find that although a number of enhancer elements contain a dHMR ($n = 145$) or dMR ($n = 75$) during the 13 weeks of liver maturation (Figure 1e), there is no correlation to the transcriptional activity of nearby genes (Supplementary Figure S4).Taken together, these results argue that the epigenetic landscape over enhancer elements (enrichment of 5hmC, depletion of 5mC) is independent of the transcriptional activity of the nearby genes over the period studied.

As a number of genes were found to contain a dHMR and/or dMR between the young and mature (+91 day) livers, we set out to investigate the relationships between the expression of these genes and their PPR

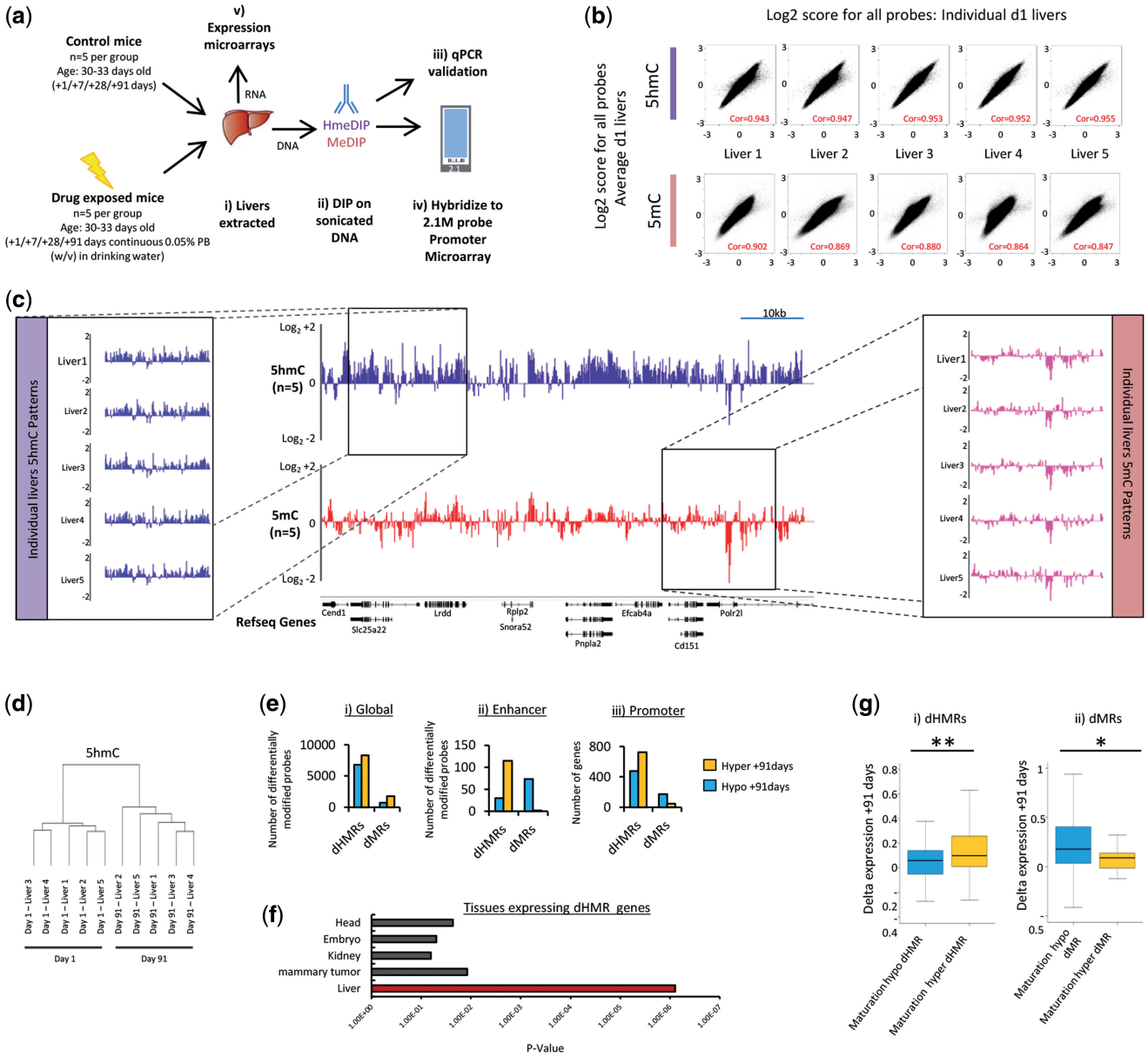

Figure 1. Global 5hmC and 5mC patterns are conserved between individuals of a similar age but changes throughout liver maturation. (**a**) Overview of experimental strategy for the identification OF PB-induced epigenetic and transcriptomic perturbations in liver tissue from control and related PB-treated B6C3F1 male mice (+1, +7, +28 and +91 days dosing). In all, 5hmC and 5mC profiles were analysed following antibody based purification techniques (hydroxymethyl-DNA immunoprecipitation: hmeDIP and methyl-DNA immunoprecipitation: meDIP) by promoter microarrays and verified by qPCR. The transcriptomes of control and PB-treated livers was investigated by parallel expression microarrays. (**b**) Global liver 5hmC and 5mC patterns are similar between animals of the same age (30–33 days old; 'day 1 mice'). Scatter plots of $\log_2$ values for all probes on the array comparing correlations between five individual livers and the average of all five. Correlation values are displayed at the bottom right of each plot in red. (**c**) Visualization of 5hmC (purple) and 5mC (red) patterns across the indicated regions in individual livers from mice aged 30–33 days old ('day 1 mice'). Boxed regions are expanded to reveal the reproducible nature of the DNA modification profiling. Data are represented by normalized $\log_2$ scores for all probes across this region. Annotated Refseq genes are represented below as black bars including gene structures. (**d**) In all, 5hmC patterns are sufficient to cluster juvenile mouse livers from mature mouse livers. Dendrogram plot analysis clustered by both Euclidian and Ward methods (see 'Materials and Methods' section). (**e**) A large number of regions of differential hydroxymethylation (dHMRs) and methylation (dMRs) occur during 13 week liver maturation (i). In addition to occurring over a number of enhancer elements (ii), a subset of genes is also found to contain either a dHMR or dMR over their PPRs (iii). Regions gaining a mark ('hyper-') are represented by yellow bars, whereas losses ('hypo-') are blue. (**f**) Plot of *P*-values for tissue expression levels in genes containing a dHMR. Analysis carried out using the 'UP_Tissue' DAVID analysis tool (see 'Materials and Methods' section). (**g**) Boxplot of genes containing a dHMRs/dMRs reveal associated transcriptional changes during 13-week liver maturation. Genes with hyper-dHMRs are significantly elevated in their expression versus those with a 91 day hypo-dHMR, (i) whereas the relationship with 5mC is the opposite. (ii) $^{**}P$-value < 0.005, $^{*}P$-value < 0.05. 'Hyper-' events = yellow, 'hypo-' events = blue.

modification level. Although GO-term analysis failed to highlight any particular class of genes, analysis of tissue expression patterns using the Uniprot tissue annotation database reveals a significant enrichment for dHMRs over the promoters of genes, which are typically expressed in the liver (*P*-value $7.9E^{-7}$, Figure 1f). As gene expression data already existed for the same mice from earlier studies (53), we were able to relate the changes in expression occurring throughout liver maturation (for juvenile mice 30–33 days and older mice 91 days later) to changes in promoter DNA modifications. Typically, genes containing a gain of PPR 5hmC in the mature mice ('91 day hyper-dHMR') were found to exhibit an increase in gene expression when compared with genes containing a loss of PPR 5hmC (Figure 1g, Willcox *P*-values <0.005). Inversely, genes losing 5mC over their PPRs in the mature mice were seen to have a small but significant increase in their gene expression change to genes which had gained PPR 5mC (Figure 1g, Willcox *P*-values <0.05).

Together, these results highlight the fact that 5hmC in particular is a useful indicator of cell state throughout liver maturation; first, owing to the highly reproducible distribution of the mark between multiple individuals of a similar age as well as its relationship with associated gene expression events, which take place during liver maturation.

PB mediated perturbations to the liver hydroxymethylome are highly reproducible between individuals

We have previously reported that 28-day exposure to the drug PB results in the induction of a subset of genes, which also show reciprocal changes in both 5hmC and 5mC levels over a subset of promoters (17,34). Here, we expand on this work by investigating temporal dependent changes to the methylome and the hydroxymethylome following both short term (+1 day) and longer term PB exposure (+7, 28 and 91 days PB) to better understand the dynamics of observed epigenetic changes in response to the drug. Furthermore, as prolonged PB exposure has been shown to promote tumour formation in the mouse liver, stable changes in both gene expression and DNA modification patterns over candidate biomarkers following longer-term PB exposure may serve to better understand this progression to a cancerous state (54).

Focussing initially on mice who had received short term (1 day) doses of the drug, we found that the induced perturbations to the hydroxymethylome were not random and were instead largely reproducible between multiple liver samples (Figure 2a). This was not true for 5mC, which revealed varying levels of perturbations following the initial 1 day PB exposure (Supplementary Figure S5). Following this result, we tested whether the DNA modification patterns were sufficiently perturbed on PB exposure to allow for independent clustering of liver samples. Similar to the clustering of 5hmC patterns during liver maturation (Figure 1d), the 5hmC patterns of both short- (1 day) and long-term (91 day) drug-exposed mice were found to be more similar to one another than to their control counterparts (Figure 2b).Global perturbations to the 5mC patterns were not able to distinguish drug treated from control livers following short-term PB exposure (Supplementary Figure S6). Once more, this may be in part due to technical issues with the antibody (46–49) but may also be due to the relative abundance of the 5mC modification through the genome, which is thought to contribute to a far greater proportion of cytosine bases than 5hmC (55,56). However, clustering of livers receiving longer doses PB (91 days) based on global 5mC patterns could better separate drug exposed mice from the relevant control set (Supplementary Figure S6). One possible explanation for such a result is that 5mC is thought to be a somewhat stable epigenetic modification when compared with 5hmC. This may represent a requirement for long-term drug exposure to strongly perturb 'locked down' 5mC levels, whereas the more dynamic 5hmC modification can be rapidly perturbed.

Long-term PB exposure results in exacerbated epigenetic perturbation

Following the observation that the PB driven perturbations were reproducible between individuals, we set out to better characterize the dynamic changes to the epigenome, which occur on prolonged PB exposure. First, focussing on the epigenomes of livers exposed to short-term PB dosing, we found that many changes had already occurred in this short time frame inferring that epigenetic changes can occur rapidly in response to a drug or external stimulus. Following this short dose of PB, we found 6009 regions of differential hydroxymethylation (dHRMs: 3258 hyper, 2751 hypo) and 1604 methylation (dMRs: 1064 hyper, 540 hypo), respectively, between drug-exposed and control mice of the same age (Figure 2c). This is particularly significant, as it reveals that a far greater number of perturbations arise through drug treatment than is seen to occur naturally between individual mice (168 total dHMRs in mice aged 30–33 days compared with 6009 dHMRs in mice of the same age +1 day PB).

Analysis of dHMRs and dMRs, which arise following longer doses of PB (+7, +28 and +91 days PB treatment, reveals a continuous increase in the levels of perturbation with prolonged drug exposure. Similar to the results of the maturation of the liver in control mice, the hydroxymethylome undergoes a greater change than the methylome in response to PB, as shown by the higher number of dHMRs than dMRs at all time points tested Figure 2c).Globally, we find around twice the number of hyper-dHMRs and dMRs than hypo- events, indicating that there is an overall detectable gain in both of the modifications following drug exposure.

Mapping of these dHMRs and dMRs to genomic features present on our array platform (containing 11 kb regions surrounding annotated transcriptional start sites for 20 718 unique genes) reveals that the distribution of these dHMRs and dMRs is dependent on the length of PB treatment (Supplementary Figure S7). With prolonged PB exposure, dHMRs are typically lost from 'promoter distal' regions and instead are found in 'genic' and 'promoter flanking' sites, whereas dMRs are dynamically

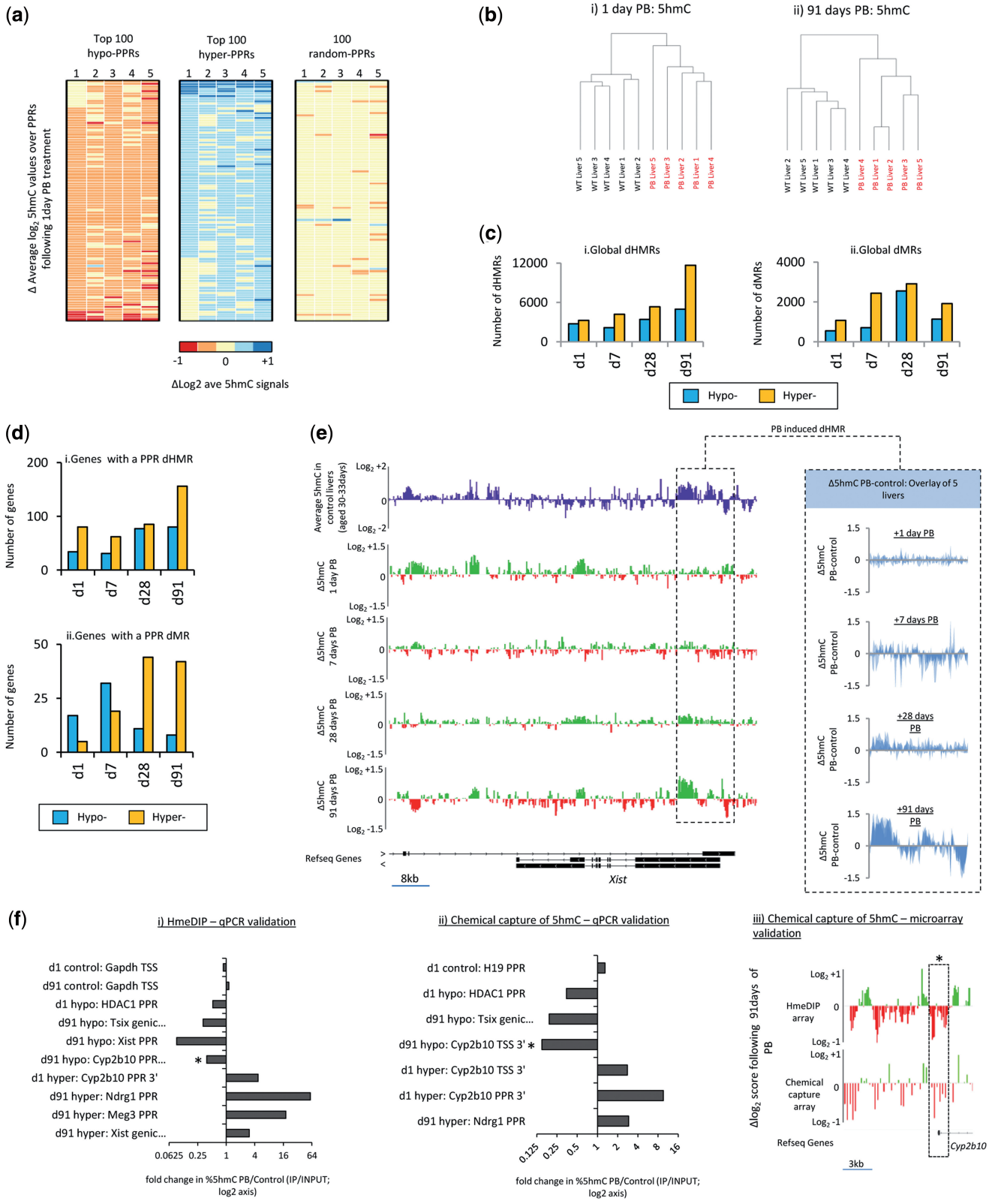

Figure 2. PB induces dose-dependant reproducible changes to the hydroxymethylome. (**a**) PB-mediated perturbations to the hydroxymethylome are reproducible between individuals. Heatmap plots of PB driven changes in the 5hmC signals over the 100 most hyper- and hypo- PB perturbed PPRs (TSS +/− 1 kb). Average changes in 5hmC levels over 100 random PPRs are also plotted. Changes in PPR 5hmC level are calculated for each liver ($n = 5$) relative to the average of the control group ($n = 5$). Plots range from $\Delta \log_2 -1$ (dark red) to $\Delta \log_2 +1$ (dark blue). These plots reveal the reproducible nature of epigenetic perturbations over PPRs strongly perturbed by PB. (**b**) In all, 5hmC patterns can stratify PB-exposed livers from control groups following both short (i) and long (ii) doses of the drug. Data sets are clustered by Euclidian and Ward methods (see 'Materials and Methods' section). (**c**) The numbers of dMRs and dHMRs in the data set vary with prolonged PB exposure. Total numbers of defined regions of

(continued)

redistributed primarily between promoter distal and genic regions (Supplementary Figure S7). Analysis of the distribution of PB driven dHMRs and dMRs at enhancer elements present on the array results in very little overlap, indicating that that enhancer elements do not undergo dynamic changes to the DNA modification landscape following PB exposure (Supplementary Figure S8). Taken together, these results highlight the fact that PB-driven epigenetic perturbations to the DNA modifications are highly dynamic and continue to change with prolonged drug exposure.

PB induces perturbations to the 5hmC and 5mC patterns over a subset of PPRs

Although it is widely recognized that the levels of 5hmC are lower at regions surrounding the transcriptional start sites and elevated in the genic regions throughout annotated genes (18,30,37,57–59), we previously suggested that changes to 5hmC over PPRs are functionally linked to changes in liver gene expression occurring upon 28 days of PB exposure (17). Here, we investigate the changes to both the 5hmC and 5mC levels over PPRs in the mouse liver following short (1 day) and prolonged (7, 28 and 91 days) PB exposure. Although 5hmC levels are typically low directly over a TSS, we find that many genes contain a dHMR in or around their PPRs spanning and flanking these TSS regions (Figure 2d, For full list of genes, see Supplementary Table S4). It also appeared that the number of genes with either a dHMR or dMR in a PPR was dependant on the duration of PB dosing (Figure 2d and e; the *Xist* gene). At these dHMRs and dMRs, we find that the PB-induced perturbations are highly reproducible between samples with many exhibiting a gradual shift in the levels of change in relation to the length of PB dosing (Figure 2e and Supplementary Figure S9). Therefore, exposure to the drug PB can lead to rapid changes in 5hmC and 5mC levels at select promoter loci following transient exposure to a NGC (1 day PB), which are intensified following longer periods (7–91 day) of drug dosing.

Validation of dHMRs and dMRs reveals strength of hmeDIP-array analysis

Following the identification of PB-mediated dHMRs and dMRs, we set out to independently validate the results of the array analysis on samples, which had received both a short (1 day) and long-term (+91 day) drug doses. qPCR validation was initially carried out on HmeDIP and meDIP enriched material, and in each case, the regions defined as losing or gaining either mark were successfully validated using this technique (Figure 2f–i and Supplementary Figure S10). Although the 5hmC antibody used in our hmeDIP has previously been shown to be highly specific (17), we also used a recently described chemical capture technique that relies on the selective enrichment of biotinylated and glycosylated 5hmC fragments by streptavidin beads to independently enrich for fragments of DNA containing the 5hmC mark (18) and then follow this by qPCR over the previously identified dHMRs (Figure 2f–ii). The results of this analysis complement those of the hmeDIP-qPCR experiments, once more validating the results of the microarrays. Finally, we couple the chemical capture technique to tiling microarrays of lower probe density to independently profile PB mediated changes following long-term (91 day) drug exposure. In doing so, we find that this new data set parallels our HmeDIP-array results (Figure 2f–iii and Supplementary Figure S11). Overall, these independent validations reveal the extent of perturbation to select regions of the epigenome following PB exposure.

Perturbation of the transcriptome following long term PB exposure

To study changes to the liver transcriptome following PB exposure, Affymetrix expression arrays were carried out on both control and PB treated mice following 1, 7, 28 and 91 days of drug exposure (53). Genes showing strong responses to PB (>$\log_2$ 1.5-fold induction) were grouped into one of five classes based on their expression dynamics (Figure 3a and Supplementary Table S5). Genes were first grouped on their relative responsiveness to PB ('rapid', 'slow', 'late' or 'no response') and then further classified by the retention of this induction; either returning to control levels at later time points of dosing ('dynamic') or remaining stably over-expressed at later time points ('prolonged'). From this classification, we found 44 genes that were strongly induced following short (1 day) PB exposure, which either remained so with prolonged drug exposure ('rapid prolonged response', $n = 17$) or were seen to return to control levels ('rapid dynamic

Figure 2. Continued
significant 5mC/5hmC change are plotted for each mark against the dosage time. 'Hyper-' = yellow bars, 'hypo-' = blue bars. (**d**) A number of genes contain an associated promoter proximal dMR/dHMR. (**e**) Visualization of 5hmC perturbation patterns over a PB-induced dHMR. The average 5hmC profiles of 30–33-day-old control mice (purple) are shown with respect to the changes in 5hmC following 1, 7, 28 or 91 days of PB dosing (PB minus control $\log_2$ values per probe; green = gain in 5hmC, red = loss of 5hmC). Annotated Refseq genes are represented below as black bars including gene structures. A region containing a dHMR is highlighted by a box in the main plot. For this dHMR, average 5hmC changes in five liver samples are overlaid on the right in the expanded box for each time point of dosing. Individual changes in 5hmC are represented by light blue plots. Patterns for the 5hmC change in each mouse (light blue) are superimposed so that darker regions highlight regions of reproducible 5hmC change. All data are represented by normalized $\log_2$ scores for all probes. (**f**) Independent validation of array defined dHMRs. qPCR validation of dHMRs carried following the hmeDIP technique used to generate the arrays (i) as well as on an antibody independent chemical capture technique (ii). Data are plotted as fold change in the percentage of 5hmC enrichment/Input for PB-treated mice versus control mice and plotted on a $\log_2$ scale to represent data accordingly. In all, 5hmC-enriched fragments generated by chemical capture were also analysed by hybridization to a tiling array (iii) and was found to compliment the perturbations seen in the hmeDIP-array dataset, specifically over the dynamic *Cyp2b10* locus (denoted by * in validation techniques i, ii and iii). Data are plotted as log2 scores for all probes across this region with annotated Refseq genes represented below as black bars including gene structures. Scale bar is shown below as black bar.

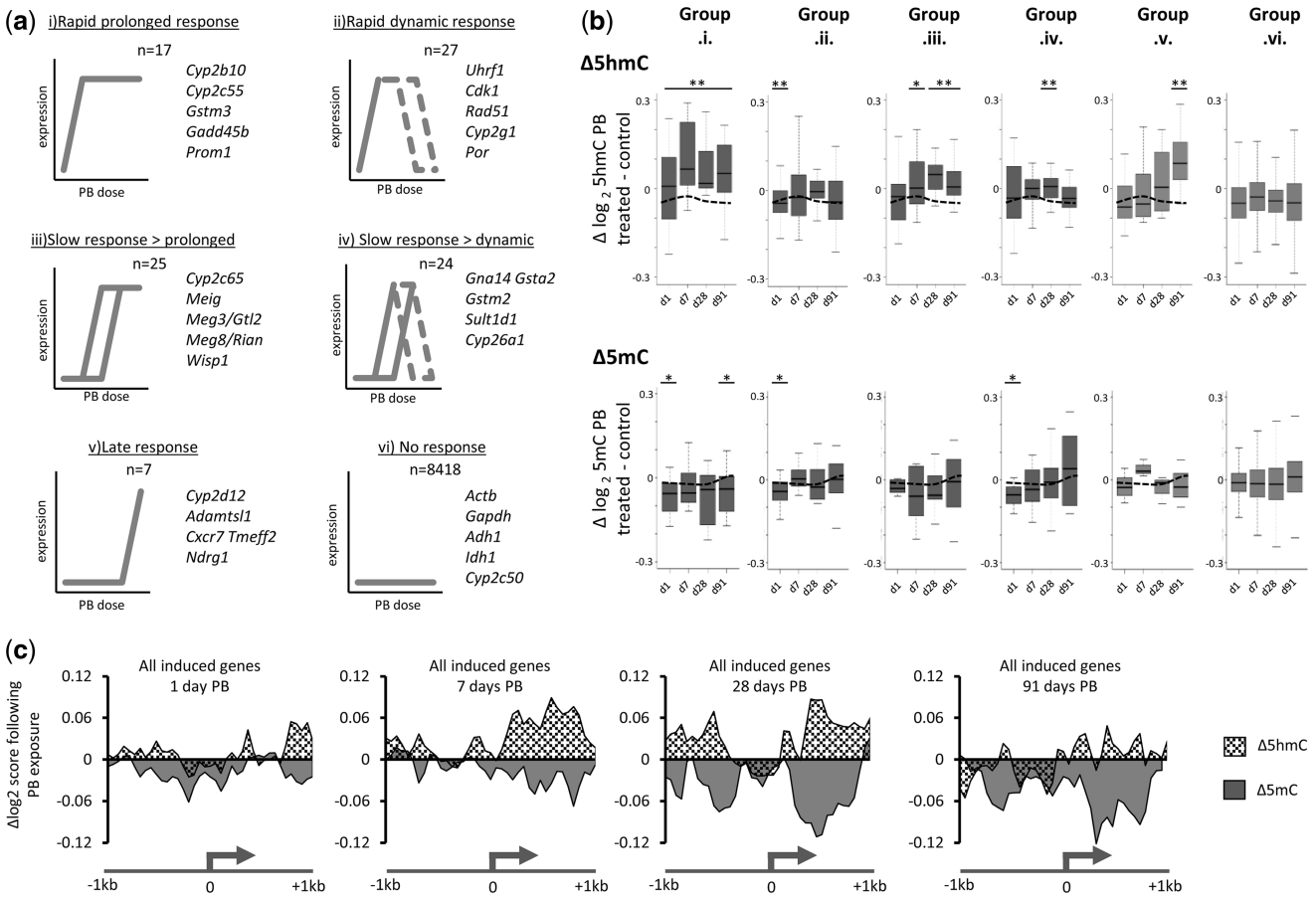

Figure 3. Strongly induced PB genes reveal elevated promoter proximal 5hmC levels around the time of initial induction, which are lost with continuous expression. (**a**) Exposure to PB results in the strong mis-regulation of a set of genes, which can be clustered on their expression dynamics. Genes showing strong levels of transcriptomic perturbation (>$\log_2$ 1.5-fold change) are first grouped based on the speed of the PB response: 'rapid' (1 day PB), 'slow' (7 and 28 days PB) or 'late' (91 days PB). Additionally, these genes were either found to maintain this strong induction ('prolonged') or return to basal levels ('dynamic').The majority of genes exhibited no PB response (<$\log_2$ 0.25–>$\log_2$ 0.25-fold change, $n = 8418$). The number of genes in each group is shown above each plot with five examples from each subset shown to the right. (**b**) Box plot analysis of the change in either 5hmC levels (top plots) or 5mC levels (lower plots) occurring over the PPRs of genes in each cluster outlined in Figure 3a. Classes are indicated by roman numerals as in Figure 3a. The dashed line denotes levels of change in either mark over genes exhibiting no PB affect. ** denotes *P*-value < 0.005 over un-changes genes at each time point, * *P*-value < 0.05 over un-changed genes). (**c**) PB-induced changes to 5mC and 5hmC patterns at transcriptionally induced genes are highly reciprocal and dynamic. Average changes to the 5mC (solid black) and 5hmC (dashed) patterns are plotted over the PPRs of strongly induced genes at each time point (>log2 1.5-fold change following PB). Although 5hmC is initially elevated in promoter flank regions, 5mC remains lost from the entire PPR of induced genes. With continued PB exposure (28 day), both marks are seen to be lost from promoter 'core' regions, whereas at prolonged (91 day) PB exposure, both marks are lost from the entire PPR.

response', $n = 27$). Genes belonging to the former group likely represent those involved in the turnover and metabolism of the drug as was verified by the fact that many members of this group are involved in the removal of xenobiotics from the cell, such as the Cytochrome P450 (*Cyp*) and Glutathione-S-transferase (*Gst*) families of genes. Genes exhibiting delayed yet prolonged induction following 7 or 28 days PB exposure are likely to be secondary events in either the removal of drugs from the liver or may be involved in pre-NGC tumour formation resulting from prolonged PB exposure. As such, it was interesting to note that several members of this group have associated roles in the progression of tumorigenesis including the gene *Wisp1* (WNT1 inducible signalling pathway protein 1) and the two imprinted non-coding RNAs *Meg3* (Maternally expressed 3, also known as *Gtl2*) and *Meg8* (Maternally expressed 8, also known as *Rian*). The later of these two genes reside within the imprinted *Dlk1-Dio3* locus; a region of the genome in which misregulated non-coding transcripts have been observed in mouse and human HCC (41,53). Additionally, these transcripts have proposed roles in the regulation of cellular pluripotency (60,61) as well as being found to define a subtype of stem cell-like tumours (41). *Gadd45b*, *Prom1* and *Rad51* are also rapidly induced on PB exposure. These genes all have proposed roles in response to DNA damage, regulation of the cell cycle and maintenance of stem cell state. Their activation in a temporal manner reinforces earlier studies that report induction of these genes in response to PB (34,53,62,63). As such, this group may represent a class of genes whose perturbation may facilitate the progression towards a tumorigenic state following NGC exposure by creating a pre-neoplastic foci before subsequent mutational change.

Finally, we found a small cohort of genes, which are only strongly induced following 13 weeks of continuous PB exposure (termed 'late response' $n = 7$, Figure 3a). In a similar manner to the group of genes exhibiting a slow and prolonged response to PB, these genes are of interest due to their potential roles in the progression to a tumour state (54,64). Within this group, we identified another Cytochrome P450 gene, *Cyp2d12*, as well as a tumour suppressor *Ndrg1* (N-myc downstream regulated 1) and the chemokine receptor *Cxcr7* both of which have been shown recently to be strongly induced in both human and mouse HCC (42,65). Taken together, the expression changes, which occur during short- and long-term PB exposure, reveal a series of genes with roles in both the removal of the drug from the liver as well as the identification of potential biomarkers for PB driven tumorigenesis.

Promoter proximal 5hmC levels are related to changes in gene expression following PB exposure

To investigate whether these changes in gene expression are related to changes in promoter DNA modification status, we investigated changes to promoter proximal 5hmC and 5mC levels in the five subclasses of PB induced genes as well as the un-induced set of genes (Figure 3b). Overall, we find that 5hmC levels are typically significantly increased over PPRs of strongly induced genes around the time of initial transcriptional activation (P-values < 0.005) with a maintenance of this elevated 5hmC found at genes with prolonged expression (Figure 3b—upper panel; 'rapid prolonged', 'slow prolonged' and 'late' plots). Interestingly, in genes whose expression levels are not maintained in an elevated state after initial gene activation (Figure 3b—upper panel; 'rapid dynamic' and 'slow dynamic'), the levels of 5hmC perturbation follow changes in the transcriptional activity and do not show a significant elevation over the un-induced genes at later time points. The relationship between 5mC over PPRs and transcriptional perturbation following PB are less clear than 5hmC; however, 5mC levels are significantly reduced at genes, which show rapid induction following PB exposure (P-values < 0.05) and to some degree remain lower with prolonged exposure (Figure 3b—lower; 'rapid prolonged').Taken together, these results highlight the epigenetic changes occurring over the promoters of PB-induced genes and reveal the intimate relationship between 5hmC levels and gene expression state.

Due to the fact that 5hmC is a derivative of the 5mC modification itself (15,59), we set out to test the relationship between 5hmC and 5mC induced changes over these PPRs. Plots of the average changes in each modification across the PPRs of strongly induced genes ($>\log_2$ 1.5-fold change) at each time point of analysis (+1, 7, 28, 91 days PB) reveal that the patterns of 5hmC and 5mC change are both dynamic and reciprocal (Figure 3c). In all, 5mC levels are found to be rapidly lost following transient (1 day) PB exposure whilst there is an elevation in 5hmC levels over the promoter flanking regions. These perturbations appear to be at their most dynamic following 7 or 28 days of continuous drug exposure (strong gain of 5hmC/strong loss of 5mC) before loss of both modifications following prolonged (91 day) drug exposure. Due to the proposed role of 5hmC as a demethylation intermediate, this may represent a completion of demethylation over these PPRs following gene induction, which would likely create an open and accessible region around the promoter that can facilitate the engagement of the transcriptional machinery and accessory factors.

Cytochrome P450 family members display rapid and dynamic epigenetic perturbations following NGC exposure

A common family of genes, which appear both transcriptionally induced as well as routinely epigenetically perturbed are that of the cytochrome P450 (*Cyp*) genes. The CAR in particular plays an essential role in PB-induced hepatocarcinogenesis in rodents (66); however, it is uncertain whether increased cytochrome P450 enzyme activity itself plays a direct role in tumour formation (67). The response of several of the *Cyp* genes have been widely studied following PB exposure (34,35,37,68–70), as this family of genes are involved in the detoxification of electrophilic compounds, including carcinogens (71–73). We previously found that 28 days of continuous exposure to PB resulted in epigenetic and transcriptomic perturbations to these *Cyp* genes (17).

We find that many of the *Cyp* genes are strongly induced following continuous PB exposure (Figure 3a) with four members belonging to the 'rapid prolonged' response group of genes (*Cyp2b10, Cyp2c55, Cyp2b13 and Cyp2g1*) and five more found to be in either the 'slow' or 'late' response classes (*Cyp2c65, Cypc37, Cypb9, Cyp26a1* and *Cypd12;* Figure 3a). Due to the fact that 5hmC shows dynamic and strong perturbations in response to drug exposure, we investigated the changes over the PPRs of these genes in each of the five individual livers to investigate the reproducibility of epigenetic perturbation at these loci. In agreement with our earlier observations (Figures 1b and c), we find that the changes to the hydroxymethylome are highly conserved between individuals exposed to PB depending on the length of dosing (Figure 4a). We also observe that these PPRs undergo a directional change in the 5hmC patterns depending on the length of drug dosing. Similar to the results described earlier (Figure 3c), these *Cyp* genes initially gain 5hmC at flanking sites outside of the promoter core region before progressing to a general loss of the mark, which appears to originate from the TSS. As described above for genes strongly induced following long-term PB exposure, this may represent active demethylation through a 5hmC intermediate (17).

In agreement with our earlier 28 day focused study (17), we find that the PPRs of the *Cyp2b10* gene exhibit some of the most dramatic perturbations to both the hydroxymethylation and methylation patterns coupled to strong PB-mediated induction of expression at all time points tested. *Cyp2b10* mis-expression has also been observed in a subset of liver tumours that occur after the initial inducer compound has been withdrawn or arose in the absence of exposure to PB (66,74,75) and liver tumours that are glutamine synthetase positive and mutated in β-catenin show concomitantly elevated

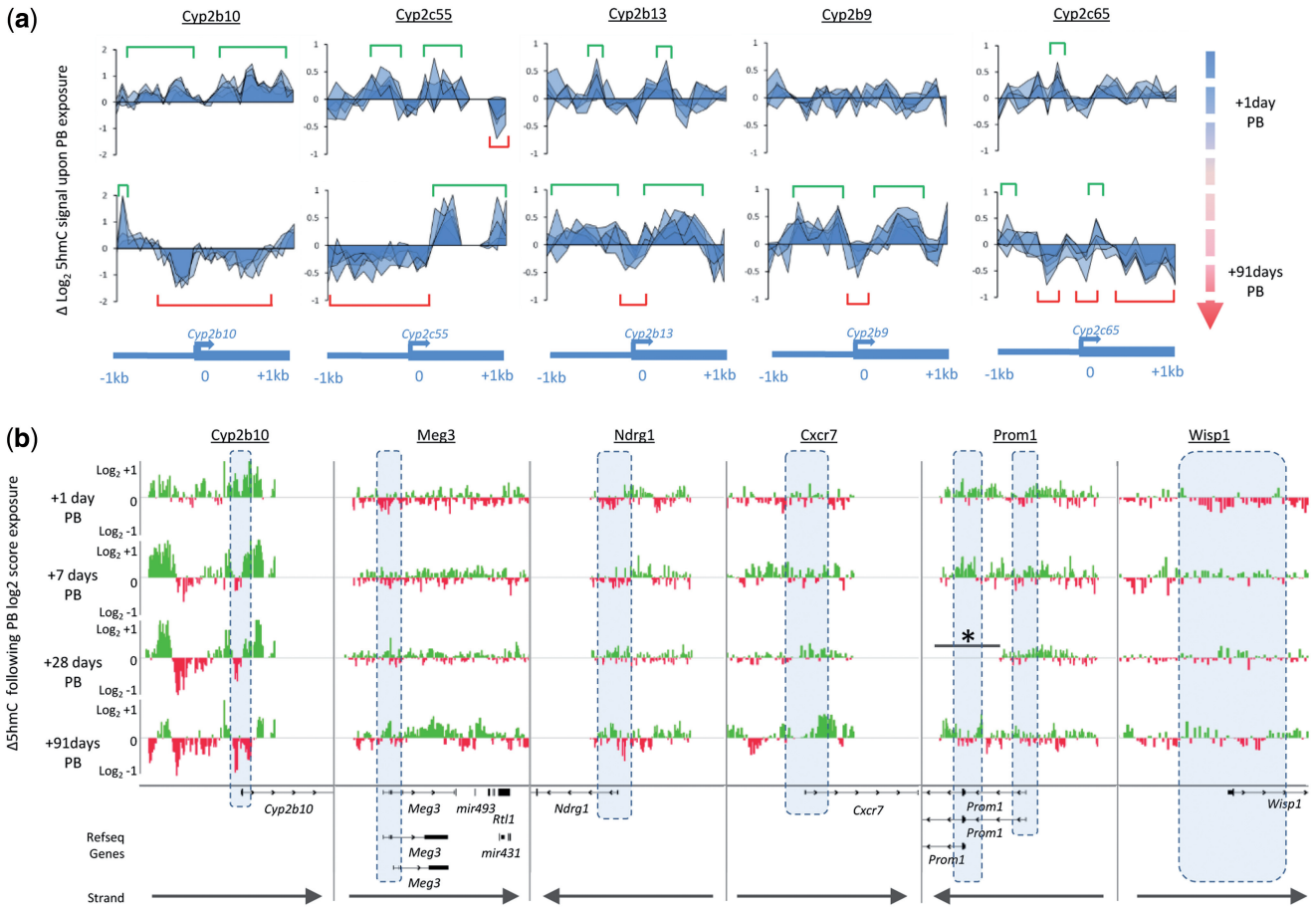

Figure 4. Analysis of candidate PB exposure biomarkers with potential roles in the progression of tumorigenesis (**a**). The cytochrome P450 family of genes reveal reproducibly dynamic levels of perturbation to their promoter hydroxymethylation profiles following both short- and long-term PB exposure. Changes in the patterns of five PB treated mice (+1 and +91 days PB) relative to the average of the control mice ($n = 5$) are plotted over the PPRs of a set of strongly induced *Cyp* genes (Δ $\log_2$ values for each probe over these regions). Patterns for each mouse are superimposed so that darker regions highlight regions of reproducible 5hmC change. Regions of strong 5hmC gain (>$\log_2$ 1.5-fold increase) are denoted by green brackets, whereas red brackets denote regions of strong 5hmC loss (>$\log_2$ 1.5-fold decrease).Structures of the promoter regions are shown below in blue. (**b**) Following PB exposure, perturbations to the 5hmC patterns in a subset of strongly induced genes facilitate the identification of candidate biomarkers for the progression of NGC. Plots show changes in the 5hmC $\log_2$ scores at the four time points of PB dosing over six potential biomarkers. Plots below the zero line indicate net loss of 5hmC following PB exposure (red bars), whereas a gain represents an increase in the modification (green bars). Dashed boxes indicate PPRs (TSS $\pm$ 1 kb) with an overlapping PPR found at an alternative promoter at the Meg3 locus, which is boxed as one region. Refseq annotated genes are displayed below with the direction of the gene indicated with an arrow. Asterisk denotes a region downstream of the Prom1 gene missing from the 28 day data set due to a change in probe locations on the microarrays.

Cyp2b10 expression (38). Thus, strong and prolonged upregulation of this gene may be of critical importance for the progression towards tumorigenesis.

Combined epigenomic and transcriptomic approaches identify potential biomarkers for the early progression of non-genotoxic carcinogenesis

Through our combined epigenomic and transcriptomic analysis, we find a series of interesting candidates, which may be used as biomarkers for the progression of non-genotoxic carcinogenesis. Candidates were selected based off of their transcriptional response to PB in addition to any associated epigenetic perturbations. As pointed out earlier, many of these have roles in the regulation of the cell cycle, DNA damage response pathways as well as tumour suppression (Figure 4b and Supplementary Figure S12).

One such candidate biomarker is that of the non-coding RNA *Meg3*. This gene resides in the imprinted Dlk1/Dio3 locus whose transcripts have been found to be elevated in many liver tumours as well as in response to PB (53,76). Furthermore mis-regulation of imprinted genes thought to be a key event in carcinogenesis (77). Here, we find for *Meg3*, that the levels of 5hmC associated with this gene increase with prolonged PB exposure (91 days), particularly at regions surrounding the alternative (downstream) promoter as well as in the body and 3′ end of the gene (Figure 4b and Supplementary Figure S12). Another candidate biomarker is that of the late PB response gene and tumour suppressor *Ndrg1*, which exhibits a loss of 5hmC directly over the TSS and promoter proximal changes in 5hmC levels around the same time with which the gene becomes strongly induced (91 days PB, Figure 4b and Supplementary Figure S12). In support of our selection

of this gene as a potential biomarker for NGC progression, studies have shown *Ndrg1* to be highly expressed in many HCC with associated poor prognosis (65). The same pattern of promoter proximal 5hmC gain is also noted for the chemokine receptor *Cxcr7* (Figure 4b and Supplementary Figure S12). There was also a marked and continual upregulation in the transcriptional activity of the *Prom1* gene, which has roles in the maintenance of stem cell state and the Wnt signalling pathway gene *Wisp1* following 1 and 7days of PB exposure, respectively, with associated increases in the promoter proximal 5hmC levels at these two genes (Figure 4b). Mis-regulation of the Wisp1 gene has been previously shown to play an important role in the progression of malignant transformation in both mouse and human HCC (78), making it an ideal candidate biomarker for the progression of NGC following PB exposure. As such, these combined epigenomic and transcriptomic analyses have identified a candidate set of biomarkers, which reveal PB response genes, which may have roles in the progression to a tumour state. Further work analysing the epigenetic and transcriptomic changes across these genes in mice with PB-induced tumours may help to better characterize and validate these loci as biomarkers for the early progression of NGC.

DISCUSSION

The proposed role of 5hmC as an intermediate in DNA demethylation pathway has led to a renewed interest in the study of the DNA methylation modification and its derivatives in both human and mammalian models. Furthermore, 5hmC patterns have been demonstrated to change throughout development alongside transcriptional changes (13). As such, the mark is an ideal identifier of tissue and cell state. Here, we present a detailed study of promoter and promoter proximal 5hmC patterns in the mouse liver between a large number of mice over a range of ages (30–124 days old). From this analysis on high density promoter microarrays, we show for the first time that the 5hmC patterns are highly reproducible between individuals of a similar age, and that these patterns change as the mouse liver ages. Furthermore, these changes in promoter 5hmC were seen to correspond to genes whose expression state also changed during late liver development.

In addition to the study of normal 5hmC perturbations, which arise during development, we also study the effects that both short- and long-term exposure of the widely studied NGC, PB exhibits on the liver hydroxymethylomes and methylomes. In the rodent liver, the anticonvulsant PB is widely used as an established model to study the progression of non-genotoxic carcinogenesis, with prolonged PB exposure resulting in liver tumour formation (38,54,63,79). As this progression to a tumour state initially occurs in the absence of genetic perturbation, changes to the epigenome likely underpin this process. With this in mind, we set out to better understand PB driven epigenetic perturbations, which may ultimately result in the progression to a tumorigenic state. Following the study of the control mice, we show that the 5hmC patterns alone undergo reproducible levels of perturbation between individuals exposed to PB. Furthermore, these perturbations allow the independent clustering of samples based on PB exposure, which fits with the notion that 5hmC patterns can be used as a cell state/tissue identifier (16). We then defined regions of differential hydroxymethylation and methylation, which we independently validated, and then mapped these to the genome to reveal that the global distribution of these loci change with prolonged drug exposure. Several studies have previously investigated the relationships between 5hmC and transcriptional activity (4,8,9,11,16–18,57,80). We expand these findings in this study by analysing select genes, which respond strongly to PB administration and find that levels of 5hmC are elevated over the PPRs of strongly induced genes. Furthermore, we show that PPR 5hmC and 5mC patterns are dynamic and reciprocal following both short (12–24 h) and long (91 days) dosing over strongly induced genes.

As long-term PB exposure is known to induce tumorigenesis in murine models (54), the genes, which show delayed and heightened elevation in their gene expression levels following PB exposure, make ideal candidates as early biomarkers for the progression of non-genotoxic carcinogenesis. Here, we highlight a few such candidates, all of which have roles in the regulation of the cell cycle, maintenance of an ES cell state and involvement in tumour suppression of *Wnt* signalling pathways. Many of these genes have already been described as transcriptionally perturbed in liver cancers and human HCC (41,42,78).

This work allows us to propose a model in which epigenetic perturbations occur at PPRs following short- and long-term PB exposure and in particular how these may contribute to the progression of tumour formation (Figure 5). Cells receiving PB will rapidly activate a series of genes involved in the removal of xenobiotics such as the cytochrome P450 member *Cyp2b10* (68,69,79). As these genes are normally silent, activation of the gene will first rely on the conversion of any promoter 5mC to non-modified cytosine through a 5hmC intermediate; an effective example of active demethylation (Figure 5 - upper). It will be of great interest to study the epigenetic environment of this promoter region once the drug has been removed and the gene silenced once more. Prolonged drug exposure has been shown to result in tumour formation (54), and our data suggest that this may be partially due to the epigenetic perturbation of normal promoter DNA modification state and subsequent activation, which may become constitutive, of a cohort of genes with roles in cell cycle regulation and progression towards tumorigenesis (Figures 4b and 5ii). In collaboration with the Innovative Medicines Initiative MARCAR Consortium (www.imi-marcar.eu), we are currently evaluating potential tissue, gender, strain and species (mouse versus rat) differences in PB-induced perturbations of the DNA methylome. The mechanistic basis and functional significance of early PB-induced changes in specific DNA methylation biomarkers is also being explored using (i) reversibility studies; (ii) PB-promoted liver tumours;

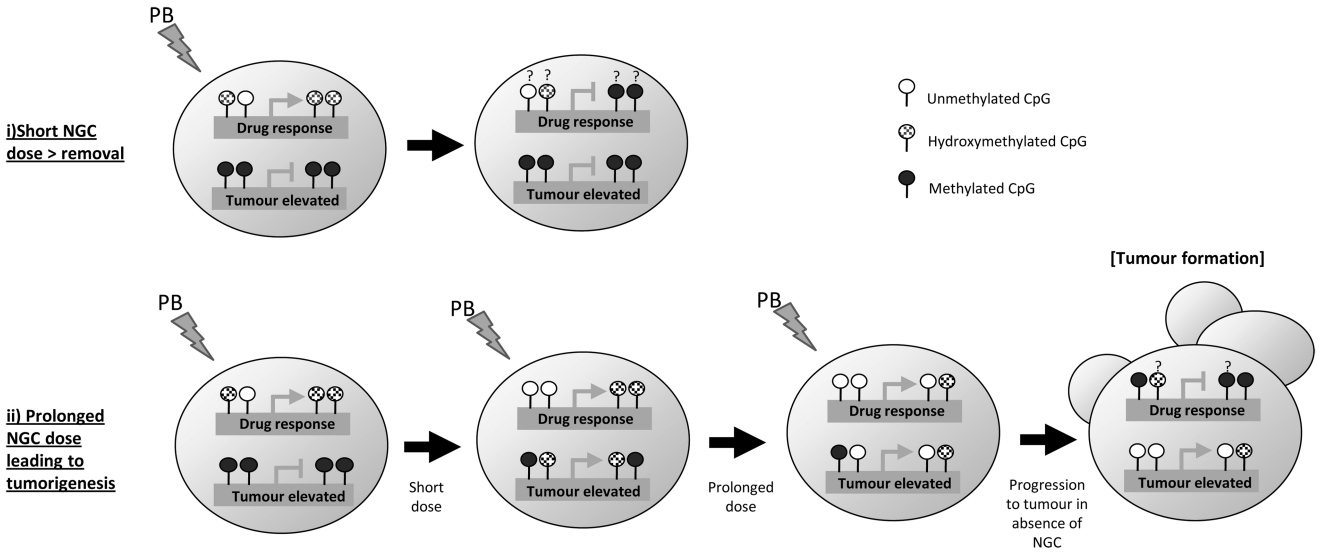

Figure 5. A model for the role of NGC induced promoter DNA modification in the progression to a tumour state. Following transient PB exposure, liver cells activate a series of PB response genes, some of which are involved in the removal of drugs. The promoters of these genes are converted from a silent 5mCpG state (black lollipops) to an active 5hmCpG (dashed lollipops) or unmodified CpG (white lollipops) rich environment (i and ii - left). On removal of this drug following short-term exposure, these response genes are turned off again and although the status of the epigenetic environment is currently untested; there may be some epigenetic memory of the activation event (denoted by a '?'). With prolonged PB exposure (ii), PB activates genes, which may contribute to tumour formation and the epigenetic status of these promoter regions also lose 5mC and gain 5hmC and un-modified cytosines. With long-term NGC exposure, there is also a demethylation of these genes through a 5hmC intermediate pushing them past an epigenetic 'point of no return'. On removal of the drug (after long-term dosing), the perturbed epigenetic landscape of these promoters aid in the constitutive activation of these genes and may lead to neoplasia.

(iii) transgenic mouse models for key nuclear receptors [CAR/Pregnane X receptor (PXR); knockout and humanized] and cancer signalling pathways (beta-catenin; liver-specific knockout); and (iv) liver tumour-sensitive versus tumour-resistant mouse strains. Furthermore, the sensitivity, specificity, dose response and reversibility of early DNA methylation biomarkers is also being explored for a range of genotoxic and non-genotoxic hepatocarcinogens with distinct modes of action (e.g. nuclear receptor mediated versus tissue injury and regeneration).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–6 and Supplementary Figures 1–12.

ACKNOWLEDGEMENTS

J.T., J.M. and R.M. conceived and designed the experiments. J.T., H.L. and J.H. performed the experiments. J.T., R.M. and A.M. analysed the data. J.T. and R.M. wrote the article. J.M., R.T., H.L. and A.M. edited the article. The authors thank Colm Nestor for his insight and assistance. All IMI-MARCAR consortium partners had a role in study design, data collection and analysis, decision to publish or preparation of the manuscript. J.T. is a recipient of IMI-MARCAR funded career development fellowships at the MRC HGU. R.M. and J.H. are supported by the Medical Research Council. Work in R.M.'s laboratory is supported by IMI-MARCAR, the BBSRC and the MRC. H.L. is the recipient of a NIBR Postdoctoral Fellowship. We thank Shaun Webb and Alastair Kerr (WTCCB) for Bioinformatic assistance.

FUNDING

Funding was provided by the Innovative Medicine Initiative Joint Undertaking (IMI JU) under grant agreement number 115001 (MARCAR project). URL: http://www.imi-marcar.eu/. Novartis and the MRC are full participants in the IMI MARCAR consortium. Novartis provide in kind financial contribution to the scientific program. Funding for open access charge: Funding from grants (MRC, BBSRC, IMI-MARCAR) to Edinburgh University.

Conflict of interest statement. None declared.

REFERENCES

1. Bird,A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev.*, **16**, 6–21.
2. Wu,H. and Zhang,Y. (2011) Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev.*, **25**, 2436–2452.
3. Hackett,J.A., Reddington,J.P., Nestor,C.E., Dunican,D.S., Branco,M.R., Reichmann,J., Reik,W., Surani,M.A., Adams,I.R. and Meehan,R.R. (2012) Promoter DNA methylation couples genome-defence mechanisms to epigenetic reprogramming in the mouse germline. *Development*, **139**, 3623–3632.
4. Wu,H., D'Alessio,A.C., Ito,S., Wang,Z., Cui,K., Zhao,K., Sun,Y.E. and Zhang,Y. (2011) Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in

transcriptional regulation in mouse embryonic stem cells. *Genes Dev.*, **25**, 679–684.
5. Wu,H., D'Alessio,A.C., Ito,S., Xia,K., Wang,Z., Cui,K., Zhao,K., Eve Sun,Y. and Zhang,Y. (2011) Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature*, **473**, 389–393.
6. Pastor,W.A., Pape,U.J., Huang,Y., Henderson,H.R., Lister,R., Ko,M., McLoughlin,E.M., Brudno,Y., Mahapatra,S., Kapranov,P. *et al.* (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature*, **473**, 394–397.
7. Williams,K., Christensen,J., Pedersen,M.T., Johansen,J.V., Cloos,P.A., Rappsilber,J. and Helin,K. (2011) TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature*, **473**, 343–348.
8. Stroud,H., Feng,S., Morey Kinney,S., Pradhan,S. and Jacobsen,S.E. (2011) 5-hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol.*, **12**, R54.
9. Xu,Y., Wu,F., Tan,L., Kong,L., Xiong,L., Deng,J., Barbera,A.J., Zheng,L., Zhang,H., Huang,S. *et al.* (2011) Genome-wide regulation of 5hmC, 5mC, and gene expression by tet1 hydroxylase in mouse embryonic stem cells. *Mol. Cell*, **42**, 451–464.
10. Yu,M., Hon,G.C., Szulwach,K.E., Song,C.X., Zhang,L., Kim,A., Li,X., Dai,Q., Shen,Y., Park,B. *et al.* (2012) Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell*, **149**, 1368–1380.
11. Ficz,G., Branco,M.R., Seisenberger,S., Santos,F., Krueger,F., Hore,T.A., Marques,C.J., Andrews,S. and Reik,W. (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature*, **473**, 398–402.
12. Serandour,A.A., Avner,S., Oger,F., Bizot,M., Percevault,F., Lucchetti-Miganeh,C., Palierne,G., Gheeraert,C., Barloy-Hubler,F., Peron,C.L. *et al.* (2012) Dynamic hydroxymethylation of deoxyribonucleic acid marks differentiation-associated enhancers. *Nucleic Acids Res.*, **40**, 8255–8265.
13. Bocker,M.T., Tuorto,F., Raddatz,G., Musch,T., Yang,F.C., Xu,M., Lyko,F. and Breiling,A. (2012) Hydroxylation of 5-methylcytosine by TET2 maintains the active state of the mammalian HOXA cluster. *Nat. Commun.*, **3**, 818.
14. Booth,M.J., Branco,M.R., Ficz,G., Oxley,D., Krueger,F., Reik,W. and Balasubramanian,S. (2012) Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science*, **336**, 934–937.
15. Tahiliani,M., Koh,K.P., Shen,Y., Pastor,W.A., Bandukwala,H., Brudno,Y., Agarwal,S., Iyer,L.M., Liu,D.R., Aravind,L. *et al.* (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, **324**, 930–935.
16. Nestor,C.E., Ottaviano,R., Reddington,J., Sproul,D., Reinhardt,D., Dunican,D., Katz,E., Dixon,J.M., Harrison,D.J. and Meehan,R.R. (2012) Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res.*, **22**, 467–477.
17. Thomson,J.P., Lempiainen,H., Hackett,J.A., Nestor,C.E., Muller,A., Bolognani,F., Oakeley,E.J., Schubeler,D., Terranova,R., Reinhardt,D. *et al.* (2012) Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. *Genome Biol.*, **13**, R93.
18. Song,C.X., Szulwach,K.E., Fu,Y., Dai,Q., Yi,C., Li,X., Li,Y., Chen,C.H., Zhang,W., Jian,X. *et al.* (2011) Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat. Biotechnol.*, **29**, 68–72.
19. Ko,M., Huang,Y., Jankowska,A.M., Pape,U.J., Tahiliani,M., Bandukwala,H.S., An,J., Lamperti,E.D., Koh,K.P., Ganetzky,R. *et al.* (2010) Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*, **468**, 839–843.
20. Koh,K.P., Yabuuchi,A., Rao,S., Huang,Y., Cunniff,K., Nardone,J., Laiho,A., Tahiliani,M., Sommer,C.A., Mostoslavsky,G. *et al.* (2011) Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell*, **8**, 200–213.
21. Wossidlo,M., Nakamura,T., Lepikhov,K., Marques,C.J., Zakhartchenko,V., Boiani,M., Arand,J., Nakano,T., Reik,W. and Walter,J. (2011) 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat. Commun.*, **2**, 241.
22. Ito,S., D'Alessio,A.C., Taranova,O.V., Hong,K., Sowers,L.C. and Zhang,Y. (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*, **466**, 1129–1133.
23. Ito,S., Shen,L., Dai,Q., Wu,S.C., Collins,L.B., Swenberg,J.A., He,C. and Zhang,Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, **333**, 1300–1303.
24. Cortellino,S., Xu,J., Sannai,M., Moore,R., Caretti,E., Cigliano,A., Le Coz,M., Devarajan,K., Wessels,A., Soprano,D. *et al.* (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell*, **146**, 67–79.
25. He,Y.F., Li,B.Z., Li,Z., Liu,P., Wang,Y., Tang,Q., Ding,J., Jia,Y., Chen,Z., Li,L. *et al.* (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, **333**, 1303–1307.
26. Inoue,A., Shen,L., Dai,Q., He,C. and Zhang,Y. (2011) Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. *Cell Res.*, **21**, 1670–1676.
27. Inoue,A. and Zhang,Y. (2011) Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science*, **334**, 194.
28. Hackett,J.A., Sengupta,R., Zylicz,J.J., Murakami,K., Lee,C., Down,T.A. and Surani,M.A. (2012) Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science*, **339**, 448–452.
29. Seisenberger,S., Andrews,S., Krueger,F., Arand,J., Walter,J., Santos,F., Popp,C., Thienpont,B., Dean,W. and Reik,W. (2012) The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol. Cell*, **48**, 849–862.
30. Lian,C.G., Xu,Y., Ceol,C., Wu,F., Larson,A., Dresser,K., Xu,W., Tan,L., Hu,Y., Zhan,Q. *et al.* (2012) Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell*, **150**, 1135–1146.
31. Moran-Crusio,K., Reavie,L., Shih,A., Abdel-Wahab,O., Ndiaye-Lobry,D., Lobry,C., Figueroa,M.E., Vasanthakumar,A., Patel,J., Zhao,X. *et al.* (2011) Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell*, **20**, 11–24.
32. Figueroa,M.E., Abdel-Wahab,O., Lu,C., Ward,P.S., Patel,J., Shih,A., Li,Y., Bhagwat,N., Vasanthakumar,A., Fernandez,H.F. *et al.* (2010) Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell*, **18**, 553–567.
33. Kallin,E.M., Rodriguez-Ubreva,J., Christensen,J., Cimmino,L., Aifantis,I., Helin,K., Ballestar,E. and Graf,T. (2012) Tet2 facilitates the derepression of myeloid target genes during cebpalpha-induced transdifferentiation of pre-B cells. *Mol. Cell*, **48**, 266–276.
34. Lempiainen,H., Muller,A., Brasa,S., Teo,S.S., Roloff,T.C., Morawiec,L., Zamurovic,N., Vicart,A., Funhoff,E., Couttet,P. *et al.* (2011) Phenobarbital mediates an epigenetic switch at the constitutive androstane receptor (CAR) target gene Cyp2b10 in the liver of B6C3F1 mice. *PLoS One*, **6**, e18216.
35. Phillips,J.M., Yamamoto,Y., Negishi,M., Maronpot,R.R. and Goodman,J.I. (2007) Orphan nuclear receptor constitutive active/androstane receptor-mediated alterations in DNA methylation during phenobarbital promotion of liver tumorigenesis. *Toxicol. Sci.*, **96**, 72–82.
36. Watson,R.E. and Goodman,J.I. (2002) Epigenetics and DNA methylation come of age in toxicology. *Toxicol. Sci.*, **67**, 11–16.
37. Thomson,J.P., Lempiainen,H., Hackett,J., Nestor,C., Muller,A., Bolognani,F., Oakeley,E.J., Schubeler,D., Terranova,R., Reinhardt,D. *et al.* (2012) Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. *Genome Biol.*, **13**, R93.
38. Loeppen,S., Schneider,D., Gaunitz,F., Gebhardt,R., Kurek,R., Buchmann,A. and Schwarz,M. (2002) Overexpression of

glutamine synthetase is associated with beta-catenin-mutations in mouse liver tumors during promotion of hepatocarcinogenesis by phenobarbital. *Cancer Res.*, **62**, 5685–5688.
39. Rencurel,F., Stenhouse,A., Hawley,S.A., Friedberg,T., Hardie,D.G., Sutherland,C. and Wolf,C.R. (2005) AMP-activated protein kinase mediates phenobarbital induction of CYP2B gene expression in hepatocytes and a newly derived human hepatoma cell line. *J. Biol. Chem.*, **280**, 4367–4373.
40. Cheng,A.S., Lau,S.S., Chen,Y., Kondo,Y., Li,M.S., Feng,H., Ching,A.K., Cheung,K.F., Wong,H.K., Tong,J.H. *et al.* (2011) EZH2-mediated concordant repression of Wnt antagonists promotes beta-catenin-dependent hepatocarcinogenesis. *Cancer Res.*, **71**, 4028–4039.
41. Luk,J.M., Burchard,J., Zhang,C., Liu,A.M., Wong,K.F., Shek,F.H., Lee,N.P., Fan,S.T., Poon,R.T., Ivanovska,I. *et al.* (2011) DLK1-DIO3 genomic imprinted microRNA cluster at 14q32.2 defines a stemlike subtype of hepatocellular carcinoma associated with poor survival. *J. Biol. Chem.*, **286**, 30706–30713.
42. Monnier,J., Boissan,M., L'Helgoualc'h,A., Lacombe,M.L., Turlin,B., Zucman-Rossi,J., Theret,N., Piquet-Pellorce,C. and Samson,M. (2012) CXCR7 is up-regulated in human and murine hepatocellular carcinoma and is specifically expressed by endothelial cells. *Eur. J. Cancer*, **48**, 138–148.
43. Smyth,G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.*, **3**, Article3.
44. Bolstad,B.M., Irizarry,R.A., Astrand,M. and Speed,T.P. (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, **19**, 185–193.
45. Shen,Y., Yue,F., McCleary,D.F., Ye,Z., Edsall,L., Kuan,S., Wagner,U., Dixon,J., Lee,L., Lobanenkov,V.V. *et al.* (2012) A map of the cis-regulatory sequences in the mouse genome. *Nature*, **488**, 116–120.
46. Weber,M., Hellmann,I., Stadler,M.B., Ramos,L., Paabo,S., Rebhan,M. and Schubeler,D. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.*, **39**, 457–466.
47. Seifert,M., Cortijo,S., Colome-Tatche,M., Johannes,F., Roudier,F. and Colot,V. (2012) MeDIP-HMM: genome-wide identification of distinct DNA methylation states from high-density tiling arrays. *Bioinformatics*, **28**, 2930–2939.
48. Down,T.A., Rakyan,V.K., Turner,D.J., Flicek,P., Li,H., Kulesha,E., Graf,S., Johnson,N., Herrero,J., Tomazou,E.M. *et al.* (2008) A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat. Biotechnol.*, **26**, 779–785.
49. Nair,S.S., Coolen,M.W., Stirzaker,C., Song,J.Z., Statham,A.L., Strbenac,D., Robinson,M.D. and Clark,S.J. (2011) Comparison of methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding domain (MBD) protein capture for genome-wide DNA methylation analysis reveal CpG sequence coverage bias. *Epigenetics*, **6**, 34–44.
50. Sasaki,K. and Sonoda,Y. (2000) Histometrical and three-dimensional analyses of liver hematopoiesis in the mouse embryo. *Arch. Histol. Cytol.*, **63**, 137–146.
51. Apte,U., Zeng,G., Thompson,M.D., Muller,P., Micsenyi,A., Cieply,B., Kaestner,K.H. and Monga,S.P. (2007) beta-Catenin is critical for early postnatal liver growth. *Am. J. Physiol. Gastrointest. Liver Physiol.*, **292**, G1578–G1585.
52. Otu,H.H., Naxerova,K., Ho,K., Can,H., Nesbitt,N., Libermann,T.A. and Karp,S.J. (2007) Restoration of liver mass after injury requires proliferative and not embryonic transcriptional patterns. *J. Biol. Chem.*, **282**, 11197–11204.
53. Lempiainen,H., Couttet,P., Bolognani,F., Muller,A., Dubost,V., Luisier,R., Del Rio Espinola,A., Vitry,V., Unterberger,E., Thomson,J.P. *et al.* (2013) Identification of Dlk1-Dio3 imprinted gene cluster non-coding RNAs as novel candidate biomarkers for liver tumor promotion. *Toxicol. Sci.*, **131**, 375–386.
54. Aydinlik,H., Nguyen,T.D., Moennikes,O., Buchmann,A. and Schwarz,M. (2001) Selective pressure during tumor promotion by phenobarbital leads to clonal outgrowth of beta-catenin-mutated mouse liver tumors. *Oncogene*, **20**, 7812–7816.
55. Huang,Y., Pastor,W.A., Shen,Y., Tahiliani,M., Liu,D.R. and Rao,A. (2010) The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One*, **5**, e8888.
56. Song,C.X., Yi,C. and He,C. (2012) Mapping recently identified nucleotide variants in the genome and transcriptome. *Nat. Biotechnol.*, **30**, 1107–1116.
57. Jin,S.G., Wu,X., Li,A.X. and Pfeifer,G.P. (2011) Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic Acids Res.*, **39**, 5015–5024.
58. Nestor,C., Ottaviano,R., Reinhardt,D., Sproul,D., Dunican,D.S., Reddington,J.P., Katz,E., Dixon,M., Harrison,D.J. and Meehan,R. (2011) Tissue-type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res.*, **22**, 467–477.
59. Kriaucionis,S. and Heintz,N. (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*, **324**, 929–930.
60. Liu,L., Luo,G.Z., Yang,W., Zhao,X., Zheng,Q., Lv,Z., Li,W., Wu,H.J., Wang,L., Wang,X.J. *et al.* (2010) Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. *J. Biol. Chem.*, **285**, 19483–19490.
61. Stadtfeld,M., Apostolou,E., Akutsu,H., Fukuda,A., Follett,P., Natesan,S., Kono,T., Shioda,T. and Hochedlinger,K. (2010) Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature*, **465**, 175–181.
62. Phillips,J.M., Burgoon,L.D. and Goodman,J.I. (2009) The constitutive active/androstane receptor facilitates unique phenobarbital-induced expression changes of genes involved in key pathways in precancerous liver and liver tumors. *Toxicol. Sci.*, **110**, 319–333.
63. Phillips,J.M., Burgoon,L.D. and Goodman,J.I. (2009) Phenobarbital elicits unique, early changes in the expression of hepatic genes that affect critical pathways in tumor-prone B6C3F1 mice. *Toxicol. Sci.*, **109**, 193–205.
64. Calvisi,D.F., Ladu,S., Factor,V.M. and Thorgeirsson,S.S. (2004) Activation of beta-catenin provides proliferative and invasive advantages in c-myc/TGF-alpha hepatocarcinogenesis promoted by phenobarbital. *Carcinogenesis*, **25**, 901–908.
65. Chua,M.S., Sun,H., Cheung,S.T., Mason,V., Higgins,J., Ross,D.T., Fan,S.T. and So,S. (2007) Overexpression of NDRG1 is an indicator of poor prognosis in hepatocellular carcinoma. *Mod. Pathol.*, **20**, 76–83.
66. Yamamoto,Y., Moore,R., Goldsworthy,T.L., Negishi,M. and Maronpot,R.R. (2004) The orphan nuclear receptor constitutive active/androstane receptor is essential for liver tumor promotion by phenobarbital in mice. *Cancer Res.*, **64**, 7197–7200.
67. Holsapple,M.P., Pitot,H.C., Cohen,S.M., Boobis,A.R., Klaunig,J.E., Pastoor,T., Dellarco,V.L. and Dragan,Y.P. (2006) Mode of action in relevance of rodent liver tumors to human cancer risk. *Toxicol. Sci.*, **89**, 51–56.
68. Honkakoski,P., Zelko,I., Sueyoshi,T. and Negishi,M. (1998) The nuclear orphan receptor CAR-retinoid X receptor heterodimer activates the phenobarbital-responsive enhancer module of the CYP2B gene. *Mol. Cell. Biol.*, **18**, 5652–5658.
69. Kawamoto,T., Kakizaki,S., Yoshinari,K. and Negishi,M. (2000) Estrogen activation of the nuclear orphan receptor CAR (constitutive active receptor) in induction of the mouse Cyp2b10 gene. *Mol. Endocrinol.*, **14**, 1897–1905.
70. Konno,Y., Kamino,H., Moore,R., Lih,F., Tomer,K.B., Zeldin,D.C., Goldstein,J.A. and Negishi,M. (2010) The nuclear receptors constitutive active/androstane receptor and pregnane x receptor activate the Cyp2c55 gene in mouse liver. *Drug Metab. Dispos.*, **38**, 1177–1182.
71. Guengerich,F.P. (2008) Cytochrome p450 and chemical toxicology. *Chem. Res. Toxicol.*, **21**, 70–83.
72. Hayes,J.D., Flanagan,J.U. and Jowsey,I.R. (2005) Glutathione transferases. *Annu. Rev. Pharmacol. Toxicol.*, **45**, 51–88.
73. Xie,H.J., Yasar,U., Lundgren,S., Griskevicius,L., Terelius,Y., Hassan,M. and Rane,A. (2003) Role of polymorphic human CYP2B6 in cyclophosphamide bioactivation. *Pharmacogenomics J.*, **3**, 53–61.
74. Hoflack,J.C., Mueller,L., Fowler,S., Braendli-Baiocco,A., Flint,N., Kuhlmann,O., Singer,T. and Roth,A. (2012) Monitoring Cyp2b10

mRNA expression at cessation of 2-year carcinogenesis bioassay in mouse liver provides evidence for a carcinogenic mechanism devoid of human relevance: the dalcetrapib experience. *Toxicol. Appl. Pharmacol.*, **259**, 355–365.
75. Lahousse,S.A., Hoenerhoff,M., Collins,J., Ton,T.V., Masinde,T., Olson,D., Rebolloso,Y., Koujitani,T., Tomer,K.B., Hong,H.H. *et al.* (2011) Gene expression and mutation assessment provide clues of genetic and epigenetic mechanisms in liver tumors of oxazepam-exposed mice. *Vet. Pathol.*, **48**, 875–884.
76. Anwar,S.L., Krech,T., Hasemeier,B., Schipper,E., Schweitzer,N., Vogel,A., Kreipe,H. and Lehmann,U. (2012) Loss of imprinting and allelic switching at the DLK1-MEG3 locus in human hepatocellular carcinoma. *PLoS One*, **7**, e49462.
77. Jelinic,P. and Shaw,P. (2007) Loss of imprinting and cancer. *J. Pathol.*, **211**, 261–268.
78. Wei,W., Chua,M.S., Grepper,S. and So,S.K. (2009) Blockade of Wnt-1 signaling leads to anti-tumor effects in hepatocellular carcinoma cells. *Mol. Cancer*, **8**, 76.
79. Wei,P., Zhang,J., Egan-Hafley,M., Liang,S. and Moore,D.D. (2000) The nuclear receptor CAR mediates specific xenobiotic induction of drug metabolism. *Nature*, **407**, 920–923.
80. Pastor,W.A., Pape,U.J., Huang,Y., Henderson,H.R., Lister,R., Ko,M., McLoughlin,E.M., Brudno,Y., Mahapatra,S., Kapranov,P. *et al.* (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature*, **473**, 394–397.