CellPress
REVIEWS

Opinion

Pitfalls and Opportunities for Epigenomic Analyses Focused on Disease Diagnosis, Prognosis, and Therapy

Volker M. Lauschke,[1,†] Maxim Ivanov,[2,†] and Magnus Ingelman-Sundberg[1,*]

The Complexity of Epigenetic Analyses

Fueled by the vast amount of genomic data generated in many population-scale human sequencing projects, many current research interests and efforts focus on the role of heritable genetic factors in disease and therapy. However, non-genetic factors such as environmental stimuli, age, diseases, diet, lifestyle, and exposure to xenobiotics also contribute substantially to disease susceptibility, etiology, progression, and remission; for example, via epigenetic mechanisms [1]. Accordingly, much research is now devoted to finding epigenetic alterations as diagnostic or prognostic biomarkers to identify at-risk individuals and predict the success of drug therapy. In addition, drugs targeting the epigenetic machinery can increase the success of conventional standard-of-care treatment and are gaining momentum particularly in the area of oncology [2].

Epigenetic marks are reversible, tissue specific, and highly dynamic modifications, which requires the consideration of specific caveats regarding the planning, design, and execution of epigenetic analyses [3,4]. However, as our knowledge and methods for the study of epigenetic regulation are relatively novel, epigenetic marks are in some studies inadvertently considered as stable and tissue invariant, similar to genetic variations. Neglecting these considerations when designing an epigenetic analysis plan results in a significant risk of false conclusions. Consequently, results obtained from easily accessible surrogate samples, such as blood, require stringent validations to draw mechanistic conclusions, and great caution is advised in extrapolating results to tissues of interest. In addition, it is by now clear that epigenetic regulation is orchestrated by the interplay of a plethora of different DNA and histone modifications and is far more complex than initially thought [5–8]. There is thus a need for scientists and clinicians to be updated regarding the rapid developments in the epigenetic field as well as to be aware of the limitations and advantages of the different methods for the determination of cytosine derivates (Figure 1).

Outcomes of Epigenetic Regulation

Epigenetic regulation has been demonstrated to be of importance for fundamental biological processes such as embryonic development, cell differentiation, and transdifferentiation as well as for tissue-specific gene expression. It is also involved in many clinically relevant phenomena, including epigenetic memory of drug exposure (which in some cases can even be inherited

Trends

Many studies assess epigenetic signatures in bodily fluids and make conclusions about epigenetic patterns in inaccessible tissues of interest such as brain, thus neglecting the tissue specificity of epigenetic marks.

There is a growing lag between the forefront of epigenetics and the methods routinely applied in epigenetic studies. For instance, while techniques to detect the functionally relevant methylcytosine derivate 5-hydroxymethylcytosine were described years ago, only a few studies use such detection methods.

[1]Department of Physiology and Pharmacology, Section of Pharmacogenetics, Karolinska Institutet, SE-171 77 Stockholm, Sweden
[2]Section for Molecular Plant Biology, Department of Plant and Environmental Sciences, Copenhagen University, Thorvaldsensvej 40, DK-1871 Frederiksberg C, Denmark
[†]These authors contributed equally.

[*]Correspondence: magnus.ingelman-sundberg@ki.se (M. Ingelman-Sundberg).

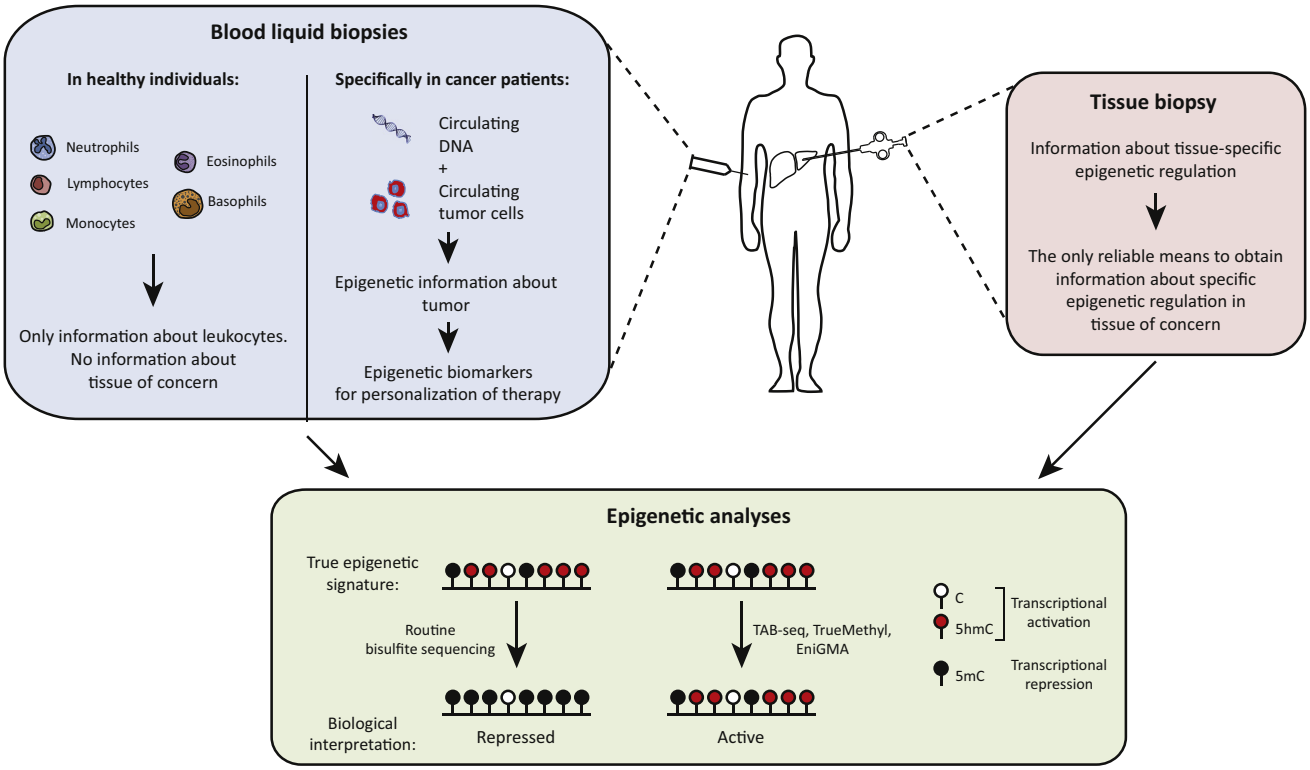

Figure 1. Important Considerations Regarding Epigenetic Biomarkers and Analysis Methods. In cancer patients, liquid biopsies from blood contain cell-free DNA and circulating tumor cells, which can inform about the epigenetic signatures in the tumor of origin. These epigenetic biomarkers can be used for patient stratification and to monitor therapy response, thus providing powerful tools that aid the personalization of treatment. By contrast, epigenetic patterns identified in blood-based liquid biopsies in healthy individuals inform only about leukocytes, and extrapolations to the epigenetic state of other tissues have to be made with great caution. Thus, while difficult to obtain, tissue biopsies provide the only means to obtain reliable data about epigenetic regulation in non-diseased tissues. In addition to these considerations regarding sampling, methods should be chosen for epigenetic analyses that allow the distinction between silencing 5-methylcytosine (5mC) and activating 5-hydroxymethylcytosine (5hmC) marks. This latter issue is especially important in 5hmC-rich tissues such as liver and brain, to arrive at correct biological interpretations of the obtained epigenetic data. Abbreviations: TAB, Tet-assisted bisulfite; EniGMA, enzyme-assisted identification of genome modification assay; C, cytosine.

transgenerationally) and the acquisition of drug resistance. Multiple convincing studies described the importance of epigenetic regulation during organ development and transdifferentiation by demonstrating, for example, the dynamic modulation of specific transcriptional programs by 5-hydroxymethylcytosine (5hmC) signatures during heart development and the remodeling of methylation signatures during stellate cell transdifferentiation into myofibroblasts in liver fibrosis [9,10]. Epigenetic factors also result in parent-specific imprinting; the most prominent examples include *IGF2* and the long noncoding RNA H19, which are exclusively transcribed from the paternal or maternal allele, respectively [11]. In the liver, the expression of more than 60 genes with importance for drug absorption, distribution, metabolism, and excretion (ADME) has been shown to be modulated by epigenetic mechanisms, with important implications for liver function and drug response [12]. Furthermore, environment- or drug-induced epigenetic alterations can have persistent effects, as shown, for example, by transient induction of the nuclear receptor CAR by its ligand TCPOBOP in neonatal mice, which induces elevated expression of the CAR-target genes *Cyp2b10* and *Cyp2c37* that persists throughout the adult life of the exposed animals and reduces their sensitivity to zoxazolamine [13]. This 'memory effect', mediated by stable alterations of histone methylation patterns, can be explained by reinforcement of the current transcriptional state of the gene with activating or repressing epigenetic signatures. In addition, during prolonged drug treatment resistance can

occur, for example, by specific induction of enzymes or transporters, events that can be counteracted by the use of so-called epidrugs that intervene in the epigenetic cellular machinery [14,15].

Technical Considerations for Epigenomic Profiling

The toolbox for epigenetic analyses has been drastically expanded in recent years [16]. Critically, however, bisulfite sequencing (BS), the conventional technique used to decode the epigenetic modification state of DNA, can only binarily distinguish between demethylation states, flagging both 5-methylcytosine (5mC) and 5hmC as methylation marks and other demethylation intermediates, such as 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), as demethylated residues [17]. Hence, BS results in a composite signal comprising a mixture of inactivating mC marks and hmC modifications that are typically associated with actively transcribed genes [18–20]. This inability to resolve mC and hmC marks is especially important in hmC-rich tissues such as liver and brain, in which up to 0.5–1% of all cytosines (i.e., 10–20% of cytosines in the CpG context) can be hydroxymethylated [7,8].

To overcome this limitation, a variety of techniques has been developed. However, results obtained with different methods are not always concordant, especially when protocols are used that rely on antibody-based quantifications, [21]. Such approaches rely on high antibody quality and the meticulous standardization of protocols to be comparable across experiments and laboratories. Therefore, the current generation of protocols is based on enzymatic protection [Tet-assisted bisulfite (TAB)-based platforms] of 5hmC or chemical conversion [oxidative bisulfite (oxBS) methods; also known as TrueMethyl] and both can be combined with downstream sequencing or array-based methods to allow the separation of 5mC and 5hmC signals with single-base resolution [22,23]. Importantly, however, extensive cross-validation of the two methodologies revealed that oxBS-based platforms exhibited increased technical variability compared with TAB-based methods and in their current state are prone to underestimating absolute 5hmC values at a fraction of CpG sites, probably due to incomplete oxidation [24]. A novel enzyme-assisted identification of genome modification assay (EniGMA) based on DNMT1 and hairpin BS was recently presented that revealed, for example, notable differences in the genomic localization of most hydroxymethylated loci between different brain regions and between individuals [25]. However, while this technique appears promising for the interrogation of candidate genes of interest, it is currently not suitable for whole-genome analyses. In addition, various LC/MS-based and nano-fluorescent techniques are being developed that may hold promise for the future development of the field [26,27]. In conclusion, despite important progress there remains a great need for novel, adaptable, high-resolution techniques compatible with global profiling for the true determination of 5mC and 5hmC in tissues, cells, and circulating DNA fragments.

Important Considerations for Epigenetic Sampling and Analyses

Epigenetic profiling of clinical samples is a challenging and technically complicated task that critically relies on well-defined experimental setups. Epigenetic modifications are quantitative and reversible in nature, whereas genetic variations are either present or not in a given individual. Thus, much higher sequencing coverage is required in NGS applications to quantify epigenetic signatures compared with regular genotyping [28,29]. In addition, the existing epigenetic profiling methodologies are often more complex than techniques for genetic analyses. For example, histone modification profiles are studied by ChIP followed by targeted or genome-wide analyses of enriched DNA (ChIP-chip, ChIP-seq), which have their own experimental and computational caveats and pitfalls [30–32]. In particular, ChIP results are highly sensitive to the quality of the input sample and the antibody used and the choice of analysis methods has to be tailored to the experimental conditions and biological question.

Most importantly, it has to be emphasized that an individual has only one genome (when not considering somatic mutations in, e.g., cancer) but myriad different epigenomes, which are driven by and reinforce tissue-specific differences in gene expression [33,34]. Moreover, epigenetic signatures differ between cells even within the same tissue with an epivariation frequency exceeding the genetic variation frequency in liver by more than 100-fold [35]. Hence, the analyzed material has to be derived from the relevant tissue and not just from the most easily available source such as peripheral blood, as demonstrated by the lack of correlation between, for example, epigenetic alterations in blood and those in liver or different parts of the central nervous system like the prefrontal cortex, entorhinal cortex, superior temporal gyrus, and cerebellum [36–38]. For instance, based on a comparison between four different tissues originating from the same individuals ($n = 75$), it was concluded that only up to 4% of the 3238 significant methylation–expression associations found in one tissue were also detectable with an identical effect direction in another tissue [37]. However, a multitude of studies have measured and correlated epigenetic signatures in bodily fluids with multifactorial and highly complex end points such as smoking behavior [39], childhood asthma due to traffic-related air pollution [40], socioeconomic status [41], and panic disorder [42]. Importantly, without mechanistic links and stringent validations these correlations are hard to interpret biologically and warrant high caution if attempts are made to translate the findings into the clinic as biomarkers or therapeutic targets.

Pharmacoepigenetic Developments towards the Clinic

Epigenetic aberrations have been implicated in a range of human diseases, including autoimmune disorders, neurodegenerative diseases, and, most importantly, cancer. Epigenetic profiling provides a powerful means to predict the origin of cancer of unknown primary, allowing more targeted therapy with significant benefits for patient survival [43]. Furthermore, targeted analyses of the epigenetic profile of specific candidate genes can indicate therapeutic targets as well as diagnostic or prognostic biomarkers [44,45].

For monitoring treatment efficacy and development of drug resistance, liquid biopsy samples are analyzed for the methylation status of various circulating DNA fragments presumably originating from the tumor. However, for such applications one has to be aware that DNA from the target tissue, such as a tumor, represents only a fraction of all cell-free DNA in human serum and specific signals of interest might be heavily diluted, depending on comorbidities and tumor stage. Importantly, however, recent methodological developments in DNA haplotype mapping and nucleosomal footprinting hold promise in providing information about the tissue of origin of identified DNA fragments in bodily fluids and thus potentially allowing conclusions about epigenetic signatures in tissues of interest from liquid biopsy data [46–48]. Furthermore, while 5hmC levels regularly decrease in tumors compared with normal tissues [49,50], the relative distribution of 5mC and 5hmC changes in different genomic regions during tumor development. However, 5mC and 5hmC levels are not resolved in the vast majority of studies that aim to identify circulating disease biomarkers. Thus, knowledge about to what extent the relative 5mC/5hmC distribution is predictive of tumor phenotype or the extent of, for example, drug resistance is in many cases incomplete.

Epigenetic mechanisms are further of importance for the acquisition of drug resistance, particularly in cancer therapy. During chemotherapy, demethylation and concomitant elevated expression of major drug export transporters such as ABCB1 and ABCG2 can occur within minutes [51–53]. Furthermore, reduced cellular uptake of oxaliplatin by epigenetic silencing of the transporter gene *SLC22A2* encoding the responsible transporter OCT2 contributes to oxaliplatin resistance [54]. Importantly, mounting preclinical evidence has demonstrated that epigenetically acquired drug resistance is reversible [14,54,55]. Furthermore, convincing clinical data have already resulted in the regulatory approval of six drugs targeting the

epigenetic machinery (azacitidine, decitabine, belinostat, panobinostat, romidepsin, and vorinostat) and many clinical trials are currently investigating the effects of including epidrugs as (re) sensitizers in the standard-of-care treatment regimen with promising results [56,57]. Overall, however, reliable conclusions about biological mechanisms and circulating biomarkers for drug resistance cannot be made without reliable resolution of DNA methylation and hydroxymethylation.

Concluding Remarks

Much effort must be made to refine and develop the methods for epigenomic analyses to identify actionable biomarkers and drug targets for clinical use (see Outstanding Questions). Currently, severe problems are evident in sampling and epigenetic analysis methodologies. Thus, there is a need for cost-effective and high-throughput-compatible novel epigenetic technologies. Furthermore, extending knowledge about the pitfalls and opportunities of epigenomic analyses among scientists is necessary to facilitate the discovery of novel biomarkers for the efficacy of disease treatment.

Outstanding Questions

Can novel epigenetic footprinting methods facilitate identification of the tissue of origin of circulating DNA fragments and increase the specificity of epigenetic liquid biopsy data in clinical practice?

Can such biomarkers contribute to disease diagnosis or the personalization of drug treatment? Will the deconvolution of mC and hmC markedly improve our understanding of epigenetic gene regulation and establish novel, high-accuracy biomarkers?

Will epigenetic therapy be established in routine treatment regimens for cancer and other diseases?

References

1. Lauschke, V.M. and Ingelman-Sundberg, M. (2016) The importance of patient-specific factors for hepatic drug response and toxicity. *Int. J. Mol. Sci.* 17, E1714
2. Lauschke, V.M. *et al.* (2017) Pharmacoepigenetics and toxicoepigenetics: novel mechanistic insights and therapeutic opportunities. *Annu. Rev. Pharmacol. Toxicol.* (in press)
3. Michels, K.B. *et al.* (2013) Recommendations for the design and analysis of epigenome-wide association studies. *Nat. Methods* 10, 949–955
4. Birney, E. *et al.* (2016) Epigenome-wide association studies and the interpretation of disease -omics. *PLoS Genet.* 12, e1006105
5. Tahiliani, M. *et al.* (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930–935
6. Kriaucionis, S. and Heintz, N. (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 324, 929–930
7. Ivanov, M. *et al.* (2013) Ontogeny, distribution and potential roles of 5-hydroxymethylcytosine in human liver function. *Genome Biol.* 14, R83
8. Globisch, D. *et al.* (2010) Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One* 5, e15367
9. Kunderfranco, P. *et al.* (2016) DNA hydroxymethylation controls cardiomyocyte gene expression in development and hypertrophy. *Nat. Commun.* 7, 12418
10. Page, A. *et al.* (2016) Hepatic stellate cell transdifferentiation involves genome-wide remodeling of the DNA methylation landscape. *J. Hepatol.* 64, 661–673
11. Barlow, D.P. and Bartolomei, M.S. (2014) Genomic imprinting in mammals. *Cold Spring Harb. Perspect. Biol.* 6, a018382
12. Fisel, P. *et al.* (2016) DNA methylation of ADME genes. *Clin. Pharmacol. Ther.* 99, 512–527
13. Chen, W.-D. *et al.* (2012) Neonatal activation of the nuclear receptor CAR results in epigenetic memory and permanent change of drug metabolism in mouse liver. *Hepatology* 56, 1499–1508
14. El-Khoury, V. *et al.* (2007) The histone deacetylase inhibitor trichostatin A downregulates human *MDR1* (*ABCB1*) gene expression by a transcription-dependent mechanism in a drug-resistant small cell lung carcinoma cell line model. *Br. J. Cancer* 97, 562–573
15. Sharma, S.V. *et al.* (2010) A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* 141, 69–80
16. Yong, W.-S. *et al.* (2016) Profiling genome-wide DNA methylation. *Epigenetics Chromatin* 9, 26
17. Huang, Y. *et al.* (2010) The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One* 5, e8888
18. Xu, Y. *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
19. Stroud, H. *et al.* (2011) 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol.* 12, R54
20. Wen, L. *et al.* (2014) Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. *Genome Biol.* 15, R49
21. Matarese, F. *et al.* (2011) 5-Hydroxymethylcytosine: a new kid on the epigenetic block? *Mol. Syst. Biol.* 7, 562–562
22. Yu, M. *et al.* (2012) Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. *Nat. Protoc.* 7, 2159–2170
23. Booth, M.J. *et al.* (2012) Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science* 336, 934–937
24. Ivanov, M. *et al.* (2016) Single base resolution analysis of 5-hydroxymethylcytosine in 188 human genes: implications for hepatic gene expression. *Nucleic Acids Res.* 44, 6756–6769
25. Kawasaki, Y. *et al.* (2017) A novel method for the simultaneous identification of methylcytosine and hydroxymethylcytosine at a single base resolution. *Nucleic Acids Res.* 45, e24
26. Amouroux, R. *et al.* (2016) *De novo* DNA methylation drives 5hmC accumulation in mouse zygotes. *Nat. Cell Biol.* 18, 225–233
27. Chen, H.-Y. *et al.* (2017) Spectroscopic quantification of 5-hydroxymethylcytosine in genomic DNA using boric acid-functionalized nano-microsphere fluorescent probes. *Biosens. Bioelectron.* 91, 328–333
28. Sims, D. *et al.* (2014) Sequencing depth and coverage: key considerations in genomic analyses. *Nat. Rev. Genet.* 15, 121–132
29. Ziller, M.J. *et al.* (2015) Coverage recommendations for methylation analysis by whole-genome bisulfite sequencing. *Nat. Methods* 12, 230–232 1 p. following 232
30. Park, P.J. (2009) ChIP-seq: advantages and challenges of a maturing technology. *Nat. Rev. Genet.* 10, 669–680
31. Kim, M.-S. *et al.* (2016) Common errors in mass spectrometry-based analysis of post-translational modifications. *Proteomics* 16, 700–714
32. Nakato, R. and Shirahige, K. (2017) Recent advances in ChIP-seq analysis: from quality management to whole-genome annotation. *Brief. Bioinform.* 18, 279–290
33. Sadikovic, B. *et al.* (2008) Cause and consequences of genetic and epigenetic alterations in human cancer. *Curr. Genomics* 9, 394–408
34. Yu, H. *et al.* (2008) Inferring causal relationships among different histone modifications and gene expression. *Genome Res.* 18, 1314–1324

35. Gravina, S. *et al.* (2016) Single-cell genome-wide bisulfite sequencing uncovers extensive heterogeneity in the mouse liver methylome. *Genome Biol.* 17, 150
36. Hannon, E. *et al.* (2015) Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics* 10, 1024–1032
37. Bonder, M. *et al.* (2014) Genetic and epigenetic regulation of gene expression in fetal and adult human livers. *BMC Genomics* 15, 860–813
38. Lunnon, K. *et al.* (2016) Variation in 5-hydroxymethylcytosine across human cortex and cerebellum. *Genome Biol.* 17, 27
39. Tiili, E.M. *et al.* (2015) Effect of genotype and methylation of CYP2D6 on smoking behaviour. *Pharmacogenet. Genomics* 25, 531–540
40. Somineni, H.K. *et al.* (2016) Ten-eleven translocation 1 (TET1) methylation is associated with childhood asthma and traffic-related air pollution. *J. Allergy Clin. Immunol.* 137, 797–805.e5
41. Swartz, J.R. *et al.* (2017) An epigenetic mechanism links socio-economic status to changes in depression-related brain function in high-risk adolescents. *Mol. Psychiatry* 22 (2), 209–214
42. Shimada-Sugimoto, M. *et al.* (2017) Epigenome-wide association study of DNA methylation in panic disorder. *Clin. Epigenetics* 9, 6
43. Moran, S. *et al.* (2016) Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. *Lancet Oncol.* 17, 1386–1395
44. Bram, E.E. *et al.* (2009) Chemotherapeutic drug-induced ABCG2 promoter demethylation as a novel mechanism of acquired multidrug resistance. *Neoplasia* 11, 1359–1370
45. Fisel, P. *et al.* (2013) DNA methylation of the SLC16A3 promoter regulates expression of the human lactate transporter MCT4 in renal cancer with consequences for clinical outcome. *Clin. Cancer Res.* 19, 5170–5181
46. Snyder, M.W. *et al.* (2016) Cell-free DNA comprises an *in vivo* nucleosome footprint that informs its tissues-of-origin. *Cell* 164, 57–68
47. Lehmann-Werman, R. *et al.* (2016) Identification of tissue-specific cell death using methylation patterns of circulating DNA. *Proc. Natl. Acad. Sci. U. S. A.* 113, E1826–E1834
48. Guo, S. *et al.* (2017) Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA. *Nat. Genet.* 49, 635–642
49. Lian, C.G. *et al.* (2012) Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell* 150, 1135–1146
50. Yang, H. *et al.* (2012) Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation. *Oncogene* 32, 663–669
51. Abolhoda, A. *et al.* (1999) Rapid activation of *MDR1* gene expression in human metastatic sarcoma after *in vivo* exposure to doxorubicin. *Clin. Cancer Res.* 5, 3352–3356
52. Tada, Y. *et al.* (2000) *MDR1* gene overexpression and altered degree of methylation at the promoter region in bladder cancer during chemotherapeutic treatment. *Clin. Cancer Res.* 6, 4618–4627
53. Bram, E.E. *et al.* (2009) Chemotherapeutic drug-induced ABCG2 promoter demethylation as a novel mechanism of acquired multidrug resistance. *Neoplasia* 11, 1359–1370
54. Liu, Y. *et al.* (2016) Epigenetic activation of the drug transporter OCT2 sensitizes renal cell carcinoma to oxaliplatin. *Sci. Transl. Med.* 8, 348ra97
55. Knoechel, B. *et al.* (2014) An epigenetic mechanism of resistance to targeted therapy in T cell acute lymphoblastic leukemia. *Nat. Genet.* 46, 364–370
56. Coronel, J. *et al.* (2010) A double-blind, placebo-controlled, randomized Phase III trial of chemotherapy plus epigenetic therapy with hydralazine valproate for advanced cervical cancer. Preliminary results. *Med. Oncol.* 28 (Suppl. 1), S540–S546
57. San-Miguel, J.F. *et al.* (2014) Panobinostat plus bortezomib and dexamethasone versus placebo plus bortezomib and dexamethasone in patients with relapsed or relapsed and refractory multiple myeloma: a multicentre, randomised, double-blind Phase 3 trial. *Lancet Oncol.* 15, 1195–1206