DNA methylation as a genomic marker of exposure to chemical and environmental agents.

Citation for published version:

Digital Object Identifier (DOI):
10.1016/j.cbpa.2018.02.006

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published in:
Current opinion in chemical biology

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
DNA methylation as a genomic marker of exposure to chemical and environmental agents
Richard R Meehan¹, John P Thomson¹, Antonio Lentini², Colm E Nestor² and Sari Pennings³

Recent progress in interpreting comprehensive genetic and epigenetic profiles for human cellular states has contributed new insights into the developmental origins of disease, elucidated novel signalling pathways and enhanced drug discovery programs. A similar comprehensive approach to decoding the epigenetic readouts from chemical challenges in vivo would yield new paradigms for monitoring and assessing environmental exposure in model systems and humans.

Addresses
¹ MRC Human Genetics Unit, Institute of Molecular Medicine, University of Edinburgh, Crewe Road, Edinburgh EH4 2XU, UK
² Department of Clinical and Experimental Medicine, Linköping University, Linköping SE 58183, Sweden
³ Centre for Cardiovascular Science, Queen’s Medical Research Institute, University of Edinburgh, 47 Little France Crescent, EH16 4TJ, UK

Corresponding authors: Meehan, Richard R (richard.meehan@igm.m ed.ac.uk), Nestor, Colm E (colm.nestor@liu.se), Pennings, Sari (sari. pennings@ed.ac.uk)

DNA methylation dynamics in the genome
Epigenomic profiling of multiple tissues and many disease states has enhanced drug discovery programs, cancer diagnosis and analysis, population based studies and regenerative medicine. The relative ease of genomic mapping of DNA methylation was facilitated by an epigenetic mark that is highly stable and can be retrieved from archival samples [1–4]. Genomic DNA methylation profiles are tissue specific hallmarks of cell identity [5]. The modification occurs at the 5th position of cytosine to generate 5-methylcytosine (5mC), typically in the context of the dinucleotide CpG on both DNA strands. In somatic tissues >70% of CpGs are constitutively methylated [6]. The propagation and maintenance of DNA methylation patterns during development, by combined action of DNA methyltransferases (DNMTs) and DNA demethylation pathways, creates an epigenetic ‘landscape’ that supports cell type specific gene regulatory networks, regulates imprinted gene activity, represses transposon activity and enhances genome integrity [6].

DNA methylation is an epigenetic repressive mark in regulatory regions (gene promoters and enhancers), generally associated with their functional inhibition as determined by sequence context [6,7]. DNA methylation reprogramming can result from inhibition of DNMTs or de novo DNMT activity. In addition, TET enzymes contribute to DNA methylation dynamics by oxidation of 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylycysine (5fC) and 5-carboxylcytosine (5caC) as intermediates in DNA demethylation pathways. These additional DNA modifications may also act as unique epigenetic signals (Figure 1a) [8]. Although 5hmC, 5fC and 5caC are less abundant than 5mC, they contribute to a sensitive and dynamic read-out of cell state, as their profiles are in part determined by active gene-body transcription and enhancer activity, which are rapidly altered upon environmental challenge [8,9**.10]. 5mC profiles can exhibit age associated changes that may predicate more dramatic changes upon cellular transformation [11] (Figure 1b). Conversely, 5hmC profiles, in concert with other epigenomic profiles, are a highly useful tool for charting environmental exposure and integrated pathway analysis to reveal a drug’s mode of action (MoA) [12] (Figure 2).

DNA modification ‘barcoding’ in risk assessment
In an environmental exposure context, a baseline of epigenomic profiles (including transcription) for the target tissue in the organism can be used as a starting point for dose–response, temporal and transient exposure studies with model compounds (inorganic Arsenic, Vinclozolin and Triclosan) (Figure 2) [13]. It is essential to use models that can be extrapolated to humans, as responses to a drug or environmental challenge may be species specific [13]. Furthermore environmental epigenetics has to be allied with classical toxicology studies in order to be fully informative [14*,15*]

Elucidating the xenobiotic-induced physiological changes in animal models and human cohorts is important for hazard (including cancer) identification and risk assessment [15,16]. Concern over bio-reactive chemical and
The 5mC as a marker of cell state. (a) In the proposed mammalian DNA methylation pathway, canonical methylcytosine (5mC) is formed by addition of a methyl group to unmodified cytosine by DNA methyltransferase enzymes (DNMTs). Members of the TET family of di-oxigenases can then convert 5mC to 5-hydroxymethylcytosine (5hmC), using iron and α-ketoglutarate as cofactors. The DNA methylation cycle is completed by conversion of 5hmC to unmodified cytosine by (i) passive de-methylation upon DNA replication as methyltransferases do not recognise hydroxymethylated DNA, or (ii) TET-mediated oxidation of 5hmC to 5-formylcytosine (5fC) and then carboxyl- cytosine (5caC), followed by their removal through base-excision repair. (b) The development of epigenetic identity in mammalian cells reflects the establishment and activity of underlying gene regulatory networks (GRNs) that specify epigenetic transitions in dynamic and reversible systems that become more restricted as development proceeds.

Environmental agents has led to better understanding of the MoAs underlying toxic effects of chemical exposure [17]. Testing of chemicals depends on long-term rodent bioassays, which are costly and time-consuming, especially dose response studies to evaluate minimal observed adverse effect levels as points of departure (PoD) for hazard assessment. Recent progress has demonstrated that the integrated analysis of combined gene expression and epigenetic profiling of target tissues in model organisms exposed to chemical compounds can reveal underlying MoAs, and may act as predictors of atypical endpoints and adverse outcome pathways (AOPs) where cellular homeostasis is over-ridden [18,19]. The application of epigenomics profiling in concert with bioassays can further inform decisions in chemical risk assessment by identifying the critically sensitive biological pathways. As epigenomic signatures can be propagated through cell divisions, they can potentially persist even after removal of the exposure agent [20]. However, altered epigenomic profiles may also signify changes in cellular populations, altered by exposure [3]. This record of chronic epigenomic alterations induced by multiple environmental exposures may be interpreted in terms of a ‘barcode/identifier’ of risk and consequences that can be developed into alternative screening systems, such as microfluidic organs-on-a-chip technologies [20,21].

**5hmC mapping: a tool to link exposure events to disease states**

A wealth of data has recorded changes in transcriptional states following compound exposure [22]. However, due to the highly dynamic nature of transcriptional changes following chemical exposure, it is challenging to tell cause from effect in these studies. The application of dynamic 5hmC profiling can enhance the information extracted from exposure studies, especially if it can be performed on archived material and integrated with previous data profiles. 5hmC patterns are linked with gene expression status and can rapidly alter following short term (24 hours) exposure — highlighting the sensitivity of such assays [9**]. 5hmC alterations are linked to the length of drug exposure, and a number of 5hmC changes induced by the non-genotoxic carcinogen, phenobarbital (PB), were shown to persist in PB driven tumours, highlighting that such changes may be early exposure biomarkers [13]. It is also noteworthy that in age associated DNA methylation signatures, identified in humans and mice, attenuation of these changes by altered diet is associated with increased longevity [11,23]. Thus, it may be possible to decode 5hmC signatures that are unique to a given drug exposure/disease state, in which loci exhibiting reproducible 5hmC changes can be predictive of downstream consequences. 5hmC profiles in circulating cell free DNA can already aid in the identification of the tissue origin of tumours based on their abnormal DNA modification profiles [24**]. The large scale application of genome-wide sequencing approaches in these studies will lead to a greater understanding of the molecular events occurring following compound exposure and toxicity, enabling a more defined study of linkage to disease outcomes.
Potential application of epigenetic profiling to chemical challenge studies. A generic experimental design encompassing an overview of the stages relating to toxicity pathway responses, mode of action pathways (MoAs) and adverse outcome pathways (AOP) — with the examples below of inorganic arsenite and endocrine disrupting compound (EDCs: Vinclozolin and Triclosan) mediated perturbation of target organs including liver and CD4+ T-cells.
specific sequencing following bisulphite treatment or candidate locus specific PCR following epigenetic antibody enrichment strategies [15*,25*]. Progress in the assessment is summarised here and in Table 1.

Arsenic (As) contaminated drinking water at levels associated with a variety of adverse health effects and shortened lifespan are consumed by an estimated 200 million people worldwide [26]. The mechanisms underlying arsenic toxicity are not completely clarified; however, epigenetic events in concert with epithelial-to-mesenchymal transformation have been hypothesized to underpin its MoA [27]. Reduced representative DNA methylation analysis has identified differentially methylated (for 5mC and 5hmC) regions between normal and iAs-transformed cultured cells but further genome-wide efforts are required to assess in vivo arsenic-exposure risk, and dose severity, as this strongly impacts upon the observed epigenetic perturbations [26,28]. For instance, assessments of in vivo exposure by recent epigenome-wide association studies (EWAS) have identified reproducible, smoking-associated DNA methylation differences in whole blood DNA, even from short-term low-dose exposure [29–31].

The harmful effects of endocrine disrupting chemicals (EDCs) in the environment are based on their ability to act as endogenous hormone mimics [15*]. Multi-layered epigenomic analysis combined with toxicological assessments such as the EDC tests provided by the Organisation for Economic Co-operation and Development (OECD) will improve exposure monitoring and derive endpoints that are predictive/protective of all life stages in different organisms, which may have altered dependencies on epigenomic pathways [32*,33]. The mode of action of EDCs is not only through reproduction impairments [15*]; many liver specific cytochrome P450 detoxification enzymes are induced by and metabolise a broad range of steroidal and EDC compounds, leading to perturbed transcriptional and epigenetic outcomes [34]. Recent studies on the antibacterial agent, Triclosan, for instance suggest it may alter the liver methylome and act as a tumour promoter [35,36]. Comprehensive epigenomic modelling of the effect of toxic compounds and EDCs on the liver in multiple organisms can therefore underpin environmental exposure epigenetic studies; a generic experimental design is outlined in Figure 2.

Studies investigating drug induced epigenetic perturbations have focused on exposure to carcinogenic agents in rodent bioassays either by genotoxic carcinogens (GCs) or non-genotoxic carcinogens (NGCs). NGCs are able to promote carcinogenesis without inducing a direct mutation to the DNA sequence and thus can be viewed as inducers of changed epigenetic states that are propagated long after the initial exposure event [13,37]. PB is a well-established model of a non-genotoxic rodent hepatocarcinogen increasing the incidence of spontaneously and

### Table 1

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Mode of action</th>
<th>Areas of uncertainty</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heavy metals: arsenic, cadmium</td>
<td>Differentially methylated regions detected following exposure in cultured cells correlate with changes in gene expression, alternative splicing, epithelial to mesenchymal transition; differential methylation in T-cells; altered p16 expression.</td>
<td>In vivo exposure, dose severity</td>
<td>[26,27,48–52]</td>
</tr>
<tr>
<td>Airborne aerosols and cigarette smoke</td>
<td>DNA methylation differences detected in whole blood DNA after in vivo exposure locate at CpGs in or near genes including any hydrocarbon receptor (AHRR)</td>
<td>Variance in exposure levels (i.e. dose level, frequency of exposure, chemical composition differences between exposures)</td>
<td>[29–31,53,54*]</td>
</tr>
<tr>
<td>Endocrine disrupting chemicals (EDCs)</td>
<td>Mimic endogenous hormones, impairing reproduction; nuclear receptor signalling dependent activation of liver specific cytochrome P450 detoxification enzymes leads to transcriptional and epigenetic perturbation</td>
<td>Difference between organisms in their dependencies on epigenomic pathways</td>
<td>[15*,34]</td>
</tr>
<tr>
<td>Non-genotoxic-carcinogens: phenobarbital</td>
<td>Promotes nuclear translocation of the constitutive androstane receptor (CAR); changes the epigenome and transcriptome, leading to liver cancer (genetic change)</td>
<td>Diversity in MoAs and lack of genotoxicity make it hard to predict carcinogenic potential in short term tests.</td>
<td>[37,38,39**]</td>
</tr>
<tr>
<td>Germline exposure: vinclozolin, bisphenol A, or di-(2-ethylhexyl) phthalate</td>
<td>Changes in transcription and methylation in fetal cells after maternal exposure</td>
<td>Differences between studies due to genetic variability; extent of transgenerational inheritance</td>
<td>[40,41,44–46]</td>
</tr>
</tbody>
</table>
chemically induced liver tumours [38]. Previous work has shown that exposure to certain xenobiotic agents such PB results in reproducible changes in both the epigenome and transcriptome in both mouse and rat livers in a time-dependent manner — ultimately resulting in genetically defined liver tumour formation (Figure 3) [39**]. Comparative analyses of these omics landscapes provides a novel insight into the molecular events and perturbation of gene networks following both acute and chronic exposure events providing early biomarkers for the development of AOPs such as tumour formation [13].

Another area that requires deeper study is the analysis of compounds implicated in germline changes causing transgenerational outcomes [40,41]. Many groups have shown that the sperm methylome can be perturbed by environmental influences including diet; however, stochastic epigenetic variation can affect the mouse sperm methylome to a greater extent than diet and would be hard to reconcile with specific transgenerational outcomes that depend on fertilization by a single sperm [42*,43]. Vinclozolin, an anti-androgenic fungicide, is suggested to induce transgenerational phenotypes associated with perturbed epigenetic profiles [40,41]. However, independent groups were unable to replicate the phenotypic results following intraperitoneal or oral exposure routes [44–46], perhaps due to genetic variability in rat strains and other factors. In a genome-wide methylation study of prosperous-matogonia of male offspring from mothers treated with vinclozolin, bisphenol A, or di-(2-ethylhexyl)phthalate, changes in transcription and methylation in the G1 germ-line observed after EDC exposure did not persist into the G2 germline [47**]. This suggests that EDCs can exert direct epigenetic effects in exposed fetal germ cells, but these may be corrected by reprogramming events in the next generation which protect against transgenerational outcomes.

**T-cell methylation as an optimal barcode of environmental and chemical exposure in humans**

An ideal biomarker should be sensitive, specific, accessible, cause minimal discomfort to the patient and be robust to variations in laboratory handling. DNA methylation fulfils many of these criteria. However, although several chemical and environmental agents are known to affect the DNA methylation profiles of mammalian tissues, profiling *in vivo* methylation at internal sites of environmental insult is rarely possible. CD4+ T-cells patrol the entire human body, including the central nervous system, and upon exposure to their cognate antigen differentiate into short-lived effector T-cells or long-lived memory T-cells [55]. Importantly, as the chemical micro-environment of T-cells at the time of activation can affect their epigenetic profile [31,48–53,54], memory T-cells can carry epigenetic markers of both the type and location of exposure to chemical or environmental agents [57,58]. Moreover, as these cells can live for up to 40 years, our memory T-cell population carries a collective epigenetic record of our total environmental and chemical exposure over time (Figure 4).

Several studies have reported consistent alterations in T-cell methylation resulting from environmental or
chemical exposure to many of the agents reviewed above. Smoking, air pollution (polycyclic aromatic hydrocarbons, and ambient fine particles) [31,53], arsenic [48–50], pollen [54*], milk protein [56] and certain vitamins [53] have been associated with altered CD4+ T-cell methylation, at the time of exposure and afterwards [54*]. The observed changes in T-cell methylation or in T-cell population composition may result from direct effects upon the differentiation program of the T-cells or from their response to tissue injury, or both. Thus, whereas methylation profiles of total CD4+ T-cell populations may identify large changes in DNA methylation induced by exposure to chemical and non-chemical environmental agents [54*], the ability to profile thousands of single T-cell methylomes simultaneously will allow identification of individual sub-populations of memory T-cells that may have been affected by different exposures at different time points; this complex information may be codified into an epigenetic barcode of system-wide and organ-specific historical environmental and chemical exposures (Figure 4).

**Future perspectives**
The application of integrated epigenomic and transcriptomic profiling of xenobiotic exposure in animal models has enabled enhanced mechanistic interpretation and novel early biomarker discovery [12,13,59*,60]. To assess the relevance of epigenetic modifications identified in vitro for drug safety science, it is essential to investigate such changes in putative target organs in vivo. To determine the epigenetic changes that signify a toxic response, one of the future challenges will be in understanding the mechanisms by which these dynamic multi-dimensional epigenetic landscapes are regulated and altered in response to chemical exposure [14*]. It is therefore essential to establish both the dose response, particularly at low exposures, and reversibility, to aid risk assessment and dose setting (Figures 2 and 4). Another challenge will be understanding how toxicant exposure affects the composition and differentiation status of cell types in a given tissue — which may be addressed through the adaptation of new technologies such as single cell transcriptomic and epigenomic analyses. Progress in these areas will require interdisciplinary research encompassing toxicology, epigenomics data generation, bioinformatics and clinical and whole animal studies.

**Acknowledgements**
Work in RRM’s lab is supported by an LRI grant from CEFIC, the Medical Research Council (ref: MC_PC_U127574433), and previously by the MARCAR consortium. Work in the lab of CEN is funded by grants from the Swedish Research Council and the Swedish Cancer Society. Research in SP’s lab is supported by the BBSRC and BHF.
References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


This paper describes a systematic analysis of liver DNA modification changes occurring over a 91 day exposure study. 5-Hydroxymethylcytosine (5hmC) profiling was used as an indicator of cell state during organ maturation and drug-induced responses; providing novel epigenetic signatures for non-genotoxic carcinogen exposure.


Good review with a vision for toxicity testing in the 21st century: predicting that a combination of state-of-the-art in vitro assays allied with epigenomic profiling will provide new paradigms for testing and predicting MoAs for chemical hazards.


Excellent review on the history of endocrine disrupting chemicals (EDC’s) and how they may impact on epigenomic programming during development, contains excellent summary tables of diverse DNA methylation studies.


Out-standing paper that utilizes 5-hydroxymethylcytosine (5hmC) as a diagnostic identifier in circulating cell-free DNA (cfDNA) of cancer patients. They found distinct features that could be used to predict cancer types and stages with high accuracy. The authors suggest that cell-free SmC signatures may potentially be used not only to identify cancer types but also to track tumour stages. Similar methodologies may be utilised in chemical challenge studies perhaos using specific T-cell populations as the target tissue.


A systematic review of epidemiologic studies evaluating the association between environmental chemicals and DNA methylation levels in adults: incorporating measurements by microarray technologies, methylation-specific quantitative PCR, Lumimetric Methylation Assay (LUMA) and bisulfite pyrosequencing. A good preview for subsequent next generation, genome wide studies.


DNA methylation as a genomic marker Meehan et al. 55


A comprehensive report describing widespread DNA demethylation of enhancers (via 5-hydroxymethylcytosine (5hmC)) during the phyloptic period in multiple vertebrates. The study highlights a regulatory module associated with the most conserved phase of vertebrate embryogenesis that may be sensitive to perturbation by chemical exposure.


By integrative analysis of multiple data sets, this paper charts the progression from Phenobarbital (PB) exposure to end-point tumours: identifying changes in 5-hydroxymethylcytosine (5hmC) at CpG island (CI) genes that presage their hypermethylation in liver tumours. Dynamic changes in H3K27me3 promoter deposition correlated strongly with tumour-specific activation and repression of transcription. The authors suggest that loss of promoter-associated 5hmC in liver tumors licenses reprogramming of DNA methylation at silent CGI during liver cancer progression.


Intriguing paper on a potential mechanism by which paternal diet affects offspring metabolism. The authors suggest that transfer RNA (tRNA) fragments may regulate endogenous retroelements in the preimplantation embryo. Interfering with specific tRNAs by anti-sense oligos resulted in dramatic up-regulation of genes normally highly expressed in preimplantation embryos and are regulated by the long terminal repeat (LTR) of the endogenous retroelement, MERVL.


45. Gray JLE, Furr J: Vinclozolin treatment induces reproductive malformations and infertility in male rats when administered during sexual but not gonadal differentiation; however, the effects are not transmitted to the subsequent generations. Biol Reprod 2008, 78:227-228.


Interesting paper that generated a lot of debate, the authors suggest that endocrine-disrupting chemicals (vinclozolin, bisphenol A, and di-(2-ethylhexyl)phthalate) can exert direct epigenetic effects in exposed fetal germ cells, but these are corrected by reprogramming events in the next generation: thus preventing transgenerational inheritance of environmentally-caused epigenetic aberrations. DNA methylation analysis utilised the methylated CpG island recovery assay (MIRA) and hybridisation on CpG-promoter and custom-designed tile arrays.


Using Illumina’s HumanMethylation450 and HT-12 expression microarrays, the authors could clearly and robustly separate allergic patients from healthy controls. Importantly, they found that allergic patients had different proportions of T-cell sub-types compared to healthy controls. The changes in T-cell proportions underpin the difference in the DNA methylation profiles observed between patients and controls. This study represents a molecular classification of allergy using CD4+ T cells.


58. **Johnson MO, Siska PJ, Contreras DC, Rathmell JC:** Nutrients and the microenvironment to feed a T cell army. *Semin Immunol* 2016, **28**:505-513.
